Mastering Public Health

A postgraduate guide to examinations and revalidation

Geraint H Lewis
Jessica Sheringham
Kanwal Kalim
Tim JB Crayford

The Royal Society of Medicine Press Limited

Mastering Public Health

A postgraduate guide to examinations and revalidation

Mastering Public Health

A postgraduate guide to examinations and revalidation

Geraint H Lewis MA MB BChir MSc MRCP MFPH
Specialist Registrar in Public Health, London Deanery

Jessica Sheringham MA MSc MFPH
Specialist Trainee in Public Health, London Deanery

Kanwal Kalim MB ChB MSc
Specialist Registrar in Public Health, London Deanery

Tim JB Crayford MB BS MSc FFPH
Director of Public Health, Croydon PCT;
President, UK Association of Directors of Public Health

The Royal Society of Medicine Press Limited

Published by the Royal Society of Medicine Press Ltd
1 Wimpole Street, London W1G 0AE, UK
Tel: +44 (0)20 7290 2921
Fax: +44 (0)20 7290 2929
Email: publishing@rsm.ac.uk
Website: www.rsmpress.co.uk

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN-13 978-1-85315-781-3

Distribution in Europe and Rest of World:

Marston Book Services Ltd
PO Box 269
Abingdon
Oxon OX14 4YN, UK
Tel: +44 (0)1235 465500
Fax: +44 (0)1235 465555
Email: direct.order@marston.co.uk

Distribution in the USA and Canada:

Royal Society of Medicine Press Ltd
c/o BookMasters Inc
30 Amberwood Parkway
Ashland, OH 44805, USA
Tel: +1 800 247 6553/+1 800 266 5564
Fax: +1 419 281 6883
Email: order@bookmasters.com

Distribution in Australia and New Zealand:

Elsevier Australia
30–52 Smidmore Street
Marrickville NSW 2204, Australia
Tel: +61 2 9517 8999
Fax: +61 2 9517 2249
Email: service@elsevier.com.au

Typeset by Phoenix Photosetting, Chatham, Kent
A Eurasia production in China

CONTENTS

FOREWORD

'Health for All' is much more than just a slogan. It is a reminder of the enormous scope of public health. Its practitioners must be able to get to grips with the explosion of knowledge on the bewildering array of health determinants, drawing on insights from, among others, genetics, biology, economics and political science. As if this were not enough, they must be experts in the critical appraisal of evidence, so that they can judge what is likely to work in a given set of circumstances and what not, and must then deploy the managerial and organisational skills to engage with a wide range of actors so as to turn their vision into reality. The scale of the canvas on which they must work is so great, and the environment in which they work so dynamic, that the acquisition of the skills and knowledge that they require can be achieved only through a process of life-long learning. It is not possible for anyone to be an expert in every aspect of public health; some degree of specialisation is required but this must be built on a common foundation of knowledge and skills that will equip the next generation of public health specialists to set out on this journey.

In the United Kingdom this foundation has been defined by the curriculum for the examinations of the Faculty of Public Health. Recently revised, it comes in two parts. The first tests candidates' understanding of the scientific principles of public health; the second, the Objective Structured Public Health Examination, tests their ability to apply this understanding.

Whatever one's disciplinary background, there is an enormous amount of new information for the candidate to absorb. Public health requires an understanding of quantitative and qualitative methods, and of natural and social sciences, and high levels of numeracy and literacy. There are, of course, many books that those training in public health can look to for guidance. However, it has long been apparent that there is a need for a concise single volume in which the candidate can find the key elements of knowledge that he or she will need to pass the exams. The authors of this book are to be congratulated for meeting this need. They have assembled an excellent overview of the essential knowledge required by tomorrow's public health practitioners in a clear and highly readable manner. I am confident that this will rapidly become required reading for all those taking the Faculty's exams, as well as for those undertaking training in public health in many other countries. Of course knowledge is not enough: the real challenge is to develop the skills to apply it. But that only comes with practice.

Martin McKee CBE MD DSc MSc FRCP FRCPI FFPH FMedSci
Professor of European Public Health
London School of Hygiene and Tropical Medicine

PREFACE

This is the type of book that we wish had existed when we were revising for the membership examination of the Faculty of Public Health.

It began life as a mass of revision notes based on dozens of books and lectures. Over the ensuing 18 months it has developed, thanks to the contributions of a wide range of people including our colleagues and our international editors, into this text. It is a revision guide that we hope will be of use to candidates of postgraduate public health examinations across many countries. However, it will also be of use for medical students and for people who are about to embark on a course of study such as an MPH or MSc in public health, for continual professional development and as a quick, practical reference text for practitioners of the specialty.

For better or worse, we decided that the book should follow strictly the structure of the UK Faculty's Part A examination syllabus (which is also used by the Australian Faculty of Public Health). This consists of five areas of public health knowledge plus a sixth section relating to public health skills. In order to avoid excessive duplication, we have at times cross-referenced the reader to other parts of the book where the syllabus covers the same material. Occasionally (as in Section 6 – public health skills), the overlapping content has been kept in order to reflect different emphases drawn out by different parts of the syllabus.

Our aim throughout the book has been to strike a balance somewhere between brief lecture notes and a full-prose textbook. We hope that this compromise will be suited both to revision and to quick reference needs. Being a postgraduate revision book, we have assumed a certain degree of prior knowledge, and have prioritised breadth of content over depth. The appendices offer the candidate of the Part A examination some revision strategies, essay frameworks and lists of key public health names and concepts that are specifically designed to support exam preparation. Text given in italics under headings indicate the exact syllabus item.

We acknowledge the assistance and guidance of our international editors, who generously offered their time and expertise to provide us with international perspectives and comments on our draft chapters. We have included the examples that they provided (indicated by national symbols) but recognise that the book remains Anglo-centric. For this reason we would be delighted to receive further international details from readers for any future edition of the book. Particular thanks go to Drs David Pencheon and Paul Crook for their thoughtful comments.

We and the publishers have attempted to seek permission from, and to acknowledge, all those whose work we used in compiling this revision guide. If, however, if we have unwittingly omitted any acknowledgements then do please contact the RSM Press.

GH Lewis
J Sheringham
K Kalim
TJB Crayford

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contributions of the international editorial panel:

AUSTRALIA: Professor Kerry Kirke, Clinical Associate Professor, University of Adelaide, The Australasian Faculty of Public Health Medicine, Royal Australasian College of Physicians

ENGLAND: Professor David Strachan, Chair, MFPH Part A examiners, Professor of Epidemiology, Division of Community Health Sciences, St George's Hospital, University of London

HONG KONG: Professor Sian Griffiths, Professor of Public Health, Director, School of Public Health, Faculty of Medicine, Chairman, Department of Community and Family Medicine, The Chinese University of Hong Kong

HONG KONG: Professor Emily YY Chan, Assistant Professor, School of Public Health, Faculty of Medicine, The Chinese University of Hong Kong

IRELAND: Dr Emer Shelley, Population Health Directorate Health Services Executive, Stewarts Sports Centre, Dublin

NEW ZEALAND: Dr Michael Baker, Senior Lecturer, Department of Public Health, Wellington School of Medicine and Health Sciences

NORTHERN IRELAND: Dr John WG Yarnell, Reader in Cardiovascular Epidemiology, Epidemiology and Public Health, Queen's University Belfast

SCOTLAND: Dr Anne Maree Wallace, Scottish Director of Training in Public Health

SOUTH AFRICA: Dr Stephen Knight, Lecturer/Public Health Medicine Specialist, School of Family and Public Health Medicine, Nelson R Mandela School of Medicine, University of KwaZulu-Natal

WALES: Dr Nigel Monaghan, Consultant in Public Health NPHS Wales/Public Health Director, Rhondda Cynon Taff Local Health Board, National Public Health Service, Cardiff

In the preparations for their Part A exams and in writing this book, the authors are indebted to the teaching provided by:

London School of Hygiene and Tropical Medicine, Masters in Public Health, 2004–06, in particular:

Judy Green, Kara Hansen, Helen Hogan, Valerie Iles, Stephen Jan, Karen Lock, Martin McKee, Barbara McPake, Ellen Nolte, Justin Parkhurst, Carine Ronsmans, Colin Sanderson and Rosalind Raine

Edmund Jessop, Part A Examination Course and Manual, Winchester, December 2005

North Central London Strategic Health Authority, Part A Revision Course, September–December 2005, in particular:

Marion Deacon, Meic Goodyear, Andy Hall and Oliver Morgan

Gwyn Bevan, Public Policy Course, Commission for Health Improvement, 2003–04

We would also like to thank the following for their valuable advice and detailed comments on sections of the book:

Jamila Aboulhab, Abdul Azad, Glen Brice, Araceli Busby, Helen Crabbe, Daragh Fahey, Jennie Mussard, Geoff Ridgway, Abdul Roudsari, Murad Ruf, South West London Health Intelligence Team, Louisa Shillito and Chris Weston

ACKNOWLEDGEMENTS BY CHAPTER

1A

Ellen Nolte and Martin McKee, London School of Hygiene and Tropical Medicine, Issues in Public Health Course, 2004–05

Carine Ronsmans and colleagues, London School of Hygiene and Tropical Medicine, Basic Epidemiology Course, 2004–05

Colin Sanderson, London School of Hygiene and Tropical Medicine, Healthcare Evaluation Course, 2004–05

Justin Parkhurst, London School of Hygiene and Tropical Medicine, Health Power, Process and Policy Course, 2004–05

Edmund Jessop, Part A Examination Course, Winchester, December 2005

Andy Hall, North Central London Part A Examination Course, September 2005

1B

Jamila Aboulhab, Specialist Registrar in Public Health

Murad Ruf, Specialist Registrar in Public Health

Edmund Jessop, Part A Examination Course and lecture notes, Winchester, December 2005

1C

Karen Lock, London School of Hygiene and Tropical Medicine, Issues in Public Health Unit, 2004–05

Colin Sanderson, London School of Hygiene and Tropical Medicine, Public Health: Integrating Unit, 2004–05

London School of Hygiene and Tropical Medicine, Economic Evaluation Unit, 2004–05

London School of Hygiene and Tropical Medicine, Introduction to Health Economics Unit, 2004–05

Ellen Nolte, London School of Hygiene and Tropical Medicine, Health Services Linear Unit, 2004–05

Andy Hall, North Central London Part A Examination Course, September 2005

Jamila Aboulhab, Specialist Registrar in Public Health, London Training Scheme

1D

London School of Hygiene and Tropical Medicine, Introduction to Social Research Unit, 2004–05

2A

Ellen Nolte, London School of Hygiene and Tropical Medicine, Issues in Public Health Course, Lecture 2, 2004–05

2G

Geoff Ridgway, Microbiologist, Senior Medical Advisor, Health Protection Division, DH

Bharat Patel, Health Protection Agency, North Central London Strategic Health Authority Part 1 revision slides

Andy Hall, London School of Hygiene and Tropical Medicine, North Central London Strategic Health Authority Part 1 revision

2H

Abdul Azad, Health Improvement and Sure Start Co-ordinator, Redbridge PCT

Carine Ronsmans and colleagues, London School of Hygiene and Tropical Medicine, Basic Epidemiology Course 2004–05

2I

London School of Hygiene and Tropical Medicine Integrating Unit in Public Health notes, 2005

3B

Meic Goodyear, North Central London course notes, December 2005

3C

Jennie Mussard, Health Intelligence Lead, Croydon PCT

South West London Health Intelligence Team

Abdul Roudsari, Head of Centre for Health Informatics, School of Informatics, City University

4A

Judy Green, London School of Hygiene and Tropical Medicine lecture notes, 2004–05

4C

Rosalind Raine, Healthcare Evaluation lecture notes, London School of Hygiene and Tropical Medicine, 2005

Stephen Jan, Economic Analysis for Management and Policy: Equity and Priority Setting lecture notes, London School of Hygiene and Tropical Medicine, 2005

Gwyn Bevan, Public Policy: Priority Setting in Oregon lecture notes, January 2004

Oliver Morgan, North Central London Part A Revision Course, 2005

4D

Daragh Fahey, North Central London Part 1 economics revision notes

5A

Valerie Iles, London School of Hygiene and Tropical Medicine, Organisational Management, 2005

5B

Oliver Morgan, North Central London Part A Revision Course, Organisational Management lecture notes, 2005

5C

Oliver Morgan, North Central London Part A Revision Course, Organisational Management presentation, 2005

6A

Edmund Jessop, Part A Examination Course, Winchester, December 2005

6C

North Central London Part A Revision Course

ABBREVIATIONS

ADL	activities of daily living
AFP	α-fetoprotein
ANOVA	analysis of variance
BMI	body mass index
COPD	chronic obstructive pulmonary disease
COSHH	Control of Substances Hazardous to Health
DALY	disability-adjusted life-year
DDD	defined daily dose
DEFRA	UK Department for Environment, Farming and Rural Affairs
DFID	UK Department for International Development
DH	Department of Health
DNA	deoxyribonucleic acid, did not attend
DRG	diagnosis-related group
DVT	deep vein thrombosis
DWI	Drinking Water Inspectorate (England & Wales)
EBM	evidence-based medicine
EIA	environmental impact assessment
FEV_1	forced expiratory volume over 1 second
FRR	familial relative risk
GDP	gross domestic product
GHS	General Household Survey
GHQ	general health questionnaire
GMC	General Medical Council
GUM	genitourinary medicine
HALE	health-adjusted life expectancy
HES	hospital episode statistics
HIA	health impact assessment
HNA	health needs assessment
HPA	UK Health Protection Agency
HPV	human papillomavirus
HRG	healthcare resource group
HRQoL	health-related quality of life
ICD	*International Classification of Diseases*

ICP	integrated care pathway
LAA	local area agreement
LR	likelihood ratio
LSP	local strategic partnership
MANOVA	multivariate analysis of variance
MHRA	Medicines and Healthcare products Regulatory Authority
MRSA	meticillin-resistant *Staphylococcus aureus*
NCEPOD	National Confidential Enquiry into Patient Outcome and Death
NICE	National Institute for Health and Clinical Excellence
NMC	Nursing and Midwifery Council
NNT	number needed to treat
NSC	UK National Screening Committee
NSSEC	National Statistics Socio-economic Classification
ONS	Office for National Statistics
PACT	prescribing analysis and cost
PALS	Patient Advice and Liaison Service
PCR	polymerase chain reaction
PCT	primary care trust
PDP	personal development plan
PEM	protein–energy malnutrition
PSA	prostate-specific antigen
QALY	quality-adjusted life-year
QMAS	quality management and analysis system
QOF	quality and outcomes framework
RCT	randomised controlled trial
RDA	recommended daily amount
ROC	receiver operating characteristic
RRR	relative risk ratio
SARS	severe acute respiratory syndrome
SD	standard deviation
SE	standard error
SIDS	sudden infant death syndrome
SMR	standardised mortality ratio
WARNER	weekly analysis report of notifications above expected rates
WHO	World Health Organization
YLL	years of life lost

Section 1
RESEARCH METHODS

Public health practitioners need to understand how health knowledge is generated not only to be able to select and use appropriate research methods for their own work, but also so that they can appraise the quality of published research and provide credible professional advice.

Section 1 outlines the research methods that underpin public health. Epidemiology is the key to public health practice: the discipline involves scrutinising data to generate meaningful inferences that can be used as the basis of health policy. The validity of public health research relies on the appropriate use of statistics, from the design and instigation of data collection, through to their analysis and interpretation.

Qualitative research methods are necessary for understanding the reasons why things happen the way that they do.

Section 1 ends with the field of health-care assessment that draws on all of these disciplines in order to evaluate the structure, function and performance of health systems and services.

1A
Epidemiology

Epidemiology is the study of the distribution and determinants of health and disease in populations. It is a collection of techniques for studying the characteristics of many different people. Epidemiology is the science of the '*who*', the '*what*', the '*where*' and the '*when*', and it is the tool for exploring the underlying questions of '*how*' and '*why*' as they relate to health.

It is therefore fundamental to public health practice, for it uses data from large populations to generate meaningful information that can be used as the basis of health policy.

The techniques covered in this chapter are needed for the interpretation of the scientific literature, and for providing credible professional advice. Public health practitioners need to learn how to apply these techniques well enough to design and analyse with confidence their own basic studies.

1A.1 HEALTH STATISTICS

Use of routine vital and health statistics to describe the distribution of disease in time, place and person

Information about a defined population that is collected in a consistent manner for administrative reasons is often called **routine data**. These data may be used to describe the needs of and services provided to different population groups. The sources of routine health data, the information collected and the frequency of collection vary between different countries. However, almost all countries process their data into **vital statistics**. These concern the important events in human life, such as births, deaths and migrations.

Two standard and important vital statistics used across the globe for assessing a population's health are life-expectancy and infant mortality.

- Life-expectancy is the expectation of life at birth. It is defined as the period after which half of all persons born have died.
- Infant mortality is the number of children per 1000 live births who die in their first year of life.

Strengths and weaknesses of vital statistics are shown in Box 1A.1.1.

Box 1A.1.1

Strengths of vital statistics	Weaknesses of vital statistics
• Cheap and readily available • Almost complete data recording • Contemporary • Can be used for ecological studies to develop hypotheses • Recorded at regular intervals – can be used for following trends	• Incomplete • Potential for biases (e.g. postmortem inflation of socioeconomic status; diseases with stigma under-reported) • Currency – can become out of date (e.g. census data only recorded every 10 years)

Improving the reliability, validity and completeness of routine data is important to avoid waste and to maximise the use of resources. There should therefore be a good reason to begin or to stop collecting each item of data.

The quality of data can be improved as shown in Box 1A.1.2.

Box 1A.1.2

Computerised data collation and analysis	Improves the accuracy and timeliness of the preparation and dissemination of information
Feedback	Improving feedback of collated data to providers is essential if their interest is to be maintained and their attention to providing quality data sustained
Presentation	Data should be presented in a variety of ways which are meaningful to policy makers, the media, professionals and the lay public
Training	Training the coders and those responsible for data entry in the use of standard definitions, terminology, etc

ROUTINE STATISTICS IN THE UK

Eng The key sources of routine health data in England are listed in Table 1A.1.1. More details are available in Section 3B.

Table 1A.1.1 Routine statistics collected in England

Type	Examples	
Demography	Census	
	General Household Survey†	
Mortality	Mortality statistics (Office for National Statistics)	
	10-yearly supplement on occupational mortality	
Morbidity	Primary care	GP disease registries, e.g. Quality and Outcomes Framework (QOF)
		GP research database
		General Household Survey
	Hospital	NHS Secondary Uses Service (SUS)
		Laboratory results
		A&E attendances
	Registers	Regional cancer registries
		National childhood cancer register
		Confidential Enquiry into Maternal and Child Health
		Congenital abnormalities
		Prostheses
		Transplants
		Confidential enquiries
	Other	Notifiable disease statistics
		Health Survey for England*

**Collects information from about 13 000 people living in private households in Great Britain. Started in 1971 and has been carried out approximately annually since then (ONS 2003).*

†Uses multistage stratified sampling with the postcode address files as the sampling frame. Comprises: questionnaire, height, weight and blood pressure measurement, plus a blood sample.

MORTALITY INDICES

UK Data on mortality are highly accurate in the UK, owing to the legal processes to which registrations of deaths are subject. Mortality indices are shown in Table 1A.1.2.

FERTILITY INDICES

A number of measures exist to describe different features of the reproductive behaviour of a population. These are commonly known as fertility indices. See Table 1A.1.3.

Table 1A.1.2 Mortality indices

Index*	Typical reference period	Numerator	Denominator
Crude mortality rate	1 year	Number of deaths	Mid-year population
Age-specific mortality rate	1 year	Number of deaths aged *X*	Mid-year population aged *X*
Child mortality rate	1 year	Number of deaths in children aged 1–4 years	Mid-year number of children aged 1–4 years
Infant mortality rate	1 year	Number of deaths under 1 year old	Number of live births
Postnatal mortality rate	1 year	Number of deaths in infants aged 4–52 weeks	Number of live births
Neonatal mortality rate	1 year	Number of deaths in the first 28 days	Number of live births
Perinatal mortality rate	1 year	Number of stillbirths + deaths <7 days	Number of live births + stillbirths
Standardised mortality ratio	See Section 1A.5		

**Commonly expressed per 1000 or per 100 000.*

Table 1A.1.3 Fertility indices

Rate	Typical reference period	Numerator	Denominator	Advantages/limitations
Crude birth rate	1 year	Number of live births	1000 total population	Poor indicator of fertility: denominator includes men, children and postmenopausal women
General fertility rate	1 year	Number of live births	1000 women aged 15–44 years	Denominator includes only women, typically of childbearing age
Age-specific fertility rate	1 year	Number of live births in 1 year	1000 women within a particular age band, e.g. 20–24 years	More precise – takes into account differences in fertility at different ages
Total period fertility rate	1 year	Sum of the age-specific rates across an average woman's reproductive lifetime. UK = 1.66 in 2005		Enables comparisons between countries over time

DESCRIPTIVE EPIDEMIOLOGY

Epidemiology is the study of the patterns, causes and control of disease in groups of people. In descriptive epidemiology, the three major dimensions used to describe the occurrence of disease are **time**, **place** and **person**.

TIME

Considers **when** the disease occurs and how it changes/has changed over time, described by:

- Epidemic curves: acute increases in disease frequency
- Seasonal variation: cyclical patterns in disease frequency, e.g. seasonal influenza
- Secular trends: trends over decades and centuries
- Point events: sudden emergence of disease at a particular time.

PLACE

Describes **where** the incidence is high/low, where incidence is changing/has changed, considering:

- International: ecological comparisons may suggest hypotheses regarding causation
- National: highlights urban–rural patterns and patterns related to deprivation
- Small area: compare census data/index of multiple deprivation/town-centre data.

PERSON

Describes **who** is affected, taking into account characteristics such as:

- Age
- Sex
- Occupation/social class
- Ethnicity
- Behaviour/lifestyle.

1A.2 NUMERATORS, DENOMINATORS AND POPULATIONS AT RISK

In epidemiology, a **numerator** is a feature that has been counted (e.g. number of deaths); it forms the upper part of a fraction. A **denominator**, for epidemiological purposes, is usually the population from which the numerator was drawn. The denominator is the lower part of a fraction and is often restricted to a particular time period.

A '**population at risk**' is a subgroup of a wider population that has been subjected to the exposure of interest and which therefore has a propensity to have been affected by that exposure. For example, the population at risk of prostate cancer is all males who have not previously had a prostatectomy.

The numerator and denominator can be combined into a **ratio**, a **proportion** or a **rate**, as detailed below. Some of the most frequent mistakes in applying epidemiological principles to real-world problems come from a failure to define and to count numerators and denominators accurately.

RATIO

$$\text{Ratio} = \frac{n_1}{n_2} = n_1{:}n_2$$

where n_1 and n_2 are numbers.

For example,

$$\text{Ratio of males to females in a population} = \frac{\text{Number of males in a population}}{\text{Number of females in a population}}$$

A ratio is often expressed as odds, which is a single number. For example, if a bag contains 2 white balls (n_1), 3 black balls (n_2) and 5 grey balls (n_3), then the ratio of black balls to white balls is 3:2 (three to two). This can be simplified by dividing both the numerator and the denominator by the denominator to give 1.5:1 (one and a half to one). Expressed in odds, it is simply 1½.

PROPORTION

$$\text{Proportion } \frac{n_1}{N}$$

where n_1 is a subpopulation of the whole study population, N.

If a bag contains 2 white balls (n_1), 3 black balls (n_2) and 5 grey balls (n_3), then the proportion of black balls is 3/10 or 0.3.

For example,

$$\text{Proportion of men in a population*} = \frac{\text{Number of males in a population}}{\text{Number of males + Number of females in a population}}$$

**Assumes that all people in the population are either male or female*

RATE

$$\text{Rate} = \frac{N}{P \times T}$$

where N is the numerator, P is the number of people in the population and T is a period of time.

This is a measure of frequency of the occurrence of a phenomenon. The denominator is constituted from both population and time.

For example,

$$\text{Incidence of epilepsy} = \frac{\text{Number of new cases of epilepsy}}{\text{Population} \times \text{Reporting period}}$$

Denominators should include only those *at risk*, and not those who cannot possibly develop the disease. For example, people who have been immunised against a disease should not be included in studies looking at the rate at which people acquire that disease, since they cannot be considered in the same way to be at risk as those who have not been immunised.

1A.3 TIME AT RISK

Time at risk describes the total amount of time that individuals within a study spend at risk of developing the disease of interest. The concept is particularly important in the analysis of cohort studies. One of the problems with conducting cohort studies is that some subjects will join the study after data collection has started. Others will leave the at-risk population early because they have:

- Died
- Moved away
- Become a case
- Been lost to follow-up for another reason (e.g. withdrawn from the study)
- Been censored (see below).

To correct for this variation, the 'total person-time at risk' is used as the denominator in calculations of morbidity or mortality relating to cohort studies. Person-time represents the sum of all the individual times at risk.

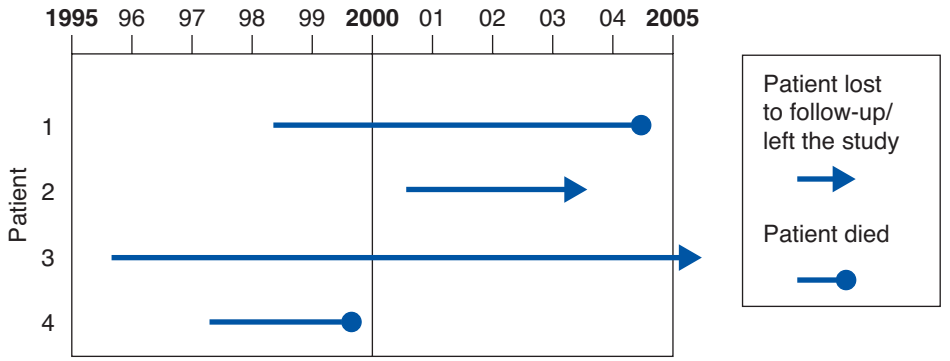
$$\text{Incidence rate} = \frac{\text{Number of new cases in a specified period}}{\text{Total person-time at risk}}$$

Box 1A.3.1 represents a small cohort study, and illustrates the concept of time at risk. As time progresses, the number of individuals at risk will fall as people die, become a case or are lost to follow-up. Censoring occurs when the value of an observation is only partially known. At the end of this study, patient 3 was still alive and therefore spent *at least* 10 years at risk. This is known as **right-censoring**. **Left-censoring** (where a data point is below a certain value but it is unknown by how much) and **interval censoring** (where a data point is somewhere in the interval between two values) are also possible.

Sometimes, participants are replaced by new people recruited into the study. For studies of short duration and where few individuals leave or join the at-risk population, it is reasonably accurate to use the number of individuals at the start of the study as the denominator in incidence calculations.

Box 1A.3.1 Cohort study illustrating the concept of time at risk

	Patient 1	Patient 2	Patient 3	Patient 4
Date of entry to study	1998	2000	1995	1997
Event	Death	Loss	Censored	Death
Date of event	2004	2003	2005	1997
Person-years in study	6	3	10	2



CALCULATION OF MORTALITY RATE USING PERSON-TIME AT RISK

Where the study duration is longer or the likelihood of individuals leaving the at-risk population is greater, it is preferable to use a more accurate calculation of the incidence rate using person-time at risk. See Box 1A.3.2.

Box 1A.3.2

Calendar time	Person-years	Events	Death rate (number/year)
1995–1999	8	1	0.125
2000–2004	12.5	1	0.080
(2005–2009)	(0.5)	(0)	(0)

1A.4 METHODS FOR SUMMARISING DATA

See also Section 1B.7.

Data are considered as either **qualitative** (non-numeric) or **quantitative** (numeric). Subtypes of each type of data are listed below.

QUALITATIVE DATA

Qualitative data are non-numeric. Some qualitative data are categorical: they describe different categories or states that a subject may fall into. These may be further categorised into nominal data and ordinal data.

NOMINAL DATA

Nominal data have no order and thus only give **n**ames or labels to various categories. Examples include the ABO blood groups (A, B, AB, O), the names of colours (red, yellow, orange, green, violet, purple) and types of hospital ward (rehabilitation, surgical, medical). See Box 1A.4.1.

ORDINAL DATA

Ordinal data have **o**rder, but the interval between measurements is not meaningful.

Some qualitative data are ordinal data: there is a natural order to the states, but no clear numerical relationship between them. Examples are [poor, fair, good, better, best] and [very quiet, quiet, normal volume, loud, loudest].

Although ordinal data should not be used for calculations, it is not uncommon to find averages calculated from data collected of the type strongly disagree, disagree, neither agree nor disagree, agree, strongly agree. See Box 1A.4.1.

QUANTITATIVE DATA

Quantitative data are numeric, and they are further classified as either discrete or continuous. Data may come from one or more of the following categories.

DISCRETE DATA

Discrete data have a finite number of possible numerical values. Examples include the number of children with brown eyes in a class of 30 children, or the number of times someone has been admitted to hospital in their lifetime.

CONTINUOUS DATA

Continuous data include measurable quantities of length, volume, time, mass, etc. They frequently have an upper or lower limit, e.g. height cannot be <0.

INTERVAL DATA

Interval data have meaningful intervals between measurements, e.g. the age groups 0–4, 5–9, 10–14 ... 90+. They are typically displayed as a table or histogram.

RATIO DATA

Ratio data are the most flexible data to work with, since they contain the most information of any data type. It becomes meaningful to say not only that A scored 1 and B scored 2, but that B is twice as good as A. Ratio data are ideal for use as outcome variables in regression. See Box 1A.4.1.

BINARY DATA

Binary data are a special type of data that have just two values. Depending on how they are analysed, they may be considered to have properties of ordinal data, interval data or ratio data. They are very common data in epidemiology, since they accurately describe many of the states that are of interest. For example, did a patient improve following a particular treatment – or not? Is the case alive or dead? Was the case in the treatment group or the control group?

Binary data often take values such as True/False or Male/Female. For analysis, they are usually transformed into values of 0 and 1. Data of this type are often used in logistic regression (see Section 1B.14), where they might be the outcome variable itself, or a binary coefficient reported as an odds ratio.

Box 1A.4.1

Example: Colours

To most people, black, brown, red, orange, yellow, green, blue, violet, grey and white are just names of colours – **nominal data**. To an electronics student familiar with colour-coded resistors, these data are in ascending order and thus represent **ordinal data**. To a physicist, red, orange, yellow, green, blue and violet correspond to specific wavelengths of light and would be considered **ratio data**.

Example: Temperatures

Celsius	Fahrenheit	Kelvin
–273.15°C	–459.67°F	0 K
0°C	32°F	273.15 K
100°C	212°F	373.15 K
–17.8°C	0°F	255.4 K

Only the Kelvin scale has a true zero and can thus be used as a ratio scale. A temperature of 200 K is twice as hot as a temperature of 100 K. This is not true for the Celsius and Fahrenheit scales.

SUMMARISING DATA

Discrete variables (e.g. blood groups) are typically summarised as **proportions**. For example, the distribution of the common ABO blood groups in England is shown in Table 1A.4.1.

Continuous variables (e.g. blood pressure) are described using both:

- A measure of **central tendency** (mean/median/mode)
- A measure of **spread** (range/variance/standard deviation).

If a population is **normally distributed**, then it can be described by both its **mean** and its **standard deviation** alone. For an example see Box 1A.4.2.

Table 1A.4.1 Prevalence of different ABO blood groups in the UK population

Blood group	Proportion
A	0.41
B	0.08
AB	0.03
O	0.48

Reproduced from www.stats.ox.ac.uk.

Box 1A.4.2

Ages of children in a childhood leukaemia cluster

1	5
2	5
2	6
2	7
2	7
3	9
3	10
4	13
4	15

$N = 18$ children

Define n as the set of individual observations

Define n_i as a particular observation

Define N as the total number of observations

Measures of **central tendency** of observations:

Mean age $= \frac{\sum n_i}{N} = 100/18 = 5.6$ years

Modal age (most frequently occurring) = 2 years

Median age = 4.5 years. This is the value midway between the 9th and 10th observations of the 18 ranked observations in the data

MEASURES OF THE SPREAD OF OBSERVATIONS

Some commonly used terms are listed in Table 1A.4.2 with further properties described in Section 1B.7.

Table 1A.4.1 Measures of spread

Measure	Description
Range	The difference between the largest and smallest value
Percentiles	The value below which p% of the observations in a population fall is called the pth percentile
Interquartile range	The difference between the value at the 25th centile and the value at the 75th centile (i.e. between the 1st and 3rd quartiles)
Standard deviation	Measure of the spread of observations about the mean of the sample. Takes the same units as the data; 1.96 standard deviations either side of the mean covers 95% of the population Used to describe the data '**D**eviation → **D**escription' Square root of the variance of the sample, i.e. standard deviation $s = \sqrt{\text{variance}}$ where variance = $s^2 = \frac{\sum(\bar{x} - n_i)^2}{n-1}$
Standard error of the mean	The error associated with measuring a sample **mean** is determined by both the standard deviation and the number of observations in the sample. The more observations, the more precise is the estimate of the mean Measure of variability of the mean of the sample Used to make *estimations* about the true mean of the population '**E**rror → **E**stimation' $\text{SE} = \frac{s}{\sqrt{n}}$
Central limit theorem	If repeated samples are taken from *any* population, then the means of these samples will tend towards a normal distribution, even if the population is not normally distributed

1A.5 INCIDENCE, PREVALENCE AND STANDARDISATION

Incidence and prevalence (direct and indirect standardisation)

Incidence and prevalence are measures of occurrence.

INCIDENCE

Incidence relates to new occurrences. There are three related measures of incidence as shown in Table 1A.5.1.

PREVALENCE

Prevalence relates to existing occurrences.

Prevalence is also called the *'point prevalence',* i.e. the proportion of a population with a disease. It is approximately equal to incidence x duration *provided that* the incidence and the death/recovery rate have been stable for the

disease over some preceding time. Period prevalence relates to the proportion of the population with a disease during a specified period. These two measures of prevalence are shown in Table 1A.5.2.

Table 1A.5.1 Measures of incidence

Concept	Alternative name	Definition
Incidence	Incidence rate*	Number of new events during a specified time period
Incidence rate	Force of morbidity Incidence density	Number of new events divided by the total person-time at risk Used to describe studies where there have been varying periods of follow-up (e.g. 3 cases during 25 person-years ≡ 12 cases per 100 person-years)
Cumulative incidence	Risk	Proportion of a population who *become* diseased in a defined time period. This is a measure of the risk that an individual will become diseased during a defined time period, e.g. the attack rate during an epidemic

Note:

1. *Incidence (unlike prevalence) is not affected by disease survival*
2. *The denominator should include only those* 'at risk'

**Rather confusingly, the term incidence rate is sometimes used to mean incidence. The incidence and incidence rate can generally be considered as the same concept, with the number of person-years being more or less accurately measured.*

Table 1A.5.2 Measures of prevalence

Concept	Definition
Point prevalence	Proportion of a population with a disease *at* a particular point in time
Period prevalence	Proportion of a population who had the disease *during* a specified period

Note: since period prevalence considers both cases at the start and new cases arising during the time period, it includes features of both prevalence and incidence.

STANDARDISATION

Also known as **adjustment**, this technique is required to make comparisons between populations of differing demographic structures, where crude mortality rates would be misleading. For example, the technique is typically the first step in addressing questions such as, 'Does town A have a higher mortality rate than town B?' or, 'Are 10 incident cancers in a workforce of 100 people over a 10-year period more or less than we would have expected had national rates applied?' The technique is mostly used to compare populations that differ in their age structure – but it may also be applied to correct for differences in gender or social class (or combinations of these and other variables).

To standardise data, populations are divided into strata based on differences that might potentially influence the comparison that is being made. For example, to compare the numbers of deaths in a retirement community with a town having a high proportion of young mothers, the mortality rates would need to be age standardised.

Two forms of standardisation that are commonly used are **direct** and **indirect**.

DIRECT STANDARDISATION

Direct standardisation may be used to compare two populations in different regions, or the same population at two different periods of time. In direct standardisation, the age- or stratum-specific mortality rates of the *observed*

population are known (e.g. rate of lung cancer in 20- to 24-year-old men). These rates are then applied to a standard or reference population. Data typically need to be available concerning a large number of subjects in order to have adequate numbers in each stratum to be confident about the estimate.

- Start by choosing a single population (e.g. one of those being compared, their average or an outside population)
- Break this single population down into individual age bands ('standard age structure')
- For the populations being compared, take the age-specific mortality rate for each age band and multiply it by the size-weighting of that age band from the standard age structure
- Sum all of these to obtain the **age-standardised mortality rate.** The actual value of this adjusted rate is meaningless but it shows the true *difference* in rates between the populations being compared.

THE EUROPEAN STANDARD POPULATION

The European Standard Population (see **Table 1A.5.3**) is used to compute directly age-standardised rates. The same population is used for males, females and all persons.

The example in Box 1A.5.1 shows how mortality data from 2 years can be compared using direct standardisation.

Table 1A.5.3 European Standard Population

Age group	Number in age group
0	1600
1–4	6400
5–9	7000
10–14	7000
15–19	7000
20–24	7000
25–29	7000
30–34	7000
35–39	7000
40–44	7000
45–49	7000
50–54	7000
55–59	6000
60–64	5000
65–69	4000
70–74	3000
75–79	2000
80–84	1000
85+	1000
TOTAL	**100 000**

Reproduced from the World Health Annual of Statistics (1991), based on Waterhouse et al (1976).

INDIRECT STANDARDISATION

With indirect standardisation, the population under study is usually small. The technique hinges on the calculation of the rate of death that would have been **expected** in the study population had a comparison rate applied instead. This permits the determination of the standardised mortality ratio (SMR).

- Start with the stratum-specific death rates of a standard population (e.g. European Standard Population if age is the effect to be standardised)
- Use these to calculate expected number of deaths in each stratum of the study population
- Add up the expected number of deaths for each age band:

  Define E as the expected number of deaths

  Define k as the number of strata

  Define n_i as the number of people in the ith stratum

  Define R_i as the rate of death in the ith stratum

  $$E = \sum_{i=1}^{k} n_i R_i$$

- Standardised mortality ratio (SMR) is then calculated as (Observed deaths/Expected deaths).

The SMR indicates whether mortality in the study group (after correction for age, sex, etc.) is unusual when compared with a standard population. In occupational mortality studies, comparisons are often made against two standard populations: (i) an unexposed population from the same occupation and (ii) the general population.

When used to estimate the association between an occupational exposure and a disease, the SMR underestimates the true magnitude of association because the general population contains both exposed and unexposed individuals.

Two SMRs should never be compared with each other: the number of expected cases in each group depends on the group's actual age/sex/race composition. Instead, direct standardisation can be used to compare the two groups. See Box 1A.5.2 for an example of this calculation.

Box 1A.5.1

Example: Comparison of 2 years' mortality data using directly age-standardised rates

Deaths from all malignant neoplasms (ICD-10 C00–C97) in persons aged under 75 in a defined population[†]

Stage 1: calculation of age-specific death rates per 100 000 population in a defined population

(a) Number of deaths by age group in the defined population

Year	<1	1–4	5–9	10–14	…	65–69	70–74
2001	0	0	1	1	…	53	83
2002	0	0	0	0	…	63	68

(b) Population by age group in the defined population

Year	<1	1–4	5–9	10–14	…	65–69	70–74
2001	1900	8200	11 100	11 800	…	9300	8300
2002	1900	8200	11 100	11 800	…	9300	8300

(c) Age-specific death rates per 100 000 in each age group/year = (number of deaths ÷ population in each age group) × 100 000

Year	<1	1–4	5–9	10–14	…	65–69	70–74
2001	0	0	9.01	8.47	…	569.89	1000.00
2002	0	0	0	0	…	677.42	819.28

Stage 2: calculation of European age-standardised annual rates

Method: find the expected number of cases in the European Standard Population had the death rate observed applied. Multiply the death rate in a given age group/year by the number of people from the correct cell of the European Standard Population. Sum these to derive the 'expected' numbers of deaths in the population as a whole.

Year	<1	1–4	5–9	10–14	…	65–69	70–74	Sum (<75 s)	Standard rates*
2001	0	0	0.63	0.59	…	22.80	30.00	112.60	117.29
2002	0	0	0	0	…	27.10	24.58	109.84	114.42

**Obtained by dividing the sum of expected numbers at ages under 75 by the European Standard Population at ages under 75 (96 000/100 000).*

[†]*Adapted from: NHS performance indicators: standardisation. Available online at: www.chi.nhs.uk/Ratings/Downloads/direct_standardn_meth.pdf.*

Box 1A.5.2

Example: Indirect standardisation of deaths from cirrhosis of the liver in British doctors (after Bland 2000)

Age group	Mortality per million men per year in the UK	Number of male doctors in the UK	Expected deaths
15–24	5.859	1080	0.0063
25–34	13.050	12 860	0.1678
35–44	46.937	11 510	0.5402
45–54	161.503	10 330	1.6683
55–64	271.358	7790	2.1139
TOTAL			4.4966

If 14 deaths were observed among British doctors, then:

Indirectly, SMR compared with the background population = O/E = 14/4.4966 = 3.11

SMR (*100 as integer) = 311

Exact 95% confidence interval = 1.70–5.22 (SMR 170–522)

1A.6 YEARS OF LIFE LOST

One way of considering the impact of a particular disease or risk factor is to consider how many years people might have expected to have lived had their lives not been curtailed by the disease. For example, deaths from road traffic collisions affect mostly young males who otherwise may have been expected to live into their 70s or 80s.

'Years of life lost (YLL)' is a measure of **premature mortality** and it explicitly places more importance upon deaths that occur in the young than upon those in the elderly. It is therefore a **value-laden** statistic, which reflects the wish of society to prevent avoidable causes of death in younger people, and to avoid the loss to society of investment in raising children and young adults.

To calculate YLL in its most simple form, an upper age limit (e.g. 75) is chosen. A person dying at the age of 60 contributes 15 YLL (75 – 60 = 15) to the calculation. A person dying at age 15 contributes 60 years. Deaths occurring in people aged over 75 are not included in the calculation. Infant deaths may or may not be included.

In more complex calculations, an actuarial assessment of the expectation of life for each person will be undertaken. Life-expectancy is calculated using current life-tables: age-specific mortality rates are applied to a hypothetical population

$$\sum N \times L$$

where:

N = number of deaths in a particular age–sex group

L = life-expectancy of this age–sex group in this population.

A specific type of YLL can be calculated for smoking:

Smoking-attributable YLL = (Age- and sex-specific mortality attributable to smoking) × (Remaining life-expectancy in this age and sex group).

Note that YLL will underestimate the burden of disease due to chronic conditions, which have a low mortality at a young age. To take this factor into account, **health-adjusted life-expectancy** (**HALE**) can be used instead of crude life-expectancy.

HALE = Σ (Number of life years lived in each age group) × (Mean health state score for that age group)

1A.7 MEASURES OF DISEASE BURDEN

Measures of disease burden (event and time based) and population attributable risks, including identification of comparison groups appropriate to public health

See Section 1A.10 for risks.

A number of measures exist to assess the weight of disease affecting a community, taking into account both the **number** of events and the **impact** of these events on the population. They each depend on the availability of data.

MEASURES OF DISEASE BURDEN BASED ON EVENTS

These studies of **incidence** require data from routine sources:

- Death certificates (if the disease burden is in mortality rather than morbidity)
- Hospital episode data (if the disease results in hospital admissions)
- Disease register (if a dedicated register for that disease exists)
- Statutory notifications: infectious disease, procedures, e.g. abortion.

MEASURES OF DISEASE BURDEN BASED ON TIME

Where there are no routine event-based data collected, a **prevalence** ('cross-sectional') survey may be needed. Figure 1A.7.1 is a graph that illustrates the relative contributions of various conditions to the sum of 'disability-adjusted life-years' (DALYs) for the UK.

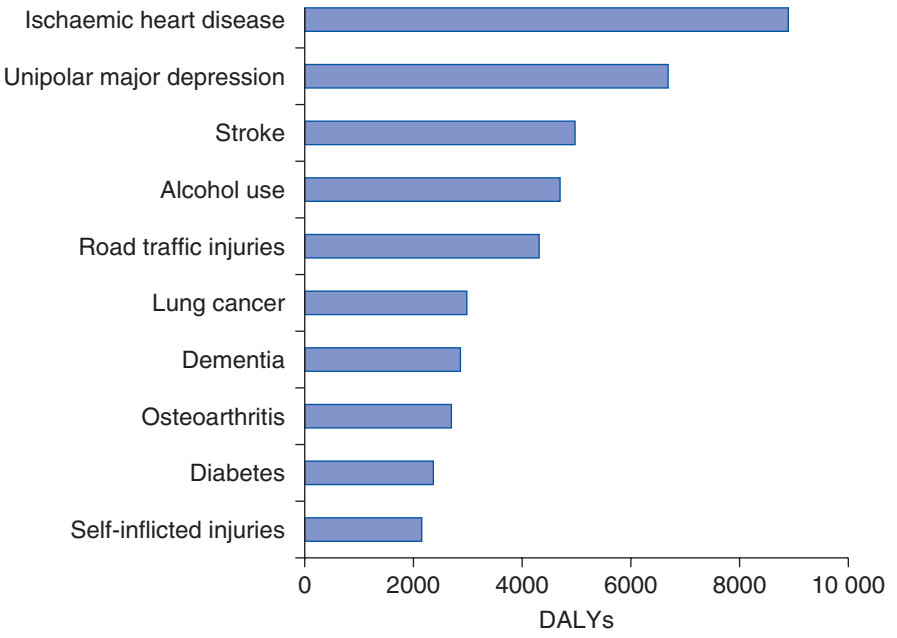


Figure 1A.7.1 Disability burden of a range of conditions in the UK. *Reproduced from the WHO*

1A.8 VARIATION

Sources of variation, its measurement and control

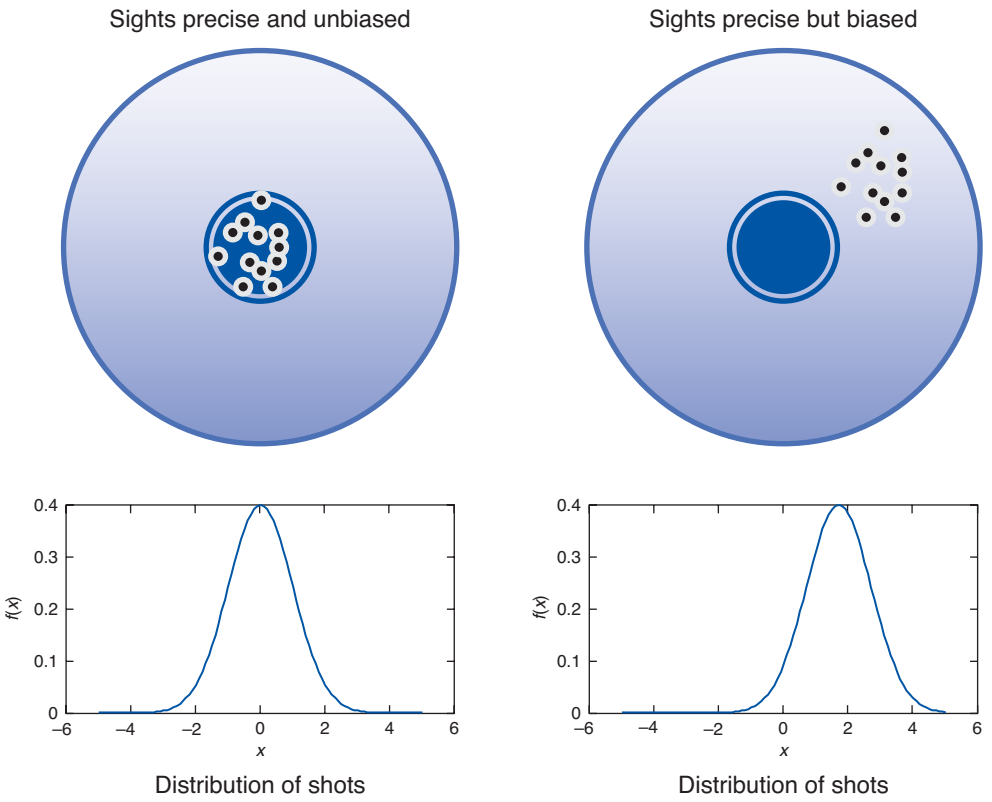
Variation arises from differences between populations or between individuals within a population. It may be due to random or non-random factors.

Measurements can differ for many reasons. Some differences will be due to genuine dissimilarities between subjects. This sort of variation is often random, and follows a normal distribution. Some people are taller than average, some are shorter than average. This is everyday, random variation.

Other differences will be due to discrepancies in the way that measurements were made or recorded. A ruler may measure only to the nearest 5 cm. It judges a person's height as 175 cm and another's as 170 cm. A more precise ruler, accurate to 0.1 cm, would confirm that the first subject is in fact 176.1 cm tall and the second is 172.4 cm tall. The aim for any investigator is to minimise the error associated with making the measurements, thereby focusing attention on the *real* differences between individuals.

Measurements can be biased or unbiased, precise or imprecise, or some combination of these.

By way of analogy, consider a markswoman firing shots at a target. Her sights might be on-target or off-target, and she might be a good shot or a bad shot. The diagrams in Figure 1A.8.1 illustrate the problems that she might encounter when aiming at a target. Making measurements in epidemiology is much the same.



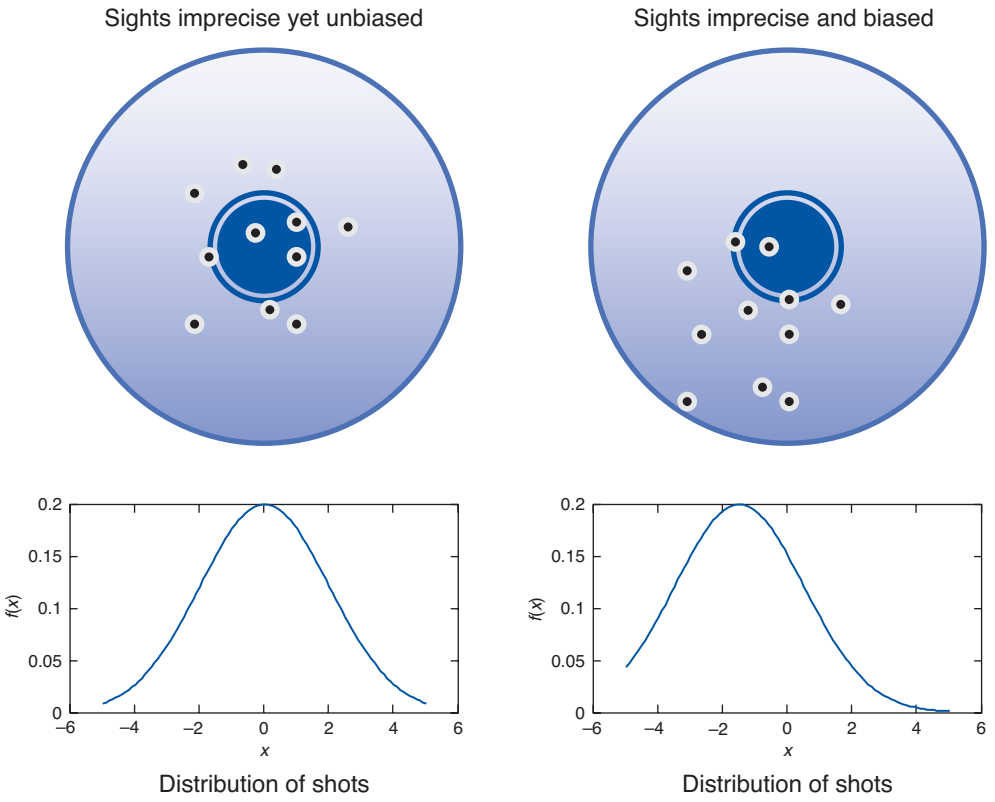


Figure 1A.8.1 Imprecision and bias for a gun and a target

Sometimes a group of people differs systematically from another. One subject might belong to an African pygmy tribe and another subject might be from a tall population, e.g. the Netherlands. The observed differences in height between these two individuals are not due to random variation but are attributable to systematic genetic differences.

It is this sort of variation that is of interest to epidemiologists. Most epidemiological techniques aim to control random variation in order to establish whether the systematic variation is significant.

1A.9 ERRORS IN EPIDEMIOLOGICAL MEASUREMENT

Common errors in epidemiological measurement, their effect on numerator and denominator data, and their avoidance

See also Section 1B.10.

Measurement error can be defined as any mistake that occurs during the process of applying a standard set of values (i.e. a measurement scale) to a set of observations. In epidemiology, such errors can lead to the misclassification of cases and controls. Acknowledgement of measurement error is important in research, and leads to more robust and defensible scientific results.

Measurement errors are either **random** (i.e. occurring due to chance) or **systematic** (i.e. persistent, non-random, differences between the observed measurement and its true value). These errors will lead respectively to **non-**

differential errors (affecting all groups equally) and **differential** errors (affecting one group more than another). Table 1A.9.1 compares random and systematic errors in epidemiological measurement.

Table 1A.9.1 Random and systematic errors in epidemiological measurement

Random error	Systematic error
Occurs due to chance	Occurs due to non-random factors
Affects all groups equally (non-differential)	Affects some groups more than others (differential)
Exposure: misclassification of exposures is *equal* for cases and controls (e.g. 20% of cases and 20% of controls are misclassified as being exposed)	**Exposure:** misclassification of exposures *differs* between cases and controls (e.g. 20% of cases misclassified as having been exposed but only 5% of controls misclassified)
Outcome: misclassification of outcomes is *equal* for exposed and non-exposed (e.g. 30% of exposed and 30% of non-exposed are misclassified as having the disease)	**Outcome:** misclassification of outcomes *differs* between exposed and non-exposed (e.g. 19% of the exposed are misclassified as having the disease but only 7% of non-exposed are misclassified as having the disease)
Bias is towards the null-hypothesis (i.e. dilution of the study's findings)	Bias can be in any direction
Less threatening for a study than systematic error	More threatening for a study than random error
Example: a study of lung cancer in relation to proximity of residence to a coke oven classifies subjects (cases and populations) by distance of residence from the oven at the time of follow-up. Here there is misclassification due to migration (not all people living near the oven at the time of follow-up will have lived there at the aetiologically relevant time) but this error occurs randomly	**Example:** a study assessing the association between visual display unit (VDU) usage and spontaneous abortion. Here, cases are more likely to recall VDU usage compared with controls, particularly if there has been media interest in this hypothesis. Therefore the association measured is likely to be greater than the true association

EFFECT OF MEASUREMENT ERRORS

Measurement errors can affect:

- Dependent variables (outcomes such as disease)
- Independent variables (risk factors such as exposures)
- Confounder variables: see Section 1A.15
- Effect modifiers: see Section 1A.15.

With regard to the measurement of confounder variables, an apparent association between the exposure and the disease may persist after statistical adjustment. Such residual confounding is particularly problematic where the variable being measured is difficult to quantify (e.g. measurement of socioeconomic status). For this reason, residual confounding should always be considered where an association persists following statistical adjustment for a known confounder.

AVOIDANCE OF MEASUREMENT ERROR

Measurement error can be avoided or accounted for at various stages of the study. See Table 1A.9.2.

Table 1A.9.2 Avoiding and accounting for measurement error

Stage of study	Strategy for dealing with measurement error
Design	Set out to measure reliability: correlation coefficients (continuous variables) or Cohen's kappa (categorical variables) Blinding Standardised measurement instruments Use multiple sources of information (questionnaires, direct measurements, registries, case records) Use multiple controls
Data collection	Administer instruments equally to: cases/controls exposed/unexposed
Data analysis	Perform a sensitivity analysis to test the robustness of the findings (see Section 4D.5)
Reporting	Consider and mention potential random and systematic errors

1A.10 CONCEPTS AND MEASURES OF RISK

See also Sections 1A.5, 1A.7 and 2F.2.

RISK

Also called 'cumulative incidence'.

Risk, in an epidemiological sense, is the likelihood of some future event. For example, the risk that a 30-year-old male will die at some point in the future is 100%. The risk that this 30-year-old male will die in the next year is about 1 in a 1000.

ATTRIBUTABLE RISK

Also called 'absolute risk' or 'excess risk', the **attributable risk** is the difference in the risk or rate of disease between the exposed group, I_e (pronounced I sub e), and the unexposed group, I_0 (I sub zero), assuming causality (see Section 1A.13).

$I_e - I_0$

The **attributable risk per cent**, also called proportional attributable risk or aetiological fraction when expressed as a proportion, is the percentage of disease in the exposed group, which is attributable to the exposure.

$\frac{I_e - I_0}{I_e} \times 100\%$ or $\frac{RR - 1}{RR} \times 100\%$ where data on disease incidence is not available RR = relative risk (see page 23).

This is a useful epidemiological statistic that is readily understood by non-epidemiologists. For example, in people who smoke, about 90% of lung cancers arise *because* they smoke. See Box 1A.10.1.

Population attributable risk is the excess rate of disease in the whole population that is attributable to the exposure. Define I_T as the rate of disease in the **t**otal study population (both exposed and unexposed groups). Define I_0 as the rate in the unexposed population.

Population attributable risk = $I_T - I_0$

Box 1A.10.1

Example: Population attributable risk of smoking in lung cancer	
Mortality in whole population	= 55 per 100 000
Mortality in non-smokers	= 16 per 100 000
Population attributable risk	= Rate in population – Rate in non-smokers
	= 55 – 16
	= 39 deaths per 100 000/year

If generalising for a broader population, the true prevalence of exposure in that broader population must be known from an outside study; otherwise it will have to be assumed that the broader population has the same exposure as the study population.

POPULATION ATTRIBUTABLE (RISK) FRACTION

This statistic measures the proportion of disease in the study population that is attributable to the exposure. This is the **preventable fraction**. It is one of the most important indices in prioritising population interventions and may also be called the *'population attributable risk per cent'.* It can be calculated in two ways. See Box 1A.10.4 for an example of its calculation.

$$\text{Population attributable fraction} = \frac{\text{Population attributable risk}}{\text{Rate of disease in population}}$$

$$= \frac{\text{Prevalence in exposed population} \times (\text{RR}-1)}{[1 + \text{Prevalence in exposed population } (\text{RR}-1)]}$$

COMPARISON GROUPS USED IN PUBLIC HEALTH

Comparison groups may be used in the control arm of a study (see Section 1A.19). At the population level, comparators are used for needs assessment studies and for assessing the quality of health care. Factors to consider when choosing comparison groups include:

- Size of population
- Area characteristics (rural/urban)
- Age structure
- Ethnicity
- Socioeconomic characteristics
- Disease burden (mortality and morbidity)
- Health service usage

- Health service provision, funding and organisation (relevant to international comparisons)
- Provision of other relevant services, e.g. social care.

RELATIVE RISK (RATIO)

Relative risks are used to measure the strength of association between an exposure and an outcome. They are used to assess aetiological strength and can be calculated as the **risk ratio**, **rate ratio**, **odds ratio** or **standardised mortality ratio**.

RISK RATIO

Risk ratio = (Risk of disease in exposed) ÷ (Risk of disease in non-exposed).

See Box 1A.10.2 for an example of how to calculate a risk ratio.

Box 1A.10.2

Example: Calculating the risk ratio for hypertension among cardiac patients

A consultant believes that hypertension is unusually common in patients in her clinic who have had a coronary bypass operation.

		Hypertension		
		Yes	No	
Coronary bypass	Yes	15	467	482
	No	70	6364	6434
		85	6831	

Risk ratio = (Risk of disease in exposed) ÷ (Risk of disease in unexposed)

= (15 ÷ 482) ÷ (70 ÷ 6434)

= 0.0311 ÷ 0.0109

= 2.86

In this clinic, patients who had had a coronary bypass operation were 2.86 times as likely to have hypertension as those who had not had a bypass.

RATE RATIO

Rate ratio = (Incidence rate in exposed) ÷ (Incidence rate in non-exposed).

ODDS RATIO*

Odds ratio = (Odds of exposure in cases) ÷ (Odds of exposure in controls).

**Mainly used for case–control studies only*

Note: in cohort studies the time period must be stated (because the risk of dying will always be 100% in the long run for both groups).

See Box 1A.10.3 for an example of how to calculate an odds ratio.

STANDARDISED MORTALITY RATIO (SMR)

Standardised mortality ratio = (Number of cases observed) ÷ (Number of cases expected) × 100%.

The SMR is often used in occupational settings.

Box 1A.10.3

Example: calculating an odds ratio for risk of stroke among rheumatology patients

A consultant believes that, among the patients attending her rheumatology clinic, stroke is more common among those who have recently begun immunotherapy.

		Stroke	
		Yes	No
Immunotherapy started in last year	Yes	8	168
	No	184	9813

Odds ratio = (Odds of disease in exposed group) ÷ (Odds of disease in non-exposed group)

= (8 ÷ 168) ÷ (184 ÷ 9813)

= 0.048 ÷ 0.019

= 2.54

So, the consultant is proved right: stroke is 2.54 times more common in patients attending the rheumatology clinic who have had immunotherapy in the previous year compared with those attending the rheumatology clinic who have not had immunotherapy.

ATTACK RATE AND SECONDARY ATTACK RATE

These measures are used in outbreaks or epidemics of infectious diseases. Both are **risks**.

$$\text{Attack rate} = \frac{\text{Number of new cases}}{\text{Population at risk in a defined time period}}$$

$$\text{Secondary attack rate} = \frac{\text{Number of cases from exposure to primary case}}{\text{Total number exposed}}$$

1A.11 THE ODDS RATIO

See Section 1A.10.

Box 1A.10.4

Example: Attributable risk fraction of lung cancer associated with smoking

34% of the population smoke. The incidence of lung cancer is 376.8/100 000 per year among smokers and is 13.4/100 000 per year among non-smokers.

The key point to remember is that smokers suffer from lung cancer because of the effect of smoking *and* because of the background causes which non-smokers also experience.

So, in 100 000 of the population there will be:

34 000 smokers (who will have the risk from smoking plus the background risk); and

66 000 non-smokers (who will just have the background causes)

(a) Number of cases of lung cancer among smokers that are due to their smoking habit:

$= [(376.8 - 13.4) \div 100\,000] \times 34\,000$

= 123.6 cases (this is the population attributable risk)

(b) Number of cases of lung cancer among smokers due to background risk:

$= (13.4 \div 100\,000) \times 34\,000$

= 4.6 cases

(c) Number of cases of lung cancer among non-smokers due to background risk:

$= (13.4 \div 100\,000) \times 66\,000$

= 8.8 cases

Number of cases of lung cancer in population:

= (a) + (b) + (c)

= 123.6 + 4.6 + 8.8

= 137.0

Attributable risk fraction:

$$= \frac{\text{Population attributable risk}}{\text{Rate of disease in population}} \times 100\%$$

$$= \frac{(a)}{(a) + (b) + (c)} \times 100\%$$

$= (123.6 \div 137.0) \times 100\%$

= 90.2%

1A.12 RATE RATIO AND RISK RATIO (RELATIVE RISK)

See Section 1A.10.

1A.13 ASSOCIATION AND CAUSATION

The concepts of association and causation are fundamental to the science of epidemiology. An association is a statistical link between two variables. It is assessed using a statistical test that calculates how unlikely it is that the finding occurred due to chance alone. An association does *not* imply that one factor brought about the other. For example, there is an association between eating ice-cream and hot weather, but eating ice-cream does not cause hot weather.

Association = statistical link

Association ≠ cause and effect.

Causality implies that one factor caused the other. Once an association has been established, then determining causality involves a subsequent judgement of evidence. Causality is determined according to a number of set criteria (see below). An epidemiologist must never confuse association with causality.

It is important to be cognisant of three phenomena that can erroneously make an association appear to be present when no such association exists in reality. These three phenomena are **chance**, **bias** and **confounding**.

CHANCE

Epidemiological studies make inferences regarding the wider population based on observations from a sample. However, an association might be observed in the sample due to luck of the draw, i.e. the sample does not truly reflect the wider population. The chance of observing such an unrepresentative association falls as the sample size increases.

The null hypothesis states that the association observed in the sample was due to chance alone. So the probability that the null hypothesis is true (called the *p* value) is the probability that an association at least as extreme as that observed could have occurred due to chance alone. If the *p* value is sufficiently low, i.e. sufficiently unlikely, then the null hypothesis is rejected in favour of the alternative hypothesis (which states that the association observed in the sample truly exists in the wider population).

Box 1A.13.1

Type I error (α)	Mistaken rejection of the null hypothesis when it was in fact true Experiment says that the treatments are different, when the truth is that they are the same (false positive) Value of α is the significance level (*p* value) of the test; it is decided before data are collected
Type II error (β)	Mistaken rejection of the alternative hypothesis when it was in fact true Experiment says that the treatments are not different when the truth is that they are different (false negative) Occurs when the sample size is too small
Power ($1 - \beta$)	Probability of correctly rejecting the null hypothesis, usually set at 80%

BIAS

For details of different types of bias see Section 1A.14. Estimates of rates and risks rely on obtaining an accurate denominator figure for the population at risk. Where the denominator is not available (or is systematically inflated or reduced), then errors in epidemiological measurement occur. For example, GP practice lists often include people

who have in fact moved out of the area or registered with another practice, commonly known as *ghosts*. When such a list is used as a denominator, results may become biased.

MINIMISING BIAS

See also Section 1A.14.

Bias can be reduced in a number of ways:

- **Randomisation** with true **concealment** ensures that investigators are unable to predict or affect group allocation. **Random** allocation (in intervention studies) effectively removes selection bias.
- **Blindness**: double blindness (investigator and subject both unaware of allocation) if possible, but if no placebo is possible then ensure that the person assessing outcome is blinded to the exposure status. Blinding of participants avoids respondent bias; blinding of practitioners who are treating study participants avoids instrument bias; and blinding of observers who collect measurements or analyse results avoids observer bias.
- Collect **irrelevant factors** to check for bias between groups and to mask the hypothesis under investigation.
- Enhance the reliability/reproducibility of measurements, e.g. **repeating** measurements, and assure inter-observer agreement (instrument bias).
- Study personnel should receive **standardised training** and have their performance monitored (e.g. time the interviews to assess the degree of probing by interviewers).
- Data collection should be by **detailed written protocol** using **closed questions** and **direct observations**.
- Choose **hospitalised controls** to increase comparability (they will have similar recall of events before admission). However, be aware of the hospital having different catchment areas for different specialties.
- Choose cohorts that can be **easily followed up** long term, e.g. alumni groups.
- Choose cohorts with **above-average risk** of disease to shorten follow-up period, and therefore reduce losses to follow-up.
- Use **multiple sources** of data to assess exposure/disease status.

Bias in questionnaires can be reduced by:

- Checking for known associations
- Seeking the same information in different ways
- Checking characteristics of data collection (time taken).

Bias in intervention studies can be assessed by:

- Self-reports
- Pill counts
- Measuring biochemical parameters
- Incorporating a safe biochemical marker in the placebo that can be detected in urine.

CONFOUNDING

A confounded association is one in which the effect of the variable of interest is inextricably combined with that of another uncontrolled variable. For example, a study that showed an association between increased level of foreign travel and decreased risk of stroke might be confounded by age (because, on average, older people travel abroad less frequently and are at higher risk of stroke than younger people). This is illustrated in Figure 1A.13.1.

Confounding can be controlled either at the **measurement** stage or the **analysis** stage (see Section 1A.16).

CAUSALITY

Once an association has been established, the issue of causality may be considered. Causality is a judgement of cause and effect. It is based on valid study data *plus* a background of other evidence. To be valid, all alternative

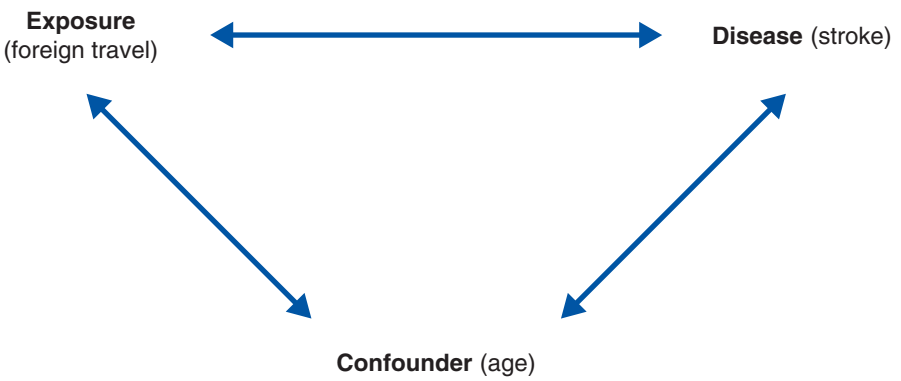


Figure 1A.13.1 Age is a confounder in the association between foreign travel and stroke

explanations for the findings (i.e. chance, bias and confounding) must be dismissed. Judgement of causality is then based on the criteria listed in Table 1A.13.1, which were first described by **Bradford Hill** (1965).

Table 1A.13.1 Bradford Hill criteria for determining causality

Strength of association	The greater the magnitude of the risk observed (positive or negative), the less likely it is to be due to confounding. The opposite is not true
Biological credibility/ plausibility	A finding is more credible if there is a known or postulated biological mechanism to explain it. However, an apparently implausible association might turn out to be a scientific discovery
Consistency of findings	Most persuasive evidence comes from several studies with different methodologies showing similar results
Temporal sequence	The exposure of interest should precede the outcome by a biologically plausible period
Dose–response	There may be a gradient of risk according to the amount of exposure. Note, however, that many biological phenomena demonstrate a threshold (rather than a linear) relationship. The relationship between alcohol and cardiovascular disease can show a 'J'-shaped curve, implying greater cardiovascular disease risk for those who drink no alcohol compared with those who drink small amounts
Specificity	If the exposure is associated with one outcome or range of outcomes, this is evidence in favour of a causal effect relationship. However, there are few such one-to-one relationships, and many exposures cause several outcomes, e.g. smoking causes a plethora of diseases
Coherence	The relationship should not conflict with the natural history of the disease
Reversibility	If the risk of the disease decreases when the exposure is removed, then this is strong evidence of causality

1A.14 BIAS

A bias is a **systematic error** that leads to a difference between the comparison groups with regard to how they are chosen, treated, measured or interpreted. This error leads to an incorrect estimate of the association between the

exposure and the risk of disease. Unlike confounding and the role of chance, the magnitude of a bias cannot be quantified.

There are five main groups of bias:

- **Misclassification**
- **Selection**
- **Measurement**
- Biases relating to **intervention studies**
- Types of bias that are associated with **screening programmes** (e.g. lead-time bias, length-time bias and volunteer bias – see Section 2C.2).

Different epidemiological study designs are particularly susceptible to the types of bias shown in Box 1A.14.1.

Box 1A.14.1

Type of study	Type of bias
All studies	Misclassification bias
Case–control	Recall bias and selection bias
Retrospective cohort	Follow-up bias and selection bias
Prospective cohort	Follow-up bias
Intervention study	Bias with comparison group, placebo, ascertainment

MISCLASSIFICATION BIAS

Misclassification will occur to some extent in all studies. It involves allocating subjects to the wrong group – either at random or in a non-random fashion.

RANDOM MISCLASSIFICATION BIAS

Also known as '*non-differential*' misclassification bias, this causes the result of the study to be diluted towards the null.

NON-RANDOM MISCLASSIFICATION BIAS

If the errors in classification are non-random (i.e. *'differential'*), then the result of the study may be biased in any direction or in none. For example, see Box 1A.14.2.

Box 1A.14.2

Example: Non-random misclassification bias among smokers

When asked, approximately 10% of smokers deny their habit. This fact can be validated through analysis of urinary cotinine. These untruthful responses can introduce a non-random misclassification bias into any study where the analysis considers the effect of smoking.

Some people have a tendency to 'yea-say', i.e. to agree with a statement rather than to disagree with it. This phenomenon may result in the misclassification of these subjects.

SELECTION BIAS

This bias occurs when the presence or absence of exposure influences either **allocation** to particular study groups or **participation** in the study. Examples of selection bias include the following.

HEALTHY WORKER EFFECT

This effect occurs when a sample drawn from a workplace is selected as representative of the general population. The general population includes those too ill to work so the workplace sample is healthier than the general population.

VOLUNTEER BIAS

This occurs because those who agree to participate in studies tend to be both healthier and more compliant than those who do not.

FOLLOW-UP BIAS

This occurs when participants who are lost to follow-up differ systematically from those who remain contactable.

MEASUREMENT BIAS

There are three types of measurement bias.

INSTRUMENT BIAS

This occurs where there are systematic inaccuracies in a test or instrument.

RECALL BIAS

It is a feature of human psychology that experiences that precede an adverse event or outcome are recalled particularly well. For this reason, participants who develop a disease will recall antecedent exposures differently from those without disease.

OBSERVER BIAS

This occurs if systematically there are differences in the ways that the exposure or outcome data are collected between the two groups. There may be inaccuracies or incompleteness in data collection that affects the groups to differing degrees. **Interviewer bias** is a type of observer bias in which there is a systematic difference in the way that the investigator collects, probes for or interprets data between groups.

BIASES IN INTERVENTION STUDIES

Intervention study designs are prone to the following types of bias.

COMPARISON GROUP

This occurs where there are systematic differences between the control group and the intervention group. Comparison group biases can be eliminated by rigorous randomisation and intention-to-treat analysis (see Section 1A.20). Differences between the intervention group and the control group are typically logged in the first table of a published study.

PLACEBO

It is a feature of human psychology that people tend to report favourable outcomes to any intervention that they receive, regardless of its physiological effect. Studies that do not employ a **placebo control** will therefore be biased.

ASCERTAINMENT

If the collection of data from one group is more accurate or complete than that from the other, then a bias will result. The longer the follow-up time, the higher the propensity to ascertainment bias. Ascertainment bias is a particular problem where the outcome being measured is subjective (e.g. degree of pain).

MINIMISING BIAS

Several measures can be taken to avoid bias (see Section 1A.9).

EVALUATING BIAS

Although bias cannot be quantified (compare confounding and chance), attempts should always be made to:

- Contemplate the **direction** of the alleged bias
- Use **internal validation**. For example, if some participants were incorrectly diagnosed initially, then, if there were no bias, these participants should have had the same exposures as those of the controls.

1A.15 CONFOUNDING AND INTERACTION

Confounding, interactions and methods for ascertainment of effect modification

See Sections 1A.13, 1A.16 and 1B.14.

CONFOUNDING

This occurs when an association arises because of differences between groups other than the exposure being studied. The observation is distorted because of the presence of a confounder, i.e. a variable that is both associated with the exposure and also independently associated with the outcome. For example, see Box 1A.15.1.

CONFOUNDERS

- A confounder must be associated with the disease *and* be independently associated with the exposure under study
- A confounder must predict disease independently of the exposure under study (including in non-exposed individuals)
- A confounder cannot simply be an intermediate step in the causal chain.

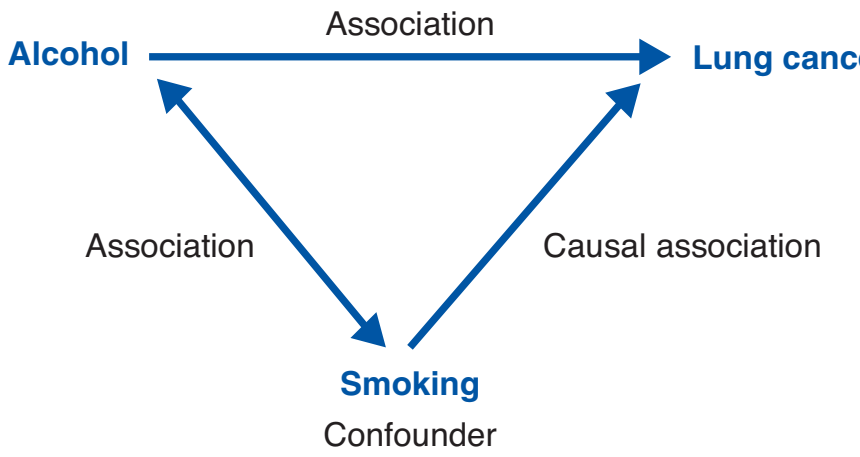
MEDIATING FACTORS

Not every factor that is associated with both the exposure and the disease is a confounding variable. A variable that is associated with both the independent and dependent variables may be an intermediate step along the causal chain. This is called a **mediating factor** and it is not a confounder. See Box 1A.15.2 for an example of a mediating factor.

Box 1A.15.1

Example: Smoking as a confounder

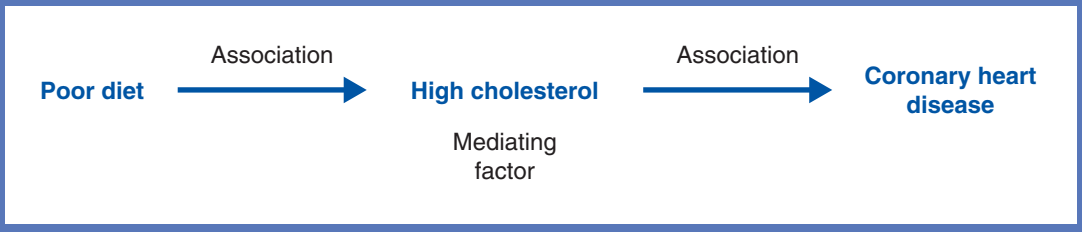
Populations with high alcohol consumption tend to have high lung cancer rates, but this is simply because people who misuse alcohol are more likely to smoke than people who do not. Smoking, in turn, is causally associated with lung cancer. Smoking therefore confounds the apparent relationship between alcohol consumption and lung cancer.



Box 1A.15.2

Example: Cholesterol as a mediating factor

There is a relationship between a poor diet and the incidence of coronary heart disease (CHD). There is also a relationship between a poor diet and a high serum cholesterol level. We also know that a high serum cholesterol level is causally associated with a higher risk of a CHD event. However, serum cholesterol is not a confounding factor here. Because a high cholesterol level may be caused by a poor diet, it is in fact on the causal pathway of the relationship between diet and CHD. It is therefore a mediating factor.



INTERACTION

An interaction occurs when the effect on the outcome of one causal factor differs according to the level of a third variable. This third variable is termed an **effect modifier.** For example, the effect of smoking on the risk of myocardial infarction is stronger in the young: age is an effect modifier in this case.

DIRECTION OF CONFOUNDING

Positive confounding makes the association appear **more pronounced** (this may be in a positive or negative direction).

Negative confounding makes the association appear **less pronounced** (i.e. diluted towards the null).

RESIDUAL CONFOUNDING

This occurs as a result of unknown confounders (i.e. those with effects that remain after the known confounders have been taken into account) or where confounders are inaccurately measured. The process of randomisation, which distributes both known and unknown confounders equally between groups, effectively eliminates all residual confounding.

1A.16 CONTROL OF CONFOUNDING

Strategies to allow/adjust for confounding in design and analysis

Confounding can be countered at the **design** stage or it can be adjusted for at the **analysis** stage.

DESIGN

At the design stage, confounding can be addressed through **restriction**, **matching** and **randomisation**. See Table 1A.16.1.

Table 1A.16.1 Techniques for dealing with confounding at the design stage

Technique	Method	Advantages	Disadvantages
Randomisation	Participants allocated to groups at random	If the sample is large enough then randomisation removes known and unknown confounders	Not always possible
Restriction	If sex and race are considered to be potential confounders, then consider only one particular group from across both factors (e.g. just black men)	Cheap	Smaller pool of potential recruits; residual confounders if restriction is insufficiently narrow; cannot assess varying levels of a factor
Matching	Nowadays only used for case–control studies: match for age, sex, race, smoking history, etc	Intuitive appeal; unique benefits of twin studies (see Section 1A.41); useful in small case series	Difficult and expensive, especially where there are multiple controls for each case; cannot explore factors that have been matched; no control for factors that have not been matched

ANALYSIS

At the analysis stage, confounding can be corrected for by means of stratification or multivariate analysis. See Table 1A.16.2.

Table 1A.16.2 Techniques for dealing with confounding at the analysis stage

Method	Details	Disadvantages
Stratification	Divide confounding variables into strata, and provide stratum-specific relative estimates (with confidence intervals) plus a weighted-average overall single estimate of the confounding effect (e.g. **Mantel–Haenszel** method)	Unable to control simultaneously for more than a few confounders because number of strata increases exponentially, so number of individuals in each stratum falls
Multivariate analysis	Mathematical models (**multiple regression**, **logistical regression**, etc)	Black box, therefore transparency is lost (a minor problem: this is usually the preferred method)

1A.17 DESCRIPTIVE AND ECOLOGICAL STUDIES

Design, applications, strengths and weaknesses of descriptive studies and ecological studies

Descriptive studies describe patterns of disease with regard to time, person and place. They include **case reports** (a description of a one-off unusual finding) and **case series** (a description of several unusual findings that are linked in some way).

Ecological studies are characterised by the unit of observation being a **group** (e.g. a population or community) rather than an individual.

Both types of study can use routinely collected data, and as a result they are relatively cheap and rapid to conduct. Both are useful for the formulation of research questions. See Table 1A.17.1.

Table 1A.17.1 Features of descriptive studies

Type	Descriptive studies (case reports/case series)	Ecological studies
Design	Astute clinician notices unusual occurrence	Describe disease pattern for an entire population with regard to another parameter Uses group level data: **aggregate** (summaries of individual data), **environmental** (measurable at individual level but easier at group level) and **global** (attributes of a place not applicable at individual level, e.g. district, legislation) Correlation coefficient (r) can range from –1 to +1
Application	Case series are particularly useful in identifying the beginning of an epidemic Hypothesis formulation	International comparisons Study of group level effects, e.g. legislation Hypothesis formulation
Strengths	Rapid reporting Low cost	Rapid results Low cost

Type	Descriptive studies (case reports/case series)	Ecological studies
Weaknesses	Case report cannot be used to test valid statistical association (may just be a coincidence) Case series difficult to test as no appropriate comparison group	Unable to control for unknown confounders Only considers average exposure so would be unable to detect a J-shaped curve Spatial autocorrelation – analysis assumes that all areas are independent but may not be Leakage of exposures through migration No information on individuals Risk of **ecological fallacy** (see below)

ECOLOGICAL FALLACY

An **ecological fallacy** is an error of logic that occurs when inferences are made regarding individuals, based on aggregate data from the population to which the individuals belong. See Box 1A.17.1.

Box 1A.17.1

The 'original' ecological fallacy

The term ecological fallacy was first used in relation to an analysis of the 1930 US census. This found that the higher the proportion of immigrants in a state, the higher its average literacy. Analysis at the individual level showed, however, that immigrants were less literate than native citizens but that they tended to settle in states where the native population was more literate. It would have been an ecological fallacy to conclude from the original analysis that immigrants were more literate people.

1A.18 SMALL AREA ANALYSIS

Analysis of health and disease in small areas

The prevalence rates for a particular disease at a small area level may be markedly different from the prevalence rates at the regional or national levels. Analysing data at these larger geographical areas would therefore mask pockets of high risk.

Similarly, it may not be valid to extrapolate health survey data (such as the Health Survey for England) down to a local level. There is, therefore, a case for high-quality research (using valid survey methods and instruments) to be commissioned at a local level, although the cost implications may be prohibitive.

See Section 3A for details of small area boundaries in the UK.

1A.19 STUDY DESIGN

Design applications, strengths and weaknesses of cross-sectional, analytical studies and intervention studies (including randomised controlled trials)

See Section 1B.9.

Study design is key to epidemiological research. An epidemiological question can often be answered using several designs: practical issues (cost, timescale, etc.) then determine which should be used.

For all analytical studies, the first step is to state the research **hypothesis** (for the statistical testing of hypotheses). This involves stating the **exposure(s)** of interest, the **outcome(s)** of interest and any **possible confounders**.

The main types of study design are:

- Cross-sectional
- Case–control
- Cohort
- Intervention.

CROSS-SECTIONAL STUDIES

In this type of study, all of the variables (exposures and outcomes) are measured at the same time. Cross-sectional studies can be descriptive, analytical or ecological. See Box 1A.19.1.

Box 1A.19.1

Descriptive	Description of the point prevalence of a disease in a population linked to health service usage data
Analytical	Comparison of several exposures with the outcome of interest
Ecological*	Neither exposure nor outcome is measured at the individual level

**Note that ecological studies need not necessarily be cross-sectional; they can also be longitudinal.*

See Table 1A.19.1 for the features of cross-sectional studies. An example of a cross-sectional ecological study is one that compared HIV prevalence and rates of male circumcision in several African countries. The study found that HIV prevalence was lower in countries where male circumcision was commoner. Interpretation of this study needs to take account of ecological fallacy (see Section 1A.17) and the potential for confounding (e.g. religious practice may be associated with higher circumcision rates *and* with lower sexual promiscuity).

Table 1A.19.1 Features of cross-sectional studies

Design	Determine the **simultaneous** prevalence of exposure and disease Measure prevalence of various diseases, characteristics and health-care usage and make comparisons
Sampling	Sampling needs to be representative of the population under study. A random sample is preferable and the sample needs to be sufficiently large
Application	Hypothesis formulation If the exposure is fixed (e.g. variables present at birth), then a cross-sectional survey is analytical and can be used to test hypotheses
Analysis	**Disease frequency:** odds or prevalence **Measure of effect:** prevalence ratio, prevalence difference, odds ratio
Strengths	Can study multiple exposures and outcomes Rapid and cheap Useful for rare diseases
Weaknesses	Because they are measuring prevalence, not incidence, findings cannot differentiate the determinants of aetiology and survival

CASE–CONTROL STUDIES

In this type of study, individuals with the outcome of interest (cases) are matched with individuals who do not have the outcome of interest (controls). See Table 1A.19.2. For example, in the study of sudden infant death syndrome (SIDS), the characteristics (e.g. sleeping position, type of mattress) of a group of children who had died from SIDS were compared with a matched group of children who had not.

Table 1A.19.2 Features of case–control studies

Design	Hypothesis: state exposure(s) and outcome Choose subjects who do and do not have the **disease** of interest and compare with respect to the exposure of interest Selection of controls is the most critical issue (see Section 1A.7) Ideal ratio of cases to controls is 1:1 but, if cases are limited, then can increase number of controls disproportionately up to 1:4 (beyond this there is little gain in power)
Sampling	Selection of cases should be representative of all cases in the population. Controls should be selected from the same population as the controls and may be matched for certain characteristics *not* being tested. In all cases, the selection of cases and controls must be independent of their exposure to the putative risk factors of interest
Application	Case–control studies can be **retrospective** (all cases of disease have been diagnosed when study starts) or **prospective** (new cases identified during lifetime of the study)
Analysis	First compare cases and controls for other baseline differences **Disease frequency:** not possible to determine from this design; cases chosen precisely because they have disease **Measure of effect:** odds ratio
Strengths	Rapid and cheap Ideal for **rare diseases/outcomes** Useful for diseases with long latent periods Can simultaneously examine a large number of potential exposures
Weaknesses	Selection bias (exposure and disease have already occurred) Temporal relationships may be difficult to establish Recall bias (of information on exposure and disease) Poor for rare exposures (unless high attributable risk percentage) Cannot compare incidence rates (unless population based) Misclassification of exposure/disease status (random ➔ null; non-random ➔ any direction) Temptation of data fishing: if multiple hypotheses are tested, then on average 1 in 20 will be significant at the p <0.05 level just because of chance (see *Bonferroni* correction, Section 1B.11)

COHORT STUDIES

In a cohort study, a group of individuals is selected who do *not* initially have the outcome of interest. A range of exposures is quantified for cohort members and at the end of the study those people who have developed the outcome of interest are compared (according to the exposure of interest) with those who have not. In retrospective cohort studies, both the beginning and end of the time period of interest are in the past. For example, the Framingham cohort studies followed residents of a town in Massachusetts. The outcomes of interest were cardiovascular endpoints (e.g. myocardial infarction) and the exposures of interest included serum cholesterol and blood pressure. See Table 1A.19.3 for the features of cohort studies.

Table 1A.19.3 Features of cohort studies

Design	Hypothesis: state exposure and outcome(s) Choose subjects on basis of **exposure** (all subjects should be disease free at the start) Follow cohort to study **temporal relationships,** so consider ease of follow-up when selecting cohort (e.g. alumni)
Sampling	Cohort members should be representative of the population of interest. If a workplace cohort is chosen, then cohort members are likely to be healthier than the general population (the 'healthy worker effect', see Section 1A.14)
Application	Able to measure incidence directly in both groups
Analysis	**Disease frequency:** rate, risk, odds, mean or median **Measure of effect:** rate/risk/odds ratio (relative), rate/risk/odds difference (absolute). Other – vaccine efficacy, difference in mean/median Must also **compare** groups to ensure **similarity** of potential confounders Can estimate effect of loss to follow-up by comparing the two extreme situations (i.e. all those lost to follow-up did develop the disease, and then that none of them did)
Strengths	Able to follow **temporal** relationships Well suited to **rare exposures** **Multiple effects** of a single exposure Minimise selection bias (prospective cohort studies) Useful for diseases with **long latency** periods (retrospective cohort studies)
Weaknesses	Time-consuming Expensive Risk of loss to follow-up (➔ poor validity [power and bias], especially if loss to follow-up >30% or disproportionate between two groups) Inefficient for rare diseases Records may be inadequate for ascertainment (retrospective cohort studies) Healthy worker effect (see Section 1A.14)

INTERVENTION STUDIES (INCLUDING RANDOMISED CONTROLLED TRIALS)

In an intervention study, the epidemiologist is able to allocate the exposure of interest across the study population. See Table 1A.19.4. For example, in the 4S Study (1994), a group of Scandinavian volunteers was randomised to receive either 20 mg of simvastatin daily or placebo. The intervention and control groups were compared according to numbers of deaths and numbers of major cardiovascular events.

Table 1A.19.4 Features of intervention studies

Design	These are epidemiological experiments in which the investigators **allocate** the exposure. The intervention is generally allocated at **random** (see Section 1A.26 for methods), but an alternative is to use **systematic** allocation **Factorial design** This allows **simultaneous** testing of two or more hypotheses at **marginal** increase in **cost** Assign to A or B and independently to X or Y There should be no known interaction between the two hypotheses, nor any effect on participant eligibility, loss to follow-up or compliance If there is actually an interaction then: • Estimate of the treatment effects will tend towards the null • Such interaction can be detected during analysis **Stopping rules** An independent group must monitor interim results to ensure the welfare of participants If there is extreme benefit or extreme harm the trial should be stopped early, but only if this is definitely not a temporary, random fluctuation in results Procedures are also needed for immediate unblinding in the event of an isolated serious complication Extremely high significance (extremely low *p* value) is needed to justify early termination of trial It is controversial whether different stringencies in significance are needed to stop a trial early because of beneficial and harmful effects **Sample size** Must be considered during planning stage so that the sample size has sufficient power to detect differences between the groups Power depends on: • Sample size • Total number of endpoints • Difference in compliance between two groups • Increased number of endpoints by selecting a **high-risk population** (notes '**healthy volunteer effect**' so number of endpoints will be lower than expected) or increased duration of **follow-up** Remember possibility of **secular changes,** i.e. whole population becoming healthier over time so fewer endpoints

Table *contd overleaf*

Table 1A.19.4 *contd*

Design (*contd*)	**Compliance** Non-compliance ➔ Two groups more similar ➔ Result tends towards the null ➔ Trial less able to detect a true effect Can improve compliance by using **run-in period** pre-randomisation during which all potential participants receive treatment/placebo to determine acceptability to patient
Sampling	Study population needs to be representative of the reference population (i.e. the population to which the results of the trial are to be extrapolated)
Application	Can investigate **therapeutic** or **preventive** exposures, at the level of either the **individual** or the **population** (e.g. fluoridation)
Analysis	**Disease frequency:** rate, risk, odds, mean or median **Measure of effect:** rate/risk/odds ratio (relative), rate/risk/odds difference (absolute). Other – vaccine efficacy, difference in mean/median Start with a table comparing the two groups for baseline characteristics Consider the **placebo effect** Analyse by **intention to treat** (see Section 1A.20)
Strengths	Evidence provided can be of extremely high quality If sample large enough, then **validity** can be guaranteed Blinding minimises observation bias (where blinding impossible then independent examiners should assess endpoints)
Weaknesses	Generalisability: Is the **study population** similar enough to the reference population for the results to be applicable? If inclusion/exclusion criteria too strict, then it may not be Is the **intervention** comparable to the treatment that would be received outside a trial (with regard to clinicians' interest and attentiveness)? **Efficacy** is the effect under ideal conditions; **effectiveness** is the real-life effect There must be sufficient doubt (**equipoise**) about the alternatives Study will be not be feasible if the treatment under investigation is already too widely accepted Study will be unethical if existing evidence suggests harm from giving or withholding the treatment Potential for enormous **cost** Potential **biases** from: • **Loss to follow-up** of many subjects (more likely with long follow-up periods) • **Unequal follow-up** (accuracy or completeness) between two groups • **Observation bias** (unlikely with mortality but more likely with cause of death – minimised by double-blinding) • **Placebo effect** (response to any therapy regardless of physiological effect: true effect of intervention is Percentage effect in treatment group – Percentage effect due to placebo)

1A.20 INTENTION-TO-TREAT ANALYSIS

While it may be intuitive to compare the groups in an intervention study according to the actual interventions that they received, such **analysis-by-treatment-received** is fundamentally flawed for the three reasons given below. Instead the analysis should always be performed based on the a priori group to which each participant was originally assigned. This is known as **intention-to-treat analysis** and it ignores the reality of whether the participant actually received the intervention or placebo assigned to that group.

ADVANTAGES

The advantages of intention-to-treat analysis over analysis by treatment received are:

- The actual research question being asked concerns the effect of *offering* this treatment to patients in the real world, as opposed to the actual physiological effect of their receiving it.
- The balance of unknown confounders achieved at randomisation is impossible to recreate once non-compliers are removed from the two groups (hence the maxim, *'once randomised, always analysed'*).
- Those who switched from their allotted groups may differ systematically from those who remained (e.g. patients who were too sick to cope with side effects may have tended to transfer to the placebo group).

1A.21 CLUSTERED DATA

Clustered data – effect on sample size and approaches to analysis

Clustered data are not fully independent of each other. Rather, they are linked according to a particular characteristic (e.g. time, place or person). Standard statistical techniques depend on the assumption that observations are independent of each other. This is not the case when data are clustered: observations in the same cluster will be more similar to each other than if they were independent. See Box 1A.21.1 for an example of a clustered study.

CLUSTERED DATA

Common types of clustered data include:

- Multiple observations on the same subject at the same time
- Repeated observations on the same subject over time
- Cluster randomised trials
- Clustered sampling surveys.

Box 1A.21.1

Example: Ten Towns Study

This project, which started in 1990, compares 10 towns in England and Wales. Five of the towns had high levels of cardiovascular disease and five had low levels. In one study within the project, researchers surveyed school children to determine, *'Whether markers of nutrition, cardiovascular health and type 2 diabetes differ between school pupils who eat school dinners and those whose school day meal is provided from home'.*

In order to take account of clustering, the researchers assigned the town of residence as a **fixed effect** and assigned the school attended as a **random effect** (see below). Because the analysis used both fixed and random effects, it is termed a 'mixed model'.

Adapted from Ten Towns Heart Health Study: www.tentowns.ac.uk; Whincup et al (2005).

In each of these circumstances, the effect of clustering must be taken into account at the statistical analysis stage, otherwise associations that do not truly exist are more likely to be detected.

SAMPLE SIZE

Clustered trials and surveys require an increased sample size to compensate for the tendency of individuals within each cluster to be more similar to each other than to individuals in other clusters.

The amount by which the sample size needs to be increased is called the **design effect** and this depends on the number of individuals per cluster and the **intraclass correlation** coefficient, i.e. the ratio of intercluster variance to total variance.

ANALYSIS

At the analysis stage there are several ways of accounting for the effect of clustering:

- Calculate **summary statistics** for each cluster and then analyse these using standard techniques
- Calculate **robust standard errors** that account for clustering
- Use more sophisticated techniques (such as **generalised estimating equations** or **random effects models**).

RANDOM AND FIXED EFFECTS

These are statistical models that can be employed within ANOVA* or regression analysis of clustered data.

- **Random effects** models explicitly model the similarities between individuals in the same cluster, i.e. a random sample of all possible values of that variable
- In contrast, a **fixed effects** model refers to assumptions about the independent variable (e.g. that it can only take the values high, medium or low). With a fixed effects model, the conclusions will be restricted to these three values.

**ANOVA (analysis of variance) is a collection of statistical procedures that compares means by dividing the overall observed variance into different parts*

1A.22 NUMBERS NEEDED TO TREAT

Numbers needed to treat – calculation, interpretation, advantages and disadvantages

The numbers needed to treat (NNT) is a useful statistic that indicates how many patients would need to be treated with a particular intervention to reduce by one the expected number of a defined outcome. It is the reciprocal of the absolute risk reduction, and is particularly useful since it takes into account the frequency of the outcome – and thus reflects the public health impact of the intervention (e.g. impact of offering low-dose aspirin in preventing cardiovascular disease).

Analogous measures can be used to compare screening programmes* and harmful exposures (see Box 1A.22.1)†.

Box 1A.22.1

Calculation	$\text{NNT} = \dfrac{1}{\text{Absolute risk reduction}}$
Interpretation	The NNT answers the question, *'How many patients would I need to treat in order for one extra patient to benefit?'* *An analogous measure in screening studies called the **number needed to screen** †If the treatment or exposure is harmful (i.e. the result is a negative number), then the minus sign is omitted and the measure is renamed the **number needed to harm**

Advantages	More intuitive expression of absolute risk (itself expressed as a percentage)
Disadvantages	It is tempting to compare NNTs for different therapies but this is justified only when the baseline risks of the disease are similar. A drug that reduced mortality rate from 90% to 80% has the same NNT as one that reduces mortality rate from 40% to 30% but the former would have greater impact Caution is needed when calculating the NNT from meta-analysis

See Box 1A.22.2 for an example of how NNT is calculated.

Box 1A.22.2

Example: The 4S Study

In the 4S Study, during a 5.4-year period, 256 of 2223 patients (11.5%) in the placebo group died, compared with 182 of 2221 patients (8.2%) in the group who were treated with simvastatin.

Absolute risk reduction = 0.115 – 0.082 = 0.033

$$\text{NNT} = \frac{1}{0.033} = 30 \text{ patients}$$

1A.23 TIME TREND ANALYSIS, TIME SERIES DESIGNS

Time series analysis considers the situation where both the exposure and the outcome are measured over time. If there are secular increases or decreases (or seasonal variations) in both variables, then the variables will be correlated with time. This is known as serial correlation, and time will therefore be a confounding factor in the analysis. For example, both rainfall and hospital admissions are higher in winter. In answering the question whether high rainfall is associated with increased hospital admissions, this seasonal correlation would first need to be accounted for.

Difficulties in attributing effect to the exposure can be due to the causes shown in Box 1A.23.1.

TIME SERIES DESIGNS

Figure 1A.23.1 shows four different time series designs. In each of the four graphs, the horizontal axis represents time and the vertical axis represents the magnitude of the effect observed.

Box 1A.23.1

Secular changes	For example, changes in age structure, classifications of disease, diagnostic techniques
Concurrent interventions/ exposures	For example, a decrease in deaths from heart disease might be observed when smoking prevalence falls and more effective therapies (e.g. *statins*) are used
Latency period	Long time for exposure to manifest its effect (e.g. smoking and cancer)
Diffuse exposure	Exposure spread out over months or years

ANALYSIS TO ESTIMATE EFFECT SIZE

1. Smooth out 'noise' (peaks and troughs) in curves using moving averages
2. Segmented regression – analysis of each separate part of an interrupted time series.

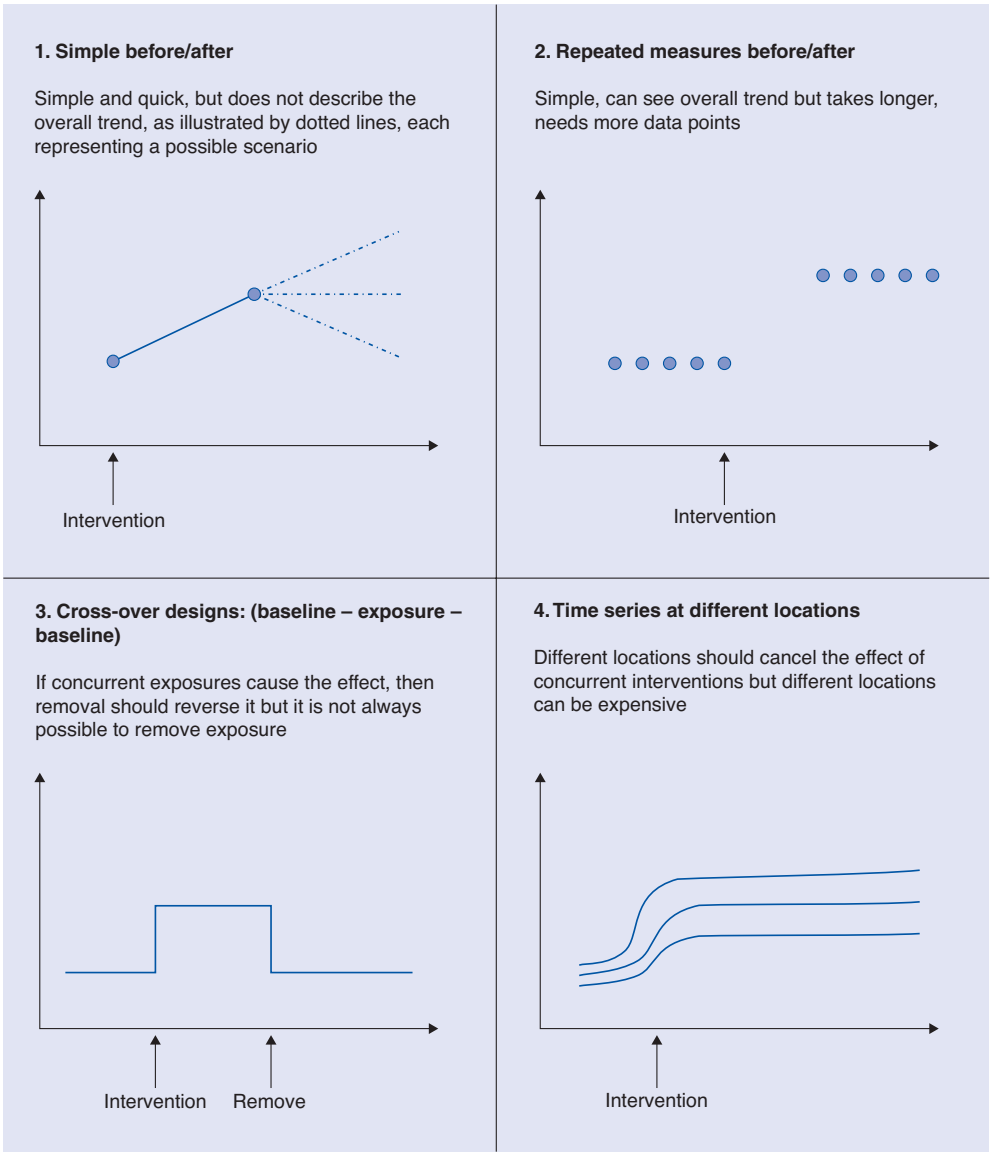


Figure 1A.23.1 Time series designs

1A.24 NESTED CASE–CONTROL STUDIES

Sometimes a case–control study is enclosed ('nested') within a cohort study. In this instance, both the cases and the controls are taken from within the population of a cohort study. Nested case–control studies can be useful when it would be **too expensive** or otherwise unfeasible to perform laboratory tests on the entire cohort.

NESTED STUDY DESIGN

A nested study might be conducted as follows:

1. Take baseline samples for all entrants into cohort but do not analyse (e.g. blood samples are frozen but not tested)
2. Run cohort study
3. Identify **cases** as members of the cohort who develop the disease
4. Choose **controls** as matched disease-free members of the cohort
5. Test frozen blood samples of cases and controls (and discard samples from the rest of the cohort).

Advantages and disadvantages of nested studies are shown in Box 1A.24.1.

Box 1A.24.1

Advantages	Disadvantages
Cost-effective Avoid recall bias by using data collected before the onset of disease Avoid selection bias by drawing cases and controls from the same cohort	Non-diseased cohort members (from whom the controls are selected) may not be fully representative of the original cohort, due to death or losses to follow-up

1A.25 METHODS OF SAMPLING FROM A POPULATION

Since investigating the entire population (by a census) is costly, sampling is a more cost-effective and convenient means of collecting information. Sampling methods produce either **probability samples** (which allow calculation of sampling error and hence allow inferences to be made regarding the population) or **non-probability samples** (which do not permit such inferences).

PROBABILITY SAMPLING

A sampling frame is required for all studies that aim to make an inference about the population (i.e. a complete list of the population from which the sample is to be drawn). Different methods of probability sampling are shown in Table 1A.25.1.

Table 1A.25.1 Methods of probability sampling

Type	Method	Advantages	Disadvantages
Random	Start with a sampling frame Sampling frames include postcode address files, electoral register, GP practice list Draw random sample from sampling frame	Purest form of probability sampling Allows calculation of sampling error	Relatively inconvenient in practice Inefficient for rare outcomes Sample frame not always easily available
Systematic	Start with sampling frame Calculate sampling interval (n): $n = \frac{\text{(number in population)}}{\text{(number in sample)}}$ Draw every nth person from the sampling frame	More convenient than random sampling Allows calculation of sampling error	Potential for bias if there are underlying patterns to the sampling frame

Table *contd overleaf*

Table 1A.25.1 *contd*

Type	Method	Advantages	Disadvantages
Stratified	Assign members of the population into relatively homogeneous subgroups ('strata') before sampling Draw random sample of subjects from within each stratum	Improves accuracy of estimation Efficient Allows calculation of sampling error	Requires accurate information about the population Choice of relevant stratification variables can be difficult Unhelpful if there are no homogeneous subgroups
Cluster	Used when there are 'natural' clusters in the population, e.g. GP practices within a borough A random sampling technique is used to choose which clusters to include in the study In single-stage cluster sampling, all the elements from each of the selected clusters are used (e.g. all patients in selected practices). In two-stage cluster sampling, elements from each of the selected clusters are selected at random (sample of patients from within the selected practices)	Convenient for fieldwork Cost-efficient Allows calculation of sampling error	Increased sampling error

NON-PROBABILITY SAMPLING

In general, for non-probability sampling, the sampling frame is not known. Different methods of non-probability sampling are shown in Table 1A.25.2.

Table 1A.25.2 Methods of non-probability sampling

Type	Method	Advantages	Disadvantages
Convenience	Subjects chosen on basis of being readily available	Useful for preliminary research as it is extremely efficient	Sampling error cannot be calculated Volunteer bias
Judgement	Subjects chosen purposively on the basis of having particular features, such as controls required for epidemiological studies	Useful for rare characteristics Useful for qualitative research	Sampling error cannot be calculated Volunteer bias

Type	Method	Advantages	Disadvantages
Quota	Select demographic characteristics of interest (e.g. age, sex, ethnicity) Select subjects to represent the proportional distribution of these characteristics within the population	Representative with regard to known characteristics	Potentially unrepresentative in terms of other characteristics Sampling error cannot be calculated Volunteer bias
Snowball	Ask subjects to recommend acquaintances who meet the sample criteria	Very cost-efficient Useful where no sample frame exists Enables researcher to reach groups that are otherwise hard to reach	Sampling error cannot be calculated Volunteer bias

CHOOSING CONTROLS FOR EPIDEMIOLOGICAL STUDIES

The selection of controls is often the most critical issue for an epidemiological study, particularly in case–control studies which are very vulnerable to selection bias. The choice of this comparison group will vary depending on the situation. Two ways of choosing controls are shown in Table 1A.25.3. Note that controls are the people who **would have been identified** as cases had they developed the disease: they are **not** the entire non-diseased population.

Table 1A.25.3 Choice of controls in case–control studies

Type of control	Sampling strategy	Advantages	Disadvantages
Controls within the health-care system	Purposive: through registers, hospital records, referral data	Easily identified By being in hospital are (like the diseased cases) more aware of antecedent events (i.e. more similar to diseased cases, therefore less recall bias) Same hospital, therefore same influences on choosing that hospital More cooperative than healthy people	Ill, therefore different from healthy population (e.g. more likely to smoke, use oral contraceptive pill and drink heavily) A given hospital may have secondary and tertiary activities, so patient selection differs Different specialties within the same hospital may have different catchment areas
Population controls	Random sampling or Purposive: friends/relatives/ neighbours of cases (may be more motivated to participate)	Healthy population	General public does not recall antecedent events as well as those who are ill Less motivated to participate Healthier people less likely to be at home during the daytime Costly and time-consuming

1A.26 METHODS OF ALLOCATION IN INTERVENTION STUDIES

Interventional studies differ from observational studies in that the epidemiologist allocates the exposures of interest. The process of allocation is therefore of pivotal importance. Ideally, subjects will be allocated in such a way that the two groups are identical, meaning that the only difference between the two groups will be the intervention under investigation.

So as to minimise allocation bias, subjects should be allocated to their groups only *after* ensuring that they are **eligible**, have **consented** to participate and have been fully **enrolled**.

Allocation may be systematic, volunteered or at random.

SYSTEMATIC ALLOCATION

Allocation is determined in advance (e.g. it is decided that recruits will be allocated alternately into groups, or that all recruits presenting on alternate days will be allocated to the same alternate group). There is much potential for **selection bias** because:

- The order is predictable so the recruiter may interfere
- There may be underlying patterns to the order in which recruits present.

VOLUNTEER ALLOCATION

Allocation is determined on the basis of which recruits volunteer to participate. This is highly unsatisfactory because of extreme **selection bias.**

RANDOMISATION

The unique advantage of random allocation is that if the groups are large enough then, on average, they will be similar with respect to all variables. This includes all confounders – both known and unknown.

If randomisation is **strict** (computer-generated or random table) and **concealed**, then selection bias is avoided. Note the difference between allocation concealment and blinding/masking: it is not always possible to **mask** the intervention from participants or assessors (e.g. surgery versus drugs), but it should always be possible for the allocation method to be **concealed** so that the person recruiting a patient does not know to which arm of the study they will be allocated.

There are several types of randomisation that may be used for allocating recruits. See Table 1A.26.1.

Table 1A.26.1 Types of randomisation

Type of randomisation	Description
Unrestricted	Participants allocated purely on the basis of chance Potential for **unequal group sizes** – insignificant in large studies
Blocked	This is a technique for obtaining equal group sizes. First the ratio of participants between allocation groups is set (e.g. 1:1). A block size is then chosen (e.g. 4). All permutations of assignments that meet the ratio are then listed for that block size (with these examples the blocks would be AABB, BBAA, ABAB, BABA). Finally, the allocation sequence is generated by random selection of blocks from this list. The block size can be varied periodically to make prediction of the allocation sequence even more difficult

Type of randomisation	Description
Stratified	Recruits are subdivided into strata (e.g. according to ethnicity or gender) and individuals within each stratum are randomised
Cluster	Groups (rather than individuals) are randomised to receive the different interventions. This needs to be taken into consideration at the analysis stage (see Section 1A.21)
Matched pair	Individuals or groups are first matched according to baseline data – matching them on as many variables as possible. The intervention is then randomly allocated to one member of the pair, with the other member of the pair receiving the control
Stepped wedge	The population is divided into groups and then the intervention is progressively introduced, in random order, across the groups until every group is receiving it. Used when other allocation methods would be unfeasible because of widespread belief that the intervention is beneficial

1A.27 RECORDING SURVEY DATA

Design of documentation for recording survey data

Note that survey/questionnaire data are **quantitative.** Either the **participant** or the **researcher** can record survey data. In both cases, the design of the survey documentation should ensure that the data recorded is:

- Complete
- Accurate
- Understandable and legible
- Stored in a format suitable for analysis
- Secure to protect respondents' **confidentiality**.

SURVEY DOCUMENTATION

Documentation of data can be:

- **Hard copy**: questionnaires for postal surveys for participants to complete or questionnaires plus prompts for researchers to complete during face-to-face or telephone surveys
- **Online**: allows documentation to be tailored to responses, can provide prompts when invalid/incomplete/no responses received when required.

Those recording survey data need to know:

- **Where** to record information
- **Type** of answer required
- **Directions** for completing the survey (if items are not relevant, respondents need to know which question to answer next, e.g. 'NO ➔ move to question 5').

Researchers recording survey data from participants also need prompts/clarifications with information on:

- **Screening**: to ensure that appropriate participant is interviewed
- Directions if there is **no response** to a particular item.

Documentation layout can help by:

- **Appropriately sized** spaces to indicate where responses are needed and in what form
- Providing **clear navigation** to guide participants through the survey.

1A.28 CONSTRUCTION OF VALID QUESTIONNAIRES

Validity is the extent to which a tool explores what it is purported to measure (i.e. provides the 'true' answer and minimises bias). A questionnaire's validity will depend on the range of factors listed in Table 1A.28.1.

Table 1A.28.1 Characteristics of valid questionnaires

Characteristic	Implications	Technique to maximise validity
Sample selected	Unrepresentative sample may lead to bias	Obtain a *sampling frame* (exhaustive list of all possible experimental units) Random selection of sample Stratification of sample
Response rate achieved	Uneven or very low response rate is potentially subject to bias	Advance warning letters Incentives for responding Data method: face to face or telephone better than postal (though more costly)
Content of questionnaire (content validity and construct validity)*	Questions should be chosen carefully to ensure that the questionnaire addresses the research question For psychometric questionnaires: conceptual model needed first to define dimensions of the construct to be measured	Researching content through: • Existing tools • Literature review • Qualitative interviews • Expert opinion
Quality of the questions asked	Participants need to understand what the questions asked require and be prepared to give this information honestly	Generate clear and non-judgemental questions: pilot and test on potential participants and revise. Data collection to suit the subject: face to face better to explain questions; postal may be better for privacy

**Note the difference between psychometric questionnaires (analysed by a summative score) and survey tools (analysed by individual question).*

Other ways to maximise validity are to:

- Use an **existing tool** if an appropriate one exists
- Maximise reliability by increasing the **sample size**
- **Triangulate** with other sources of evidence (e.g. observation).

1A.29 VALIDATING OBSERVATIONAL TECHNIQUES

Methods for validating observational techniques

Observational studies can be improved by using the techniques described below.

OBSERVATIONAL TECHNIQUES

- Participant observation, e.g. nurse studying her own patients

- Non-participant observation, e.g. researcher visiting a ward to observe effects of nurse care on patients (see also Section 1D).

ENHANCING VALIDITY

- Sampling strategy: appropriate selection of participants and settings for making observations
- Considering reflexivity: the observers' effect on what they are observing
- Enhancing reliability: recording observations using observers trained comprehensively and systematically, using a checklist to record observations.

VALIDATE OBSERVATIONAL TECHNIQUES THROUGH CROSS CHECKING

- **Repeated observations** by the same, or other, observers in the same setting
- Recorded observations (e.g. **videotape**)
- **Triangulation** with other research methods.

1A.30 STUDIES OF DISEASE PROGNOSIS

The prognosis is the expected course and outcome of a disease. Studies of prognosis estimate the frequency with which different outcomes can be expected to occur. Prognostic factors are patient characteristics that guide the prediction of outcomes. Cohort studies are used for analysing prognostic factors (e.g. the Framingham study* investigated cardiovascular prognostic factors). Disease registers (e.g. regional cancer skin cancer register) can also be used to analyse prognosis.

*See the Framingham Heart Study website www.framinghamheartstudy.org.

SURVIVAL ANALYSIS

Ideally, all events would be recorded for those patients in the cohort or on the register. In practice the information available may only be that no event had occurred by a certain date. This is known as **censoring** (discussed on page 8, Section 1A.3 in more detail) and it includes the components shown in Box 1A.30.1.

Box 1A.30.1

Fixed censoring	End of study
Loss to follow-up	Moving away or dropping out
Competing event	Death by other cause

Censoring should be assumed to be **non-informative**, i.e. that the distribution of event times is identical between censored and non-censored subjects. Also, it should be assumed that the survival probability is the same for those recruited early or late into the study.

Fixed censoring does tend to be non-informative, but loss to follow-up may be informative (e.g. patients may withdraw due to harmful effects or from being too ill). Informative censoring can bias results.

Survival analysis is generally analysed by **survival models**:

- Time to death/time to recurrence, etc
- Display by Kaplan–Meier chart
- Log rank test
- Cox's proportional hazards model for multivariate analysis.

Its uses include:

- Calculating life-expectancy
- Constructing and comparing survival curves
- Comparing survival between groups.

1A.31 STATISTICAL ANALYSIS OF EPIDEMIOLOGICAL STUDIES

Appropriate use of statistical methods in the analysis and interpretation of epidemiological studies, including life-table analysis

CHOICE OF STATISTICAL METHOD

- What is the **hypothesis** being tested?
- Are the data **independent**? Or are they matched, clustered or correlated?
- What are the **input** variables (e.g. intervention or control)?
- What are the **output** variables (e.g. measure of morbidity)?

HYPOTHESIS TESTING

The null hypothesis (H_o) states that there is no association between exposure and disease.

Assuming that the null hypothesis is true, the *p* value is the possibility of obtaining a result at least as extreme as that observed due to chance alone. Usually a two-sided *p* value is used (i.e. probability of observing such a large difference, without specifying the direction of the difference). A one-sided *p* value is used to increase the precision of an estimate where there is a strong prior belief. For example, when comparing radical mastectomy against lumpectomy there is a strong prior belief that the more radical procedure will be at least as curative as the limited procedure and the question is simply whether lumpectomy is as effective as mastectomy.

CHOICE OF OUTCOME MEASURE

The choice of statistical method and outcome measure will depend on the nature of the study. See Table 1A.31.1

CHOICE OF STATISTICAL TEST

See Section 1B.

1A.32 EPIDEMIC THEORY

Effective and basic reproduction numbers, epidemic thresholds

Techniques for analysing infectious disease data include:

- Construction and use of epidemic curves, generation of numbers
- Exceptional reporting
- Identification of significant clusters.

EPIDEMIC THEORY

An epidemic is the occurrence of a number of cases of a disease that exceeds the number of cases normally expected for that disease in that area at that time.

Compartmental epidemic models allocate individuals to specific subgroups. See Box 1A.32.1.

Table 1A.31.1 Outcome measures

Study type	Outcome measure	
Correlation study	Correlation coefficient (r)	See Section 1B.14 (linear regression and correlation coefficients)
Case–control study	Odds ratio	$\text{Odds ratio} = \frac{\text{Odds (cases)}}{\text{Odds (controls)}}$ Where: $\text{Odds} = \frac{\text{Proportion exposed}}{\text{Proportion unexposed}}$ Confidence intervals can be constructed using the natural log of standard error of the odds ratio (lnOR). Note that the confidence interval is not symmetrical around the odds ratio. If the confidence interval includes 1, then there is no significant difference between the groups
Cohort study	Risk ratio	$\text{Risk ratio} = \frac{\text{Risk (exposed)}}{\text{Risk (unexposed)}}$ Where: $\text{Risk} = \frac{\text{Number of cases}}{\text{Number at risk to start}}$
	Rate ratio	$\text{Rate ratio} = \frac{\text{Rate (exposed)}}{\text{Rate (unexposed)}}$ Where: $\text{Rate} = \frac{\text{Number of cases}}{\text{Number of person-years}}$
	SMR	$\text{SMR} = \frac{\text{Observed deaths}}{\text{Expected deaths}}$
Intervention study	Risk ratio, rate ratio, attributable risk or population attributable risk	
Life-table analysis (see Sections 1B.15 and 1B.16)	Survival probability Confidence intervals Life-expectancy Proportional hazards	Cumulative chance of death in a given time period 95% confidence intervals calculated using a formula by Kalbfleisch and Prentice Time after which half of a given cohort have died $= 0.5 + \Sigma$ (Length of survival $\times$ Cumulative chance of survival) A statistical method for comparing survival rates in different groups. If there is no difference between groups, then the ratio of hazards between the groups is constant over time (even though the underlying hazards will change) and the logged cumulative hazard curves will be parallel

Box 1A.32.1

Passively immune	Susceptible	Exposed	Infective	Recovered
M	S	E	I	R

Over time, individuals will move from one compartment to another, as illustrated by the diagram in Figure 1A.32.1 in which compartments are represented by boxes and the transitions between compartments are shown as arrows.

Different models are based on the flow patterns between these compartments, e.g. MSEIR, MSEIRS, SEIR or SIR. Many diseases (e.g. measles) can be modelled accurately using SIR. The number of individuals in any compartment will vary over time, and hence is expressed as a function of time. For example, the number of susceptible individuals at time t is expressed as $S(t)$. The different rates of transition are applied to the relevant arrows in the model.

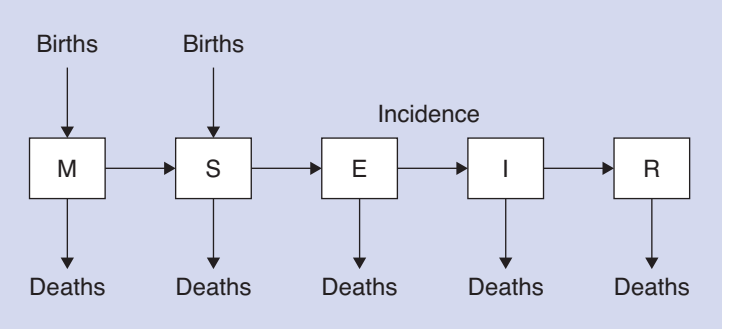


Figure 1A.32.1 A compartmental epidemic model of a disease. *Reproduced from Hethcote (2000) with permission from the Society for Industrial and Applied Mathematics*

The **basic reproduction number** (R_0) is defined as the average number of secondary infections produced when one infected individual is introduced into a host population where everyone is susceptible.

For many epidemiology models, an infection will take hold in a fully susceptible population only if $R_0 > 1$. Thus, the basic reproduction number R_0 is often considered as the threshold quantity that determines when an infection can invade and persist in a new host population.

The **epidemic threshold** is a critical value for the fraction of the population who are susceptible. Below this value an epidemic outbreak will not occur. The epidemic threshold is determined by the attack rate during the major outbreaks.

Epidemic curves are graphic representations of the number of new cases by date of onset. They are used to display an outbreak's magnitude and time trend. Features of the curve are listed in Table 1A.32.1.

Uses of an epidemic curve include:

- Determining **location** within the course of the epidemic, and possibly **projecting** its future course
- Estimating the probable **time of exposure** (if the disease and its usual incubation period are known) so that the investigation can be focused on that specific time period
- Identifying outliers (all of whom are worthy of investigation since their unusual exposure may give clues of the source)
- Inferences about the epidemic **pattern** can be drawn, e.g. whether it is an outbreak resulting from a common source exposure, from person-to-person spread, or both.

Assessment of an outbreak by **place** provides information on the geographical extent of a problem and may also show clusters or patterns that provide clues to the identity and origins of the problem. See Table 1A.32.2. A simple and useful technique for looking at geographical patterns is to plot where the affected people live or work or may have been exposed onto a 'spot map' of the area.

Table 1A.32.1 Features of an epidemic curve

Feature on epidemic curve	Phenomenon or phenomena
Sudden rise in number of cases	Common source
Steep rise then gradual fall	Single source (or 'point source') epidemic, with all cases occurring within a single incubation period
Plateau (rather than peak)	Prolonged exposure period ('continuous common source epidemic')
Series of progressively taller peaks one incubation period apart	Person-to-person spread (a 'propagated' epidemic)
Early outlier	Background (unrelated) case Or a source of the epidemic Or a person who was exposed earlier than most of the people affected
Late outlier	Unrelated to the outbreak Or a case with a long incubation period Or exposure was later than for most of the people affected Or secondary cases (i.e. the person may have become ill after being exposed to someone who was part of the initial outbreak)

Reproduced from the National Center for Chronic Disease Prevention and Health Promotion (2006).

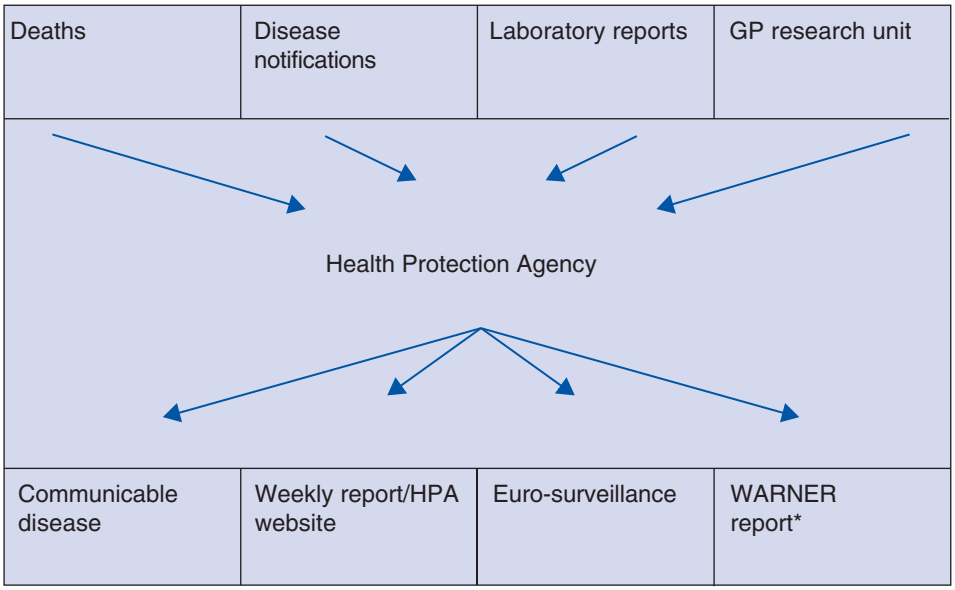
EXCEPTION REPORTING

Surveillance involves the mechanistic collection and analysis of data, and the communication of results. Early warning procedures aim to detect any divergence from usual frequency of disease or symptoms as soon as possible.

UK In the UK this work is coordinated by the Health Protection Agency (HPA). See Figure 1A.32.1.

Table 1A.32.2 Geographical patterns in epidemiology

Setting	Clusters
Community	Water supply Wind currents Proximity to a restaurant or shop
Hospital ward	Focal source (person-to-person spread) Scattered (common source, e.g. catering)
Surgical infection	Operating room Recovery room Ward



Weekly **Analysis **R**eport of **N**otifications above **E**xpected **R**ates.*

Figure 1A.32.1

SIGNIFICANT CLUSTERS

A cluster is a collection of events in space and/or time that is believed to be greater than would be expected by chance. If the population density varies between suspect areas, then a spot map will be misleading. To correct for this, the **attack rate** in each area should be calculated. The attack rate is the proportion of an exposed population at risk that becomes infected during a defined time period.

1A.33 COMBINING STUDIES

Systematic reviews, methods for combining data from several studies and meta-analyses

The process of combining the results from several studies offers a number of advantages. These are listed in Box 1A.33.1.

Box 1A.33.1

Increased power and precision	A single trial is always preferable to a combination, but often many underpowered studies will be published, all of which show a similar effect size but lack significance
Greater generalisability	Results taken from several studies may be relevant to a wider patient population
Efficiency and cost	It is quicker and cheaper to perform a systematic review than to embark on a new study

SYSTEMATIC REVIEWS

A systematic review is a summary of the medical literature conducted by using explicit methods. The process involves three steps: see Box 1A.33.2.

Box 1A.33.2

Literature search	A thorough, systematic search of the published and grey literature (see Section 1A.35)
Critical appraisal	Review of the studies found to determine which are relevant and valid
Amalgamation	Merger of the valid studies. When this is done systematically it is termed a **meta-analysis**

Two organisations that undertake systematic reviews are listed in Box 1A.33.3.

Box 1A.33.3

Examples: Systematic reviews

Cochrane Collaboration: collection of systematic reviews of medical and public health interventions. See Section 1A.39

www.cochranecollaboration.org

Campbell Collaboration: collection of systematic reviews of social and educational policies, including some health-related outcomes

www.campbellcollaboration.org

CRITICAL APPRAISAL

The critical appraisal (see Section 6A.2) of a systematic review addresses the same issues as those for individual studies, namely:

- Specific question stated
- Study population specified
- Unbiased collection of information (search strategy and data extraction)
- Valid conclusions (using appropriate statistical techniques if a meta-analysis).

META-ANALYSES

Combining trials depends on the quality of the constituent studies and on their being similar with regard to interventions and outputs. If there are enough studies with sufficiently similar characteristics that their results can be combined, then a meta-analysis is possible. Since the patients in one trial are likely to differ in a systematic way from those in another, it is wrong to compare individuals. However, since the individual studies are internally randomised, their effect sizes can be compared.

METHODS

- Fixed effects: assumes statistical homogeneity (i.e. the size of the treatment effect would be the same in all studies if there were no differences in the samples used)
- Random effects: allows for statistical heterogeneity. Will result in a more conservative estimate of effect than the fixed-effects model
- Results: summary measure (with confidence intervals produced)
- Findings presented (usually as a Forest plot or a funnel plot).

Figure 1A.33.1 shows a funnel plot of mortality results from trials of post-myocardial infarction β-blockade. As the sample size increases so the variability in odds ratios narrows.

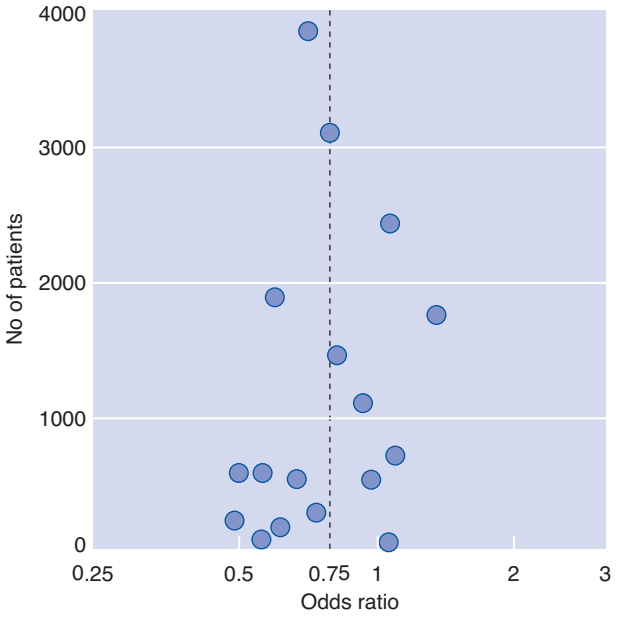


Figure 1A.33.1 Funnel plot of mortality results from trials of post-myocardial infarction β-blockade

Meta-analysis assesses the grand total of participants in all studies regarding observed and expected numbers.

If the null hypothesis is true then: **(Observed – Expected) = 0**.

If treatment is beneficial then: **(Observed – Expected) <0**.

BIAS IN META-ANALYSIS

- Poor trial quality, e.g. inadequate concealment may exaggerate treatment effect
- Publication bias: studies that show an effect are more likely to get published than those that do not. Can be examined using **funnel plots** displaying the effect size versus the study size – where the funnel would be expected to be symmetrical if no publication bias existed (see Sectio 1B.18).

RANDOMISED CONTROLLED TRIALS VERSUS OBSERVATIONAL STUDIES

Meta-analysis is often described with reference to randomised controlled trials (RCTs). It is possible to conduct meta-analyses of observational studies when it is not possible to collect data from RCTs (e.g. intervention already in common usage) but these are subject to other types of bias and the effect of confounding factors is hard to control.

FOREST PLOTS

The findings of a meta-analysis can be displayed using a Forest plot. See Figure 1A.33.2 for a Forest plot of four studies.

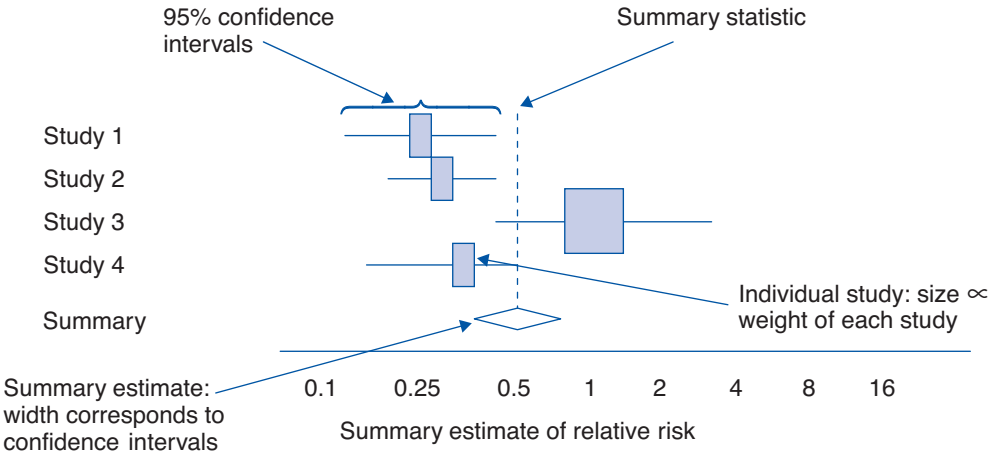


Figure 1A.33.2 Forest plot showing the estimates of relative risk for four studies

1A.34 ELECTRONIC BIBLIOGRAPHICAL DATABASES

Electronic bibliographical databases and their limitations

The extent of the health literature is colossal: there are currently over 30 000 medical journals, and the content of some of these extends over 100 years. Bibliographic databases make it possible to search these publications using Boolean operators such as AND, OR and NOT. Some databases (e.g. PubMed) contain journal material only; others (e.g. Popline) include other forms of literature such as reports, books, etc.

Databases can be searched either using free text words in the title or abstract ('keyword'), or by using a thesaurus term that may be grouped hierarchically. The latter are called 'exploded' or 'MeSH' terms and they have a number of advantages and disadvantages. See Box 1A.34.1.

Box 1A.34.1

Advantages	Automatically include all synonyms for a particular term, including: • American and British spellings • Plural and singular
Disadvantages	Prolonged time delays between publication and indexing mean that thesaurus terms may not keep pace with new areas of research

SEARCH STRATEGY

The following steps should be followed when conducting a search of the literature (Eyers 1998):

- Define the research question
- Choose which databases to search (see below)
- Define limits (e.g. time period)
- List the individual terms that constitute the research question
- Choose either keyword or thesaurus search strategy
- If keyword, then list all similar terms, spellings, etc by using a wildcard term denoted by typing an asterisk (e.g. diabet* will search for diabetes, diabetology, diabetologist and diabetic)
- If thesaurus, then identify thesaurus terms within the hierarchy
- Search for individual terms

- Combine individual terms (use AND to narrow the search and OR to broaden the search)
- Set further limits (e.g. restrict by language, or just review papers, etc)
- Save or print the papers identified
- Scan the titles for additional related terms to include in the search strategy
- Save the refined research strategy for future use.

LIMITATIONS OF DATABASES

No one database has access to all forms of publication, and no one library will have access to all databases. There is a tendency in public health to concentrate on health databases, but, being a multidisciplinary specialty, other databases should also be searched (e.g. economics, anthropology, etc.). There is also often a bias towards English language publications

Databases have a limited span of years: research that pre-dates databases is often ignored. There is often a time delay between publication and appearance in database. Using the internet can sometimes overcome this. However, using the internet as a database is risky as there are no quality controls. It does, however, offer access to the grey literature and to full-text information.

1A.35 GREY LITERATURE

Grey literature is written material issued by a body with a primary activity that is not publishing. As such, it is not readily available through traditional publication channels such as books and journals. Advantages and disadvantages of grey literature are shown in Box 1A.35.1.

Box 1A.35.1

Advantages	The internet makes this easier to access Presents less orthodox views Gives perspective to published material
Disadvantages	Traditionally difficult to access – especially publications in paper form only No quality control: the onus is on the reader to assess quality and credibility

Box 1A.35.2 gives some examples of grey literature.

Box 1A.35.2

- Technical and scientific reports
- Conference papers
- Internal reports from government and non-governmental organisations
- Government documents
- Theses
- Fact sheets
- Unpublished reports

1A.36 EVIDENCE-BASED MEDICINE AND POLICY

The term 'evidence-based medicine' (EBM) first appeared in 1992 and it relates to the **explicit** use of the current best evidence for decision-making at the level of the **individual patient**. In addition to clinical skills, a practitioner of EBM requires expertise in:

- Retrieving, ranking and interpreting the evidence
- Communicating evidence to patients
- Applying evidence to clinical decisions.

Advantages and diasadvantages of EBM are shown in Box 1A.36.1.

Box 1A.36.1

Advantages	Explicit use of best evidence
	Opinion of 'medical expert' demoted to least valid form of evidence
Disadvantages	Publication bias (failure to publish negative results)
	Retrieval bias (limitations of databases)
	Lack of evidence ≠ lack of benefit
	Lack of robust evidence for treatments other than drugs
	Evidence typically applies to populations, not necessarily to individuals
	Diminishes value of clinical nous

EVIDENCE-BASED POLICY

Beginning in the late 1990s, there has been a drive to place more explicit importance on evidence when making public policy decisions, i.e. a push for engineered policies rather than policies of conviction. The theory of evidence-based policy is that decisions should be shaped as shown in Box 1A.36.2.

Box 1A.36.2

Problem identified by policy-makers ➔ Problem solved by researchers ➔ Solution adopted as policy

Or

New knowledge ➔ Knowledge adopted into policy

In the real world there are very few examples of evidence-based policy occurring in this way. Instead, decision-makers tend to absorb the evidence which then appears unexpectedly in the future as policy: the so-called **enlightenment** model (Buse et al 2005).

1A.37 HIERARCHY OF RESEARCH EVIDENCE

Hierarchy of research evidence – from well-conducted meta-analyses down to small case series

The Centre for Evidence Based Medicine has ranked the different types of research evidence based on how likely they are to be true, as listed in Table 1A.37.1 The strongest evidence comes from a systematic review that demonstrates consistent findings from several high-quality RCTs. This is termed **level 1** evidence, and recommendations based on level 1 results are called **grade A**. Further down the hierarchy come more heterogeneous findings and evidence from less robust sources.

A similar system for ranking research evidence is used by the National Institute for Health and Clinical Excellence (NICE) in the production of guidelines and technology appraisals to indicate its relative level of confidence in the information used (see Section 1C.5).

Table 1A.37.1 Hierarchy of evidence

Level		Study type
1	A	Systematic review of RCTs
	B	Individual RCTs
	C	All or none
2	A	Systematic review of cohort studies
	B	Individual cohort study
	C	Ecological studies
3	A	Systematic review of case–control studies
	B	Individual case–control studies
4		Case series
5		Expert opinion

Reproduced from the Centre for Evidence-Based Medicine (2007).

1A.38 PUBLICATION BIAS

Publication bias is a tendency for journals to report positive results (where something was found to happen) rather than negative results (where something was found not to happen) and neutral results (no conclusive finding). It exists because researchers are less likely to submit, and publishers less likely to publish, negative and neutral results. This bias distorts meta-analyses (see Section 1A.33).

Non-reporting of RCTs is beginning to be regarded as scientific and ethical misconduct. In an effort to thwart the problem, registers such as the 'Meta Register of Controlled Trials' have been established, and medical journals have agreed to publish only registered RCTs. Other options to detect or reduce publication bias include:

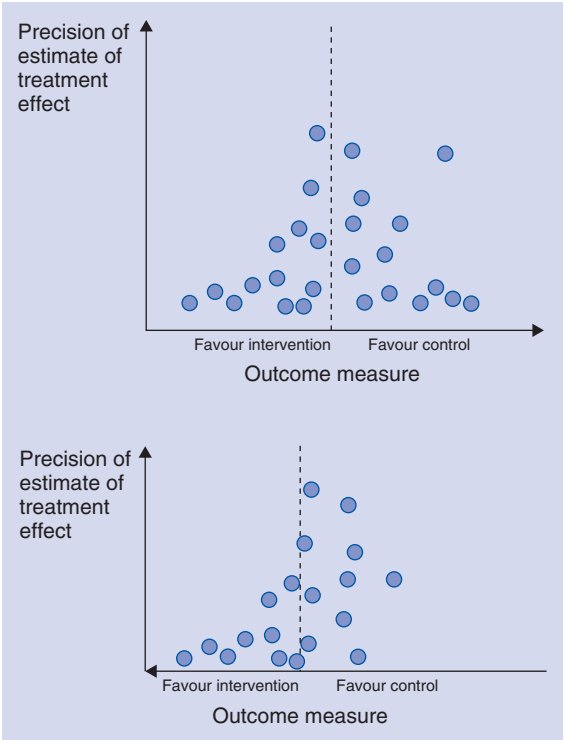
- Active discouragement of studies that do not have sufficient power to detect effects
- Examining effects in meta-analyses using funnel plots (see Figure 1A.38.1).

1A.39 THE COCHRANE COLLABORATION

The Cochrane Collaboration was established in 1993 and named after the British epidemiologist **Archie Cochrane** (1909–1988). Cochrane was a notable contributor to the development of epidemiology as a science. Between 1960 and 1974 he was Director of the Medical Research Council Epidemiology Research Unit in Cardiff.

The Cochrane Collaboration is an international, non-profit, independent organisation. It produces and disseminates systematic reviews of health-care interventions, and promotes the search for evidence in the form of clinical trials and other studies of the effects of interventions. As of 2004, there were over 11 500 people working within the Cochrane Collaboration in over 90 countries – half of whom were authors of Cochrane Reviews.

A key function of the collaboration is to produce systematic reviews (meta-analyses) of randomised controlled trials (see Section 1A.33). Cochrane Reviews are systematic assessments of evidence of the effects of health-care interventions, intended to help people to make informed decisions about health care. Cochrane Reviews ensure that health-care decisions throughout the world can be informed by high-quality, timely research evidence. These are published as part of the quarterly Cochrane Library, in the **Cochrane Database of Systematic Reviews**. Other components of the library include the Cochrane Methodology Register and the Health Technology Assessment Database.



In this non-biased funnel plot, the circles represent the point estimates of the trials. The pattern of distribution resembles an inverted funnel. Larger studies tend to be closer to the pooled estimate (the interrupted line). In this case, the effect sizes of the smaller studies are more or less symmetrically distributed around the pooled estimate

This biased funnel plot shows that the smaller studies are not symmetrically distributed around either the point estimate (dominated by the larger trials) or the results of the larger trials themselves. The trials expected in the bottom right quadrant are missing. This suggests publication bias and an overestimate of the true treatment effect

Figure 1A.38.1 Funnel plots for the assessment of publication bias. *Reproduced from Montori et al (2000)*

1A.40 EPIDEMIOLOGICAL RESEARCH ETHICS

Ethics and etiquette of epidemiological research

Ethics is the philosophical discipline concerned with understanding how human beings should act, what is good and what kind of life is best. In 1979 Beauchamp and Childress (2001) identified four principles associated with ethical medical practice: see Box 1A.40.1.

Box 1A.40.1

Autonomy	Human rights, dignity, freedom
Non-malfeasance	Do no harm
Beneficence	Do good
Justice	Equity, fairness

In public health, beneficence implies acting in the best interest of the population or society as a whole (McKeown and Weed 2002). This should be done in a just way, ensuring the fair distribution across the population of both benefits and risks. At a population level there is sometimes a tension between these principles, because of the conflicting aims associated with benefiting an individual and those of providing the optimal conditions for the wellbeing of the community. See Box 1A.40.2 for examples of such tensions.

Box 1A.40.2

Examples: conflicts between ethical principles in public health

- Fluoridation of the water supply does not permit individual informed consent. It may occur despite the opposition of people who are opposed to the intervention
- If health-care resources are redistributed with the aim of providing equitable services, it may harm those from whom resources are removed

In research studies, two ethical issues of particular importance are **informed consent** and **confidentiality**.

INFORMED CONSENT

Truly informed consent requires the features listed in Box 1A.40.3.

Box 1A.40.3

Competence	Do subjects understand what is involved? Children over a certain age can assent to take part but may lack full capacity for truly informed consent
Voluntariness	Are participants free to leave a trial at any point? Have they been put under excessive pressure to enrol in a research project? Patients need to be confident that their standard of care will not be affected by their decision to take part in a study
Understanding	Understanding the **risks, burdens** and **benefits** of the study
Documentation	Written consent is required in trial settings. In cluster-randomised trials, communities may be asked to give consent but it is not always clear who should give consent for the community

CONFIDENTIALITY

Research may involve the collection of private or sensitive information. Such confidential information should not be shared with anyone without consent except when there is a clear ethical justification (e.g. approval by a human subjects research review panel) or a legal requirement (e.g. regulations to protect children).

The use of identifiable data in research without consent requires demonstrating the importance of the research, the minimal risk to those whose information is used, the promise of benefit to society and an obligation to maintain the confidentiality of the information.

DATA PROTECTION

UK In the UK, the confidentiality of participants' information is protected by the **Data Protection Act 1998**. This contains eight principles, which state that the data must be:

1. **Processed** fairly and lawfully
2. **Obtained** and used only for specified and lawful purposes
3. **Adequate**, relevant and not excessive
4. **Accurate** and, where necessary, kept up to date
5. Kept for **no longer than necessary**
6. **Processed** in accordance with the individual's rights (as defined)
7. Kept **secure**
8. **Transferred** only to countries that offer adequate data protection.

Ire Similar principles apply in Ireland through the Data Protection Acts of 1988 and 2003. The Data Protection Commissioner provides information and guidance and ensures compliance with the legislation by those who keep personal data. Responsibility rests with data controllers, i.e. those who decide what information is to be collected or stored, to what use it is put and when it should be deleted or altered. Data controllers in public bodies and organisations specified in the 1988 Act (such as banks, insurance companies and those who keep personal data of a sensitive nature) are required to register with the Commissioner.

Aus Current privacy legislation in Australia provides for sensitive information about an individual to be collected without that individual's consent only when the information meets all three criteria below:

- The information is necessary for
  - research relevant to public health; or
  - compilation or analysis of public health statistics; or
  - management or monitoring of a health service
- The purpose cannot be served by collection of non-identified information; and
- It is impractical to seek the consent of individuals.

Australian states and territories have their own legislation that gives protection to people whose work requires them to deal with identified information, e.g. cancer registry staff, communicable disease control officers.

NZ In New Zealand, the main legislation and codes covering the privacy of personal information are:

- The Privacy Commissioners Act 1991 – established the office of Privacy Commissioner and the legal requirement for data matching
- The Privacy of Information Act 1993 – which focuses on good personal information handling practices and applies to almost every person, business and organisation in New Zealand. It includes 12 information privacy principles covering collection, holding, use and disclosure of personal information and use of unique identifiers. The Act also gives the Privacy Commissioner the power to issue codes of practice, specifies complaints mechanisms and rules governing data matching.
- The Health Information Privacy Code 1994 – which specifies a code of practice for the health sector regarding the collection, use, holding and disclosure of personal health information (i.e. that relates to identifiable individuals), access to health information and the assignment of unique identifiers. For the health sector, this code takes the place of the information privacy principles in the Act. It applies to all levels of the health system from large institutions to sole practitioners.

CALDICOTT GUARDIANSHIP

Eng Wal The Caldicott report, published in 1997, found that compliance with data protection statutes was variable. As a result, the Department of Health (DH) in England and the Welsh Assembly in Wales established a register of Caldicott guardians. Every NHS and social services institution must appoint one senior member of staff to take responsibility for protecting patient-identifiable information.

The other Caldicott recommendations (Walker 1999) are that NHS organisations should:

- Develop protocols to manage the sharing of patient information with other institutions
- Permit access to patient information only to employees who need to know that specific information
- Review and justify all uses of patient information
- Instil a culture of data protection through training, database design, etc.

Similar arrangements for protection of patient confidentiality exist in Scotland.

ELECTRONIC DATA STORAGE

UK The main information security standards that apply across the UK are listed below (Knott 2006).

BS7799

Part 1 of this British standard is known as the *'Code of Practice for Information Security Management'* and provides guidance on best practices in information security management. Part 2, the *'Specification for Information Security Management Systems'*, is the standard against which an organisation's security management systems are assessed and certified. It has been revised and now exists in the form of ISO/IEC 27001:2005.

IEC 61508

This standard of the **International Electro-technical Commission** contains requirements for ensuring that computer systems are designed, operated and maintained in ways that have sufficient integrity. Section 3 sets out the requirements on computer suppliers for new equipment.

INFORMATION GOVERNANCE

NHS organisations are subject to controls over the storage of information for both routine purposes and research. The overarching framework to assure the quality of these processes is called **information governance**.

ETHICS COMMITTEES

In research studies, ethical principles are protected by requirements of ethics local/national ethics committees. Proof of ethical approval is often a prerequisite for journals to consider publishing a study.

Eng Scot In England and Scotland, any research conducted in the NHS must receive approval from an ethics committee before commencing. A research ethics committee will consider issues such as whether the study has the potential to benefit society/participants, how participants are recruited, how participants' confidentiality will be protected, as well as any possible effects of the study on health or wellbeing and processes for obtaining informed consent.

Wal Any research requires the approval of a local ethics committee. It will subsequently be subject to local research governance arrangements that seek to ensure that the research is conducted in a manner consistent with the approval given by the ethics committee. Where the research setting is primary care, the researcher is required to seek local management approval from the local health board – which may decline approval if it believes that the research poses a risk either to patients or to delivery of NHS services.

Eng Wal Where research is not located in a single area, the NHS Central Office of Research Ethics Committees (COREC) is used. COREC also provides guidance, training and support to local committees.

Ire The Irish Council for Bioethics Guidance (2004) stipulates that all research involving or impacting on human participants requires ethics review by a research ethics committee (REC). Procedures vary for obtaining permission to undertake research in the hospital setting. Usually an application must be submitted to the hospital's REC, but national committee approval may be accepted as sufficient by some hospitals and regional approval may be accepted by smaller hospitals in the regional network. For community-based research, protocols may be submitted to the REC under the joint auspices of the Faculties of Occupational Health and of Public Health Medicine of the Royal College of Physicians of Ireland.

SA In South Africa there is a national process controlled by the *Ethics in Health Research: Principles, Structures and Processes Research Ethics Guidelines* (2004). Biomedical research ethics committees at the various universities now have to be registered and structured according to the guidelines.

Aus Ethical approval comes from a formally constituted Human Research Ethics Committee (HREC). The National Health and Medical Research Council has guidelines for the establishment and accreditation of research ethics committees. Ethics committee approval must be obtained before research on humans can go ahead.

NZ In New Zealand, all health and disability research that involves human subjects must be sent for ethical review. This includes observational research (e.g. descriptive studies using already collected personal information or information obtained by questionnaires and interviews) as well as experimental research (e.g. clinical trials). The only exceptions are certain type of audits and related activities (e.g. quality assurance and programme evaluation), public health investigations (e.g. outbreak investigations and surveillance) and where a statutory exclusion applies (e.g. official statistics).

The main system for ethical review is the set of health and disability ethics committees (HDECs). There are six regional committees and a multiregional committee for research carried out nationally or in more than one region. Research may be reviewed by an accredited institutional ethics committee (IEC) established by universities and private companies, but in most circumstances will also need to be reviewed by a HDEC. Additional review is required in some specific areas, notably: clinical trials of a pre-registration medicine, research involving assisted human reproduction and research involving manipulation of human genetic material (e.g. gene therapy).

1A.41 GENETICS

Understanding of basic issues and terminology in the design, conduct, analysis and interpretation of population-based genetic association studies, including twin studies, linkage and association studies

See Section 2D for details of public health genetics. The principles of genetic epidemiological studies (Dorak 2007) are outlined below and are summarised in Table 1A.41.1

FAMILY STUDIES

Family studies are performed in order to determine whether there is a genetic component to a particular disorder. They aim to detect a higher occurrence rate of disease in siblings or offspring of an affected person. This can be measured by the relative recurrence risk (RRR) or familial risk ratio (FRR).

RRR = Probability that a particular type of relative (sibling, cousin, etc.) of an affected individual is also affected ÷ Prevalence of the disease in the general population.

A higher RRR or FRR is a necessary (though not sufficient) attribute to decide that there is a genetic component to a disorder.

Table 1A.41.1 Genetic studies

Study question	Appropriate study designs	Unit of analysis/results obtained
Is there a genetic component to the disorder?	Family studies	RRR, FRR
What is the contribution of genetics as opposed to environment/other sources to the trait?	Family – twin, adoption	Percentage Concordance, discordance
What is the model of transmission of the genetic trait?	Family – segregation	
What is the location of the disease gene(s)?	Family – linkage	Lod score Recombination fraction
What is the allele associated with the disease susceptibility?	Population based (and family studies) – association	Linkage disequilibrium (LD)

TWIN STUDIES

Twin studies explore the relative contributions of genes and the environment by comparing identical and non-identical twins to explore the relative contributions of genes and environment to health and disease.

Identical twins can be considered to share both genes and environment. Non-identical twins act as a control: they share the same environmental factors, but are only as genetically alike as any other sibling.

Where a disease occurs with a higher probability in identical twins than in non-identical twins, this is known as *concordance* and indicates a shared genetic basis for the trait. The strength of the genetic predisposition is expressed as the heritability, e.g. asthma has a heritability of 60%. Note that heritability is a population-based statistic (i.e. measure of variance within a population), not a measure of the relative contribution of genes in an individual, nor a measure of their risk. Furthermore, the heritability of a condition is unrelated to the number of genes that influence it.

- **Pairwise** concordance is the percentage of concordant pairs (i.e. both twins affected) in a group of twins where at least one member of each pair is affected.
- **Probandwise** concordance is the percentage of twins whose twin becomes affected during the study, in a group of twins where just one member of each pair is affected.

LINKAGE STUDIES

Genetic linkage occurs when two alleles (two forms of a gene) are inherited together. In general, a parent can pass on either allele of a particular gene to his or her offspring, regardless of which alleles of other genes are passed on. This is because chromosomes are sorted randomly during meiosis. Sometimes, however, two alleles have a tendency in the population to be inherited together more often than would be expected by chance. This is called **linkage disequilibrium** and it occurs because the alleles are linked, i.e. found close to each other on the same chromosome.

When considering two genetic traits that appear to be linked, the *offspring* of people who have these two traits are studied. The lower the percentage of the offspring who do not have both traits, the smaller the physical distance between the two genes on the chromosome.

A study of the linkages between many genes enables the creation of a genetic map called a **linkage map**. Linkage may be quantified using the **lod score**, which is the logarithm of odds of linkage. Traditionally, a lod score >+3 is considered to be significant. This corresponds to a one-sided p value of 10^{-4}.

ASSOCIATION STUDIES

These studies measure the relative frequency with which a particular polymorphism occurs together with the disease of interest in a **population**, i.e. the extent to which it is associated with the disease. Designs of association studies include the following.

POPULATION-BASED CASE–CONTROL STUDIES

These are studies in which two groups are chosen from the population: unrelated affected cases and unrelated healthy controls. The relative frequencies of alleles at a single marker locus are tested statistically (using Pearson's χ^2 test or logistic regression).

FAMILY-BASED ASSOCIATION STUDIES

Families are identified that contain affected and unaffected individuals. The transmission of alleles from parents to their affected and unaffected offspring is compared statistically.

Linkage disequilibrium is measured in association studies. The closer the physical (genetic) distance between the disease and marker alleles, the lower the recombination frequency and the stronger the magnitude of linkage disequilibrium.

1B Statistics

An understanding of basic statistical principles is required in many parts of the membership examination. For the mathematically inclined we have included formulae and their derivations to assist the learning process. However, the only formulae that must be memorised for the examination are listed in Section 6D.

1B.1 ELEMENTARY PROBABILITY THEORY

Probability is a measure of the likelihood of an event. It is expressed as a positive number between 0 (event never occurs) and 1 (event is certain to occur).

There are several approaches to calculating the probability of an event: see Table 1B.1.1.

Table 1B.1.1 Approaches to calculating the probability of events

Subjective	Personal degree of belief that an event will occur, such as a doctor's opinion for clinical decision-making
Frequentist	Proportion of times an event would occur in a large number of similar repeated trials, e.g. number of 'heads' in coin tossed 100 times
A priori	Requires knowledge of the probability distribution

RULES OF PROBABILITY

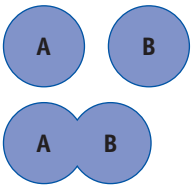
Two rules determine the ways in which two or more probabilities may be combined.

ADDITION RULE (OR)

This is used to find the probability *P* of **at least one event** occurring out of two or more possible events.

Mutually exclusive events:

P of obtaining A *or* B = $P(A \text{ or } B) = P(A) + P(B)$



Non-mutually exclusive events:

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

MULTIPLICATION RULE (AND)

The probability of the **joint occurrence** of two or more events depends on whether the events are:

- Dependent (i.e. the occurrence of one event affects the probability occurrence of the other event), or
- Independent (the occurrence of each event does not affect the other).

Independent events: *P* of obtaining A *and* B = $P(A) \times P(B)$

Dependent events: $P(A \text{ and } B) = P(A) \times P(B|A)$.

Note that $P(B|A)$ is known as conditional probability, i.e. probability of B occurring **given** that A has already occurred (see Section 1B.19).

1B.2 METHODS FOR THE QUANTIFICATION OF UNCERTAINTY

Uncertainty, in a statistical sense, refers to the fact that measurements do not always report the exact truth. A measurement may be defined as a translation of a quantity in the real world, through the use of a measuring device or instrument, into a number and a unit. Together the number and the unit represent the reality of interest. For example, a microscope might be used to estimate the size of an average bacterium on a slide (0.5 μm). A cardiologist might want to assess the amount of myocardial damage sustained by a patient by measuring the blood troponin. A researcher might want to measure whether someone has a minor psychiatric disorder using the General Health Questionnaire (GHQ). All of these measurements are associated with a degree of imprecision.

There are uncertainties inherent to **measurement** and **prediction** but this can be managed by quantifying and accounting for this uncertainty.

PROBABILITY THEORY

Probability theory, i.e. the mathematical discipline of studying probability, offers a framework for dealing with and quantifying probabilities.

An event A with probability $P(A)$ takes a numerical value between 0 and 1. Table 1B.2.1 provides a summary of the interpretation of probability values.

Table 1B.2.1 Interpretation of probability values

Value of $P(A)$	Interpretation
1	A is certain to happen
Near 1	A is very likely to happen
>0.5 and <1	A is more likely to happen than not
0.5	A is as likely to happen than not
>0 and <0.5	A is less likely to happen than not
Near 0	A is very unlikely to happen
0	A cannot happen

PROBABILITY DISTRIBUTIONS

The probability distribution is the **range of values** that a random variable can take, and the **relative frequency**

of occurrence of these values. Probability distributions are theoretical distributions, and they are expressed mathematically with a mean and a variance. If a random variable is discrete or continuous, then its probability distribution will likewise be discrete or continuous. See Box 1B.2.1.

Box 1B.2.1

	Discrete variables	Continuous variables
Distribution*	Binomial distribution Poisson distribution	Normal distribution Chi-squared (χ^2) *t*-distribution *f*-distribution
Example	Alive/dead Number of hospital admissions p.a.	Body weight Age
Sum of probabilities	Sum of all the probabilities = 1	When drawn as a curve (called the probability density function) the total area under the curve = 1

**See Section 1B.5.*

STATISTICAL INFERENCE

Statistical inference is a deduction about a population that is based on measurements made on a random sample drawn from that population. Inference includes the aspects shown in Box 1B.2.2.

Box 1B.2.2

Point estimation	Sample mean used to estimate population mean (see Section 1B.6)
Interval estimation	Confidence interval (see Section 1B.3)
Hypothesis testing	Statistical significance testing (see Section 1B.9)

1B.3 ESTIMATION OF CONFIDENCE INTERVALS

Confidence intervals are a means of making inferences about the **accuracy of a sample value** that is being used as an estimate population value.

PRECISION IN SAMPLING

A major cause of apparent differences between observations is the effect of chance, and the magnitude of this chance is determined by the size of the sample. This effect can be modelled statistically. Statistical tests give the possibility (p) of chance alone being responsible for obtaining a result at least as extreme as that observed. This is the underlying principle behind all tests of statistical significance.

Note that the value of this probability (the p value) reflects only the role of chance. It does not take account of any biases or confounding that may exist, nor does it imply causality. Every p value is a composite measure of:

- Effect size
- Sample size.

By convention, a p value of **<0.05** is taken to be suggestive of a statistically significant result, since there is less than a 1-in-20 chance of the observed result being due to chance. However, it is preferable to think of a spectrum

of evidence rather than a single cut-off: the lower the *p* value, the stronger the indication that the finding is significant (i.e. that the observed finding was not due to chance alone).

- $0.05 < p < 0.10$ (**'suggestive'** evidence)
- $0.01 < p < 0.05$ (**'quite strong'** evidence)
- $0.001 < p < 0.01$ (**'strong'** evidence)
- $p < 0.001$ (**'very strong'** evidence)

Because the *p* value depends on both the size of the effect and the sample size, given a large enough sample size, even a minute difference will be statistically significant. For this reason, confidence intervals are more informative than the *p* value alone. See Box 1B.3.1

Box 1B.3.1

***p* values and confidence intervals**

- A narrow confidence interval implies a large sample size or a small number of very similar results
- If a *p* value is non-significant and there is a **narrow** confidence interval, then it is likely that there truly is no effect
- If a *p* value is non-significant and there is a **wide** confidence interval, then this suggests that the sample size may be too small

Note that no *p* value, however small, can exclude chance completely. Likewise, no *p* value, however large, can guarantee that an association was definitely due to chance.

CALCULATION OF CONFIDENCE INTERVALS

Confidence intervals are calculated by taking a sample value, then adding and subtracting a multiple of the standard error for that value. For a 95% confidence interval ($p = 0.05$) this multiple is 1.96 (see Box 1B.3.2); for a 99% confidence interval ($p = 0.01$) it is 2.58.

Box 1B.3.2

95% confidence interval = sample value ± (1.96 × standard error)

Begin by calculating the standard error for the sample value:

For a **mean:** standard error $= \frac{s}{\sqrt{n}}$

For a **proportion**: standard error $= \sqrt{\frac{p(1-p)}{n}}$

For **two proportions:** standard error $= \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

Here s = standard deviation for the sample; n = number in the sample; p = proportion.

Boxes 1B.3.3–1B.3.5 give three examples.

Box 1B.3.3

Example 1: calculating the confidence interval of a mean

The mean pO_2 arterial blood test for a sample of 56 patients with COPD was 8.9 kPa, with a standard deviation of 0.8 kPa. What is the 95% confidence interval for the mean pO_2 of the population of all COPD patients?

Standard error $= \frac{0.8}{\sqrt{56}}$
$= 0.11$

95% confidence interval $= 8.9 \pm (1.96 \times 0.11)$
$=$ between 8.7 and 9.1 kPa

Box 1B.3.4

Example 2: Calculating the confidence interval for two proportions

57% of the 864 patients at a stop-smoking clinic had quit smoking by the end of the programme, compared with 42% of 795 patients at a neighbouring clinic.

Standard error (difference in proportions) $= \sqrt{\frac{0.57(1-0.57)}{864} + \frac{0.42(1-0.42)}{795}}$
$= 0.0243$

95% confidence interval for the difference $=$ sample value $\pm (1.96 \times$ standard error$)$
$= (0.57 - 0.42) \pm (1.96 \times 0.024)$
$=$ between 0.198 and 0.102
$=$ between 10.2% and 19.8% difference

Box 1B.3.5

Example 3: Calculating the confidence interval for a proportion

87% of the 265 patients with a deep vein thrombosis (DVT) at an anticoagulation clinic completed their 6-month course of warfarin. What is the standard error and 95% confidence interval for completion of the course?

Standard error $= \sqrt{\frac{0.87(1-0.87)}{265}}$
$= 0.02$

95% confidence interval $= 0.87 \pm (1.96 \times 0.02)$
$=$ between 0.83 and 0.91
$=$ between 83% and 91% of patients

INTERPRETATION OF CONFIDENCE INTERVALS

A confidence interval is a **range of values** indicating the **precision** with which the sample estimate is likely to represent the population from which the sample was drawn. For a 95% confidence interval, the true value for the

population at large will lie within this range 19 times out of 20. For a 99% confidence interval (which will be a wider range), the true value will lie within the range expressed 99 times out of 100.

- For measures of **absolute risk**, where a 95% confidence interval includes **zero**, there is no evidence at the $p = 0.05$ level that there is a true difference.
- For measures of **relative risk**, where a 95% confidence interval includes **one**, there is no evidence at the $p = 0.05$ level that there is a true difference.

In any diagram showing two or more point estimates and their associated confidence intervals, there are three possibilities regarding the overlap of the points and intervals. These may be interpreted as shown in Box 1B.3.6.

Box 1B.3.6

Overlap of intervals or parts	Interpretation
Confidence intervals do not overlap	Significant difference at that significance level
Confidence intervals overlap but the point estimates are outside the confidence intervals of the other	Unclear: requires calculation of a significance test
Point estimate of one sample falls within the confidence intervals of the other	No difference at that significance level

1B.4 CONDITIONAL PROBABILITY

See also Section 1B.1.

Conditional probability represents the chance that one event will occur, given that a second event has already occurred. Conditional probability allows the evaluation of how different treatments or exposures influence the probability of outcomes such as disease or mortality. It also provides a useful way to evaluate diagnostic tests. Box 1B.4.1 gives an example.

Box 1B.4.1

Example: Breastfeeding in preterm infants

In a study of breastfeeding in preterm infants, the infants were randomised into two groups. The treatment group was fed by nasogastric (NG) tube, and the control group was bottle-fed. The purpose of the study was to test whether using the NG tube for feeding rather than using a bottle would increase the likelihood of breastfeeding at discharge from hospital and afterwards.

In this table, the rows represent the group (NG tube versus bottle) and the columns represent feeding status at discharge (exclusive breastfeeding versus partial/no breastfeeding).

	Exclusive breastfeeding at discharge		
	No	**Yes**	Row total
Bottle fed	27	20	47
NG tube fed	10	32	42
Column total	37	52	89

To obtain cell probabilities, each entry in the table above is divided by the total sample size.

	Exclusive breastfeeding at discharge		
	No	**Yes**	Row total
Bottle fed	(27/89 × 100) **30.3%**	(20/89 × 100) **22.5%**	52.8%
NG tube fed	(10/89 × 100) **11.2%**	(32/89 × 100) **36%**	47.2%
Column total	41.6%	58.4%	100%

The probabilities in the row and column totals represent **unconditional probabilities**. The interior probabilities represent the probabilities of the intersection between two events.

Conditional probabilities represent the probability that an event will occur when attention is restricted to a specific row (or sometimes a specific column) of a 2 × 2 table.

	Exclusive breastfeeding at discharge		
	No	**Yes**	Row total
Bottle fed	(30.3/52.8 × 100) **57.4%**	(22.5/52.8 × 100) **42.6%**	100
NG tube fed	(11.2/47.2 × 100) **23.8%**	(36/47.2 × 100) **76.2%**	100
Column total	(41.6/100 × 100) 41.6%	(58.4/100 × 100) 58.4%	100

A vertical bar, '|' denotes conditional probability and is pronounced as the word 'given'.

For example,

P(exclusive breastfeeding | NG feeding), the probability of exclusive breastfeeding, given that NG tube feeding occurred, is 76.2%

P(exclusive breastfeeding | bottle-feeding), the probability of exclusive breastfeeding, given that bottle-feeding occurred, drops to 42.6%

The unconditional probability of breastfeeding is 58.4%, i.e. somewhere between the two conditional probabilities.

Reproduced with permission from Fleiss (1981).

1B.5 STANDARD STATISTICAL DISTRIBUTIONS

A statistical distribution is an arrangement of the values of a variable that shows how frequently each value is observed or is expected theoretically to occur.

BINOMIAL DISTRIBUTION

The binomial distribution shows the frequency of events that have two possible outcomes. It is constructed from two parameters: n (sample size) and π (true probability). When the sample size is large, the binomial distribution approximates to the normal distribution: see Figure 1B.5.1.

This distribution is used for discrete data with two outcomes (e.g. success or failure) and the sampling distribution for proportions.

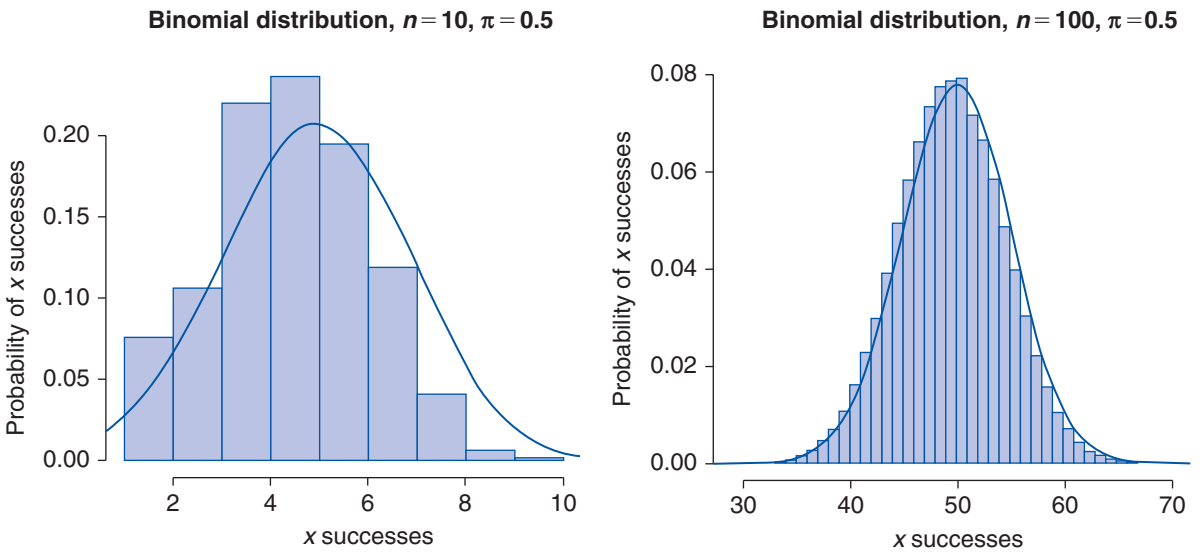


Figure 1B.5.1 Graph of binomial distribution approximating normal distribution as sample size increases, where n is the sample size, π is the probability and x represents the number of successes obtained

POISSON DISTRIBUTION

The Poisson distribution shows the frequency of events over time in which the events occur independently of each other. An example is the discharge of alpha particles during radioactive decay. In the Poisson distribution, the parameter is the variance which is equal to the mean (and therefore the standard deviation is equal to the square root of the mean). This means that small samples give asymmetrical distributions, and large samples approximate the normal distribution: see Figure 1B.5.2. The horizontal axis shows x, the number of occurrences of an event within a particular time period. The vertical axis shows the probability of obtaining x events during that period.

The distribution assumes that the data are **discrete**, occurring at **random** and **independent** of each other.

The Poisson distribution is used in the analysis of **rates** (e.g. incident rates of disease). Since it leads to a prediction of randomly occurring events, it allows determination as to whether observed events are occurring randomly or not.

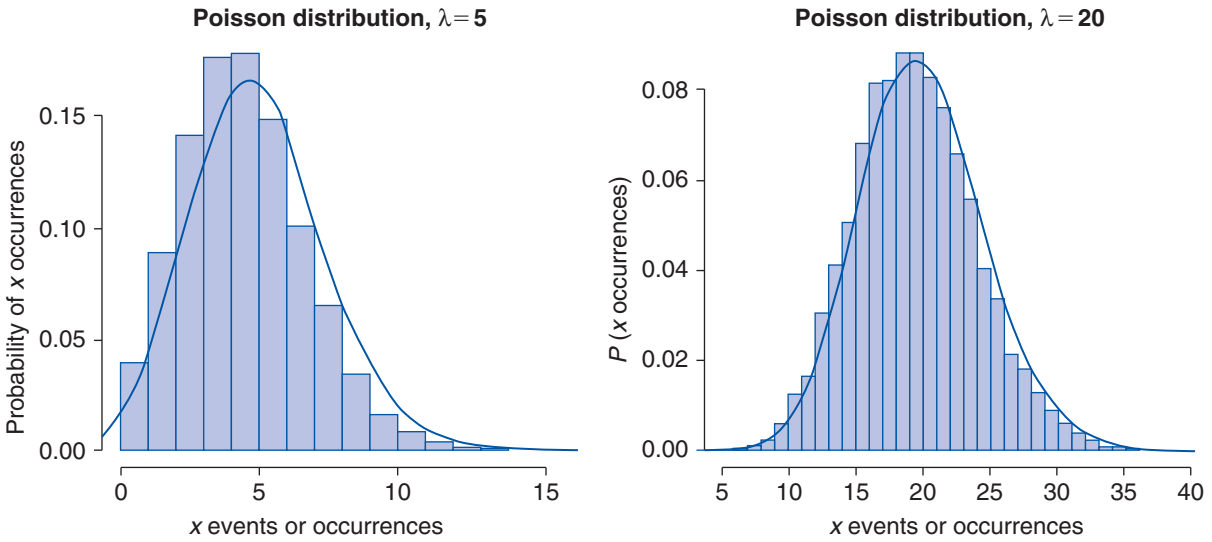


Figure 1B.5.2 Approximation of Poisson distribution to normal distribution as λ (the mean or expected value) increases. *Adapted from Zoonekynd (2007)*

CENTRAL LIMIT THEOREM

This theorem states that for any population, no matter what its distribution, if a sufficiently large number of samples is taken, then the distribution of the means of these samples will always be normally distributed, i.e. they will follow a normal (or 'gaussian') distribution.

NORMAL DISTRIBUTION

The normal distribution is a bell-shaped, symmetrical curve that is described by two parameters:

- Mean (μ)
- Variance (σ^2).

The standard normal distribution has a mean of 0 and variance of 1. In the standard normal distribution, 68% of the area under the curve is within 1 standard deviation of the mean, 95% of the area is within 1.96 standard deviations and 99% of the area is within 2.58 standard deviations: see Figure 1B.5.3.

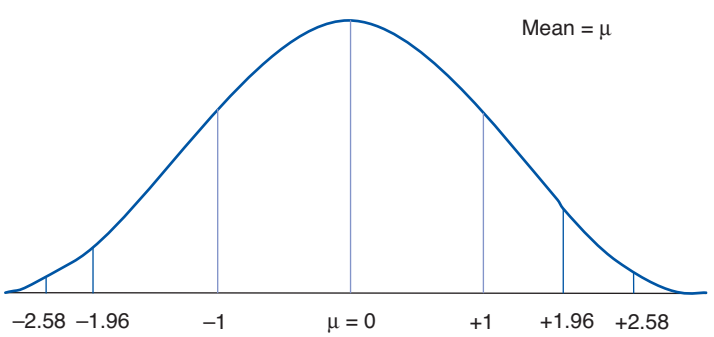


Figure 1B.5.3 The standard normal distribution

OTHER DISTRIBUTIONS

Table 1B.5.1 lists features and applications of some other distributions of continuous variables.

Table 1B.5.1 Probability distributions of continuous variables

Distribution	Features	Use and assumptions
***t*-distribution** (also called *'Student's t-distribution'*)	**Parameter:** degrees of freedom The shape is similar to the normal distribution but tails are more spread out As the degrees of freedom increases, the *t*-distribution approaches the normal distribution	**Uses:** • Estimating the mean of a normally distributed population when the sample size is small • Testing hypotheses with a single mean (small sample) • Testing hypotheses with two means (small samples) • Confidence intervals
Chi-squared	**Parameter**: degrees of freedom The shape is right-skewed, taking positive values With increasing degrees of freedom, it approximates the normal distribution	**Uses:** • Analysing categorical data, e.g. significance test for two categorical variables (comparison of observed and expected events) **Assumptions:** • If n is >40 then all x are >1 • If n is >20 but <40 all x should be >5, otherwise use **Fisher's exact test**
***f*-distribution**	• It is skewed to the right • Values are positive • It is defined by a ratio of two parameters: degree of freedom of the numerator and denominator of the ratio	**Uses:** • Useful for comparing two variances and more than two means using analysis of variance (ANOVA)

1B.6 PRINCIPLES OF MAKING INFERENCES FROM A SAMPLE TO A POPULATION

If the population is sampled in such a way that the sample correctly represents the population, then conjectures can be made about the population from the sample. Using a sample to make such inferences is more efficient in terms of time and resources than obtaining details on the whole population (which may also be impossible). However, the process of sampling introduces a degree of error. This error is minimised by taking a sufficiently large, random sample.

There are two methods of reaching a statistical inference: **estimation** and **hypothesis testing**.

ESTIMATION

An **estimate** is a measurement made from a sample, which has been drawn from a population. Estimates allow inferences to be drawn about the population, providing both a **point estimate** and a **confidence interval** (see Section 1B.5). The key to estimation is the probability with which particular values will occur during sampling – this allows the inference about the population to be made.

The values that occur are inevitably based on the sampling distribution of the population. In order to make an accurate inference about a population, **random sampling** is required, where each member of the population has the same probability of being selected for the sample (see Section 1A.25).

HYPOTHESIS TESTING

See also Section 1B.9.

Hypothesis testing involves comparing the observed findings in the sample with the findings that would have been expected theoretically (the hypothesis). The process differs according to whether a parametric test or non-parametric distribution is used: see Table 1B.6.1.

Table 1B.6.1 Differences between parametric and non-parametric tests

	Parametric	Non-parametric
Assumptions	Data follow a known distribution (e.g. normal or Poisson)	Makes no assumptions about the underlying distribution
Use	*Continuous* data are sampled from a population with an underlying *normal* distribution (or the distribution of which can be rendered normal by mathematical transformation, e.g. log or $1/x$)	When the distribution is not known or the distribution is small
Power	Powerful	Less powerful
Confidence interval calculation	Straightforward	Difficult
Standard deviation	Requires a similar standard deviation to population with which a comparison is being made	Do not need similar standard deviation
Sample size	Any	Best for <50 data cases
Examples	*t*-test ANOVA MANOVA Regression (all types) Correlation	Wilcoxon's signed rank Mann–Whitney U Rank correlation (e.g. Spearman, Kendall) Chi-squared

1B.7 MEASURES OF LOCATION AND DISPERSION, AND THEIR APPROPRIATE USES

Data are summarised according to two parameters: location and dispersion. See Table 1B.7.1.

Table 1B.7.1 Examples of measures of location and dispersion

Parameter	Description	Examples
Location	Average value for the data	Mean Median Mode
Dispersion	Spread of the data	Variance Standard deviation

MEASURES OF CENTRAL LOCATION

Three measures of central tendency are commonly used in statistics, each with its advantages and disadvantages: see Table 1B.7.2.

Table 1B.7.2 Advantages and disadvantages of measures of central location

Measure	Definition	Calculation	Advantages	Disadvantages
Mean	The average value For a sample, the mean is denoted by $\bar{x}$ (pronounced 'x-bar') For a population, the mean is denoted by the Greek letter μ (pronounced 'mu')	$\bar{x} = \frac{\sum n_i}{N}$ $\mu = \frac{\sum n_i}{N}$	Amenable to statistical analysis	Sensitive to outliers Poor for asymmetrical distributions
Median	Value at the centre of the distribution	If all observations are arranged in ascending order of value, then the median is the middle value (half the items lie above and half below)	Unaffected by extreme outliers Good for skewed distributions	Value determined solely by rank, so carries no information about any other values
Mode	Commonest value or values	Value(s) that occurs with the highest frequency	Not greatly affected by extreme values Can give extra insights (e.g. suicide is bimodal, affecting young and elderly)	Not amenable to statistical analysis Sometimes no mode exists Data containing more than one mode can be difficult to interpret

MEASURES OF DISPERSION

The principal measures of dispersion are shown in Table 1B.7.3.

Table 1B.7.3 Principal measures of dispersion

Measure	Definition and calculation	Advantages	Disadvantages
Range	Difference between maximum and minimum sample values (units are the same as data)	Intuitive Simple – only depends on two observations	Sensitive to the size of the sample

Percentiles	Values are ranked and divided into 100 groups. The 100th percentile is the largest value. The *n*th percentile will have *n*% values below it and 100 – *n*% values above it	Useful for comparing measurements, e.g. BMI in groups of similar age and sex	Comparisons made at extreme ends of distribution less informative than at the centre
Quartiles	Values are ranked and then divided into four groups, each containing 25% of the data Interquartile range: middle 50% of the sample	More stable than the range as sample size increases	Unstable for small samples and not allowing further mathematical manipulation
Quintiles	As for quartiles, but five groups rather than four		
Variance	Average squared deviation of each number from its mean Variance in a sample: $s^2 = \frac{\sum(\bar{x} - n_i)^2}{n-1}$ Population variance: $\sigma = \frac{\sum(\mu - n_i)^2}{N}$	Useful for making inferences about the population	Units are the squared units of the data
Standard deviation	$\sqrt{\text{variance}}$	Most commonly used Units are the same as the data Not sensitive to a change in units Useful for making inferences about the population	Sensitive to some extent to extreme values
Coefficient of variation	Ratio of the standard deviation to the mean to give an idea of the size of the variation relative to the size of the observation	Allows comparison of the variation of populations that have significantly different mean values	Mean value is near zero, the coefficient of variation is sensitive to change in the standard deviation
Standard error	Measure of how precisely the population mean is estimated by the sample mean $SE = s / \sqrt{n}$	Used to calculate confidence interval	Depends on sample size, i.e. small sample = large standard error
Confidence interval	Range of values in which the 'true', or population value, is likely to lie (see Section 1B.3)	Indicates the reliability of an estimate	By chance alone about 1 in 20 significant findings at the p=0.05 level will be spurious. Wide confidence intervals provide less information

1B.8 GRAPHICAL METHODS IN STATISTICS

Graphics can be an effective means of conveying statistical information. Different graphics are used for **qualitative** and **quantitative** data.

QUALITATIVE DATA

The type of chart used should be chosen carefully based on the nature of the data, the analysis and the message to be conveyed.

BAR CHARTS

These use bars to represent the frequencies (or relative frequencies), such that the height of each bar equates to the frequency or relative frequency of its category. An example is shown in Figure 1B.8.1.

- Frequencies: counts
- Relative frequencies: percentage.

PIE CHARTS

The pie is a circle divided into a number of slices, each representing a different category. The size of each of each slice is proportional to the relative frequency of that category. An example is shown in Figure 1B.8.2.

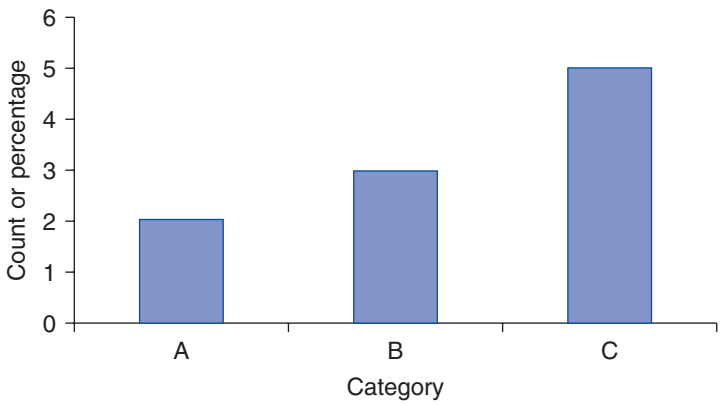


Figure 1B.8.1 Schematic of a bar chart

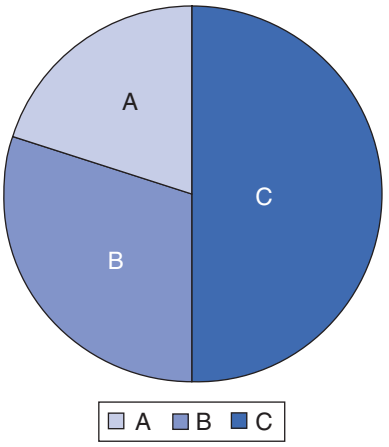


Figure 1B.8.2 Schematic of a pie chart

QUANTITATIVE DATA

The type of graphical method used to display quantitative data is chosen according to whether the data are **univariate**, **bivariate** or **multivariate**.

UNIVARIATE DATA

Here there is an observation of a single numerical variable. Graphical methods include **stem-and-leaf** displays, **boxplots** and **histograms**.

STEM-AND-LEAF DISPLAY

This is a quick technique for displaying numerical data graphically. A vertical stem is drawn, consisting of the first few significant figures. Any digit after the stem is called the leaf, i.e. the leaf is the last digit of the data value. An example is shown in Figure 1B.8.3. A back-to-back stem-and-leaf plot can be used to display data from two groups. Box 1B.8.1 lists advantages and disadvantages of stem-and-leaf display.

Box 1B.8.1 Advantages and disadvantages of stem-and-leaf display

Advantages	Disadvantages
Quick and easy to construct Actual values retained	Cumbersome for large data sets

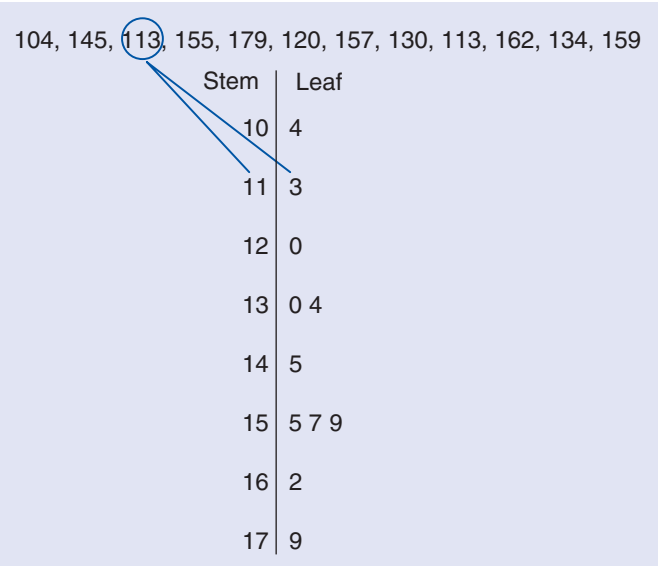


Figure 1B.8.3 Example of stem-and-leaf display constructed from 12 data points

BOXPLOTS

Also known as a box-and-whiskers plot, a boxplot shows a measure of **central location** (the median), two measures of **dispersion** (the range and interquartile range), the **skewness** (from the orientation of the median relative to the quartiles) and potential **outliers** (marked individually). An example is shown in Figure 1B.8.4, and advantages and disadvantages are listed in Box 1B.8.2.

Box 1B.8.2 Advantages and disadvantages of boxplots

Advantages	Disadvantages
The spacing between the different parts of the box can help indicate variance and skew, and identify outliers Boxplots are especially useful when comparing two or more sets of data	Exact values not retained

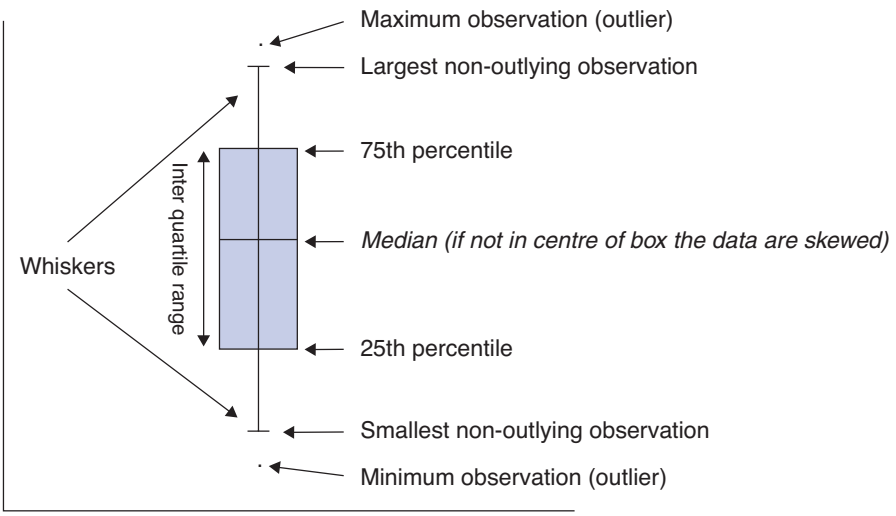


Figure 1B.8.4 Schematic of a boxplot

HISTOGRAMS

Histograms divide sample values into many intervals called 'bins'. The bars represent the number of observations falling within each bin (i.e. the bin's frequency). Examples are shown in Figure 1B.8.5. Note that, in contrast to bar charts, there are no gaps between the bars of a histogram. This is important because it is a reflection of the fact that the data are continuous. Box 1B.8.3 lists advantages and disadvantages of histograms.

Box 1B.8.3 Advantages and disadvantages of histograms

Advantages	Disadvantages
Demonstrates central tendency (mean, mode, medium) Demonstrates shape of frequency distribution (symmetrical or skewed, unimodal, bimodal or multimodal)	Cannot read exact values because data are grouped into categories More difficult to compare two data sets Use only with continuous data

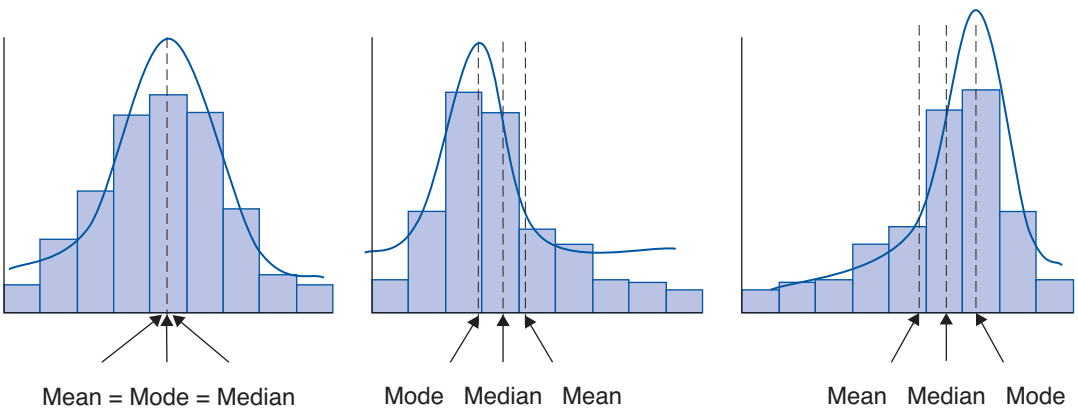


Figure 1B.8.5 Schematic of histograms, demonstrating shape of the frequency distribution

BIVARIATE DATA

Here there are two numerical variables. Bivariate data are best displayed using **scatter plots**.

SCATTER PLOT

The data from two variables are plotted one against the other to explore the association between them. A trend line is drawn to illustrate whether any correlation is:

- Positive, negative or non-existent
- Linear or non-linear
- Strong, moderate or weak.

The strength of any correlation is determined by observing the spread of the points about the line. Examples of scatter plots are shown in Figure 1B.8.6, and advantages and disadvantages are listed in Box 1B.8.4.

Box 1B.8.4 Advantages and disadvantages of scatter plots

Advantages	Disadvantages
Shows a trend in the data relationship	Hard to visualise results in large data sets
Retains exact data values and sample size	Flat trend line gives inconclusive results
Shows minimum, maximum and outliers	Data on both axes should be continuous

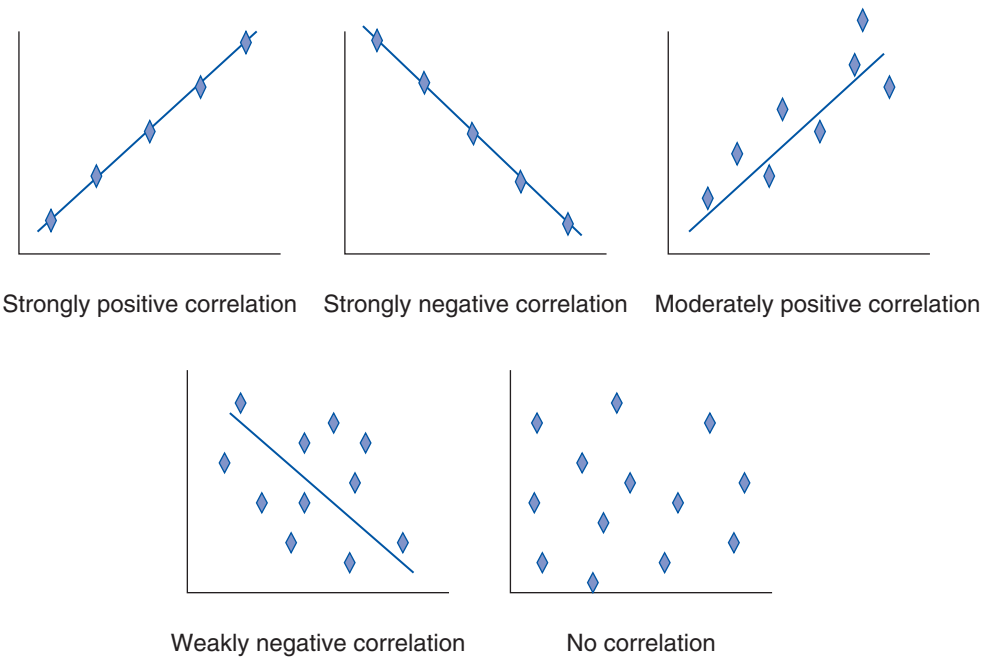


Figure 1B.8.6 Scatter plots displaying the correlation between variables

1B.9 HYPOTHESIS TESTING

To answer a statistical question, a **hypothesis** is first formulated. This is a statement that can be subjected to a test. The test will generate a probability of how likely it is that the observed outcome could have occurred due to chance alone. Depending on this probability, the hypothesis will be accepted or rejected.

STEP 1

Define a **null hypothesis** (H_0). This is a statement that there is no difference (or no relationship) between the variables being tested. For every null hypothesis, there is an **alternative hypothesis** (H_A), which assumes that a difference or relationship does exist. Both the null hypothesis and the alternative hypothesis are true/false statements that are answered in view of the significance level chosen (see Section 1B.10). For an example see Box 1B.9.1.

Box 1B.9.1

Example: Randomised controlled trial

H_0 = the rates of disease are the same in the intervention group and the control group

H_A = the rates of disease differ between the intervention group and the control group

STEP 2

Collect the data.

STEP 3

Calculate a **test statistic.**

A significance test consists of calculating the probability of obtaining a statistic as different or more different from the null hypothesis (given that the null hypothesis is correct) than the statistic obtained in the sample.

The *z* test for significance is conducted using the formula: $z = \frac{\mu}{\sigma/\sqrt{N}}$

where μ = mean, σ = SD and N = number of subjects in the sample.

STEP 4

Derive the ***p* value** for the test statistic from a known probability distribution (found in statistical tables – although note that these are not currently provided to candidates in the UK MFPH Part A examination).

STEP 5

Interpret the *p* value to **accept** or **reject** the null hypothesis, i.e. the outcome of hypothesis testing is always either '*reject* H_0' or '*do not reject* H_0'. See Box 1B.9.2.

If the null hypothesis is rejected, then the alternative hypothesis may be true. However, if the H_0 is not rejected, then this does not imply that the H_A is correct – only that there is insufficient evidence against the H_0.

Note that a very small *p* value (e.g. 0.001) does not signify a large effect. Rather, it signifies that the observed effect (whatever its size) is highly improbable given the null hypothesis. A very small *p* value can arise when an effect is small but the sample sizes are large. Conversely a larger *p* value can arise when the effect is large but the sample size is small.

Box 1B.9.2

Example: interpreting the *p* value (see also Section 1B.3)

If $p = 0.001$, there is 0.1% likelihood that, if the null hypothesis were true, this result would be obtained by chance alone. Therefore, there is **strong evidence** against the null hypothesis, suggesting that the observed effect is not due to chance.

If $p = 0.1$, there is 10% likelihood that, if the null hypothesis were true, we would obtain this result by chance. There is therefore **little evidence** to reject the null hypothesis.

1B.10 TYPE I AND TYPE II ERRORS

Findings from a study can lead to the null hypothesis being accepted or rejected. However, this acceptance or rejection may or may not reflect the 'true' situation, as seen in the contingency table shown in Table 1B.10.1, where the null hypothesis is that A = B.

Table 1B.10.1 Contingency table displaying type I and type II errors

		'True' finding	
		A = B	A ≠ B
Results from study	A = B	Correct	Type II error (β)
	A ≠ B	Type I error (α)	Correct

The magnitudes of α and β error rates are generally set in advance as the outcome of a power calculation that is used to determine the appropriate sample size of the study.

Table 1B.10.2 summarizes the differences between type I and type II errors.

Table 1B.10.2 Differences between type I and type II errors

	Type I error	Type II error
Definition	False positive	False negative
Description	Null hypothesis was wrongly rejected Study shows an effect which in reality does not exist	Null hypothesis was not rejected when it should have been Study does not detect an effect that existed in reality
Symbol	α	β
Alternative name	Significance level = α	Power = $1-\beta$
Typical value	0.05 or 0.001	0.8
Aide memoire	Type I error = *p* value	Type II (*two*) error = sample size *too* small

Type II errors are generally considered to be **less serious** than type I errors: a type II error is only an error in the sense that an opportunity to reject the null hypothesis was lost.

Note that there is a **trade-off** between type I and type II errors: the more the study restricts type I errors by setting a low level of α, the greater the chance that a type II error will occur.

SECULAR TRENDS

Erroneous findings may also result from temporal changes. A secular trend is a long-term change in the burden of a disease (as opposed to a short-term or cyclical fluctuation). It may be due to:

1. Changes in diagnostic techniques/diagnostic codes
2. Changes in accuracy/completeness of enumerating cases
3. Change in the age distribution of the population
4. Change in survival (which affects prevalence)
5. Change in the *true* incidence of the disease.*

**This reason should be considered only once items 1–4, and other important possible explanations, have been accounted for.*

1B.11 PROBLEMS OF MULTIPLE COMPARISONS

The hypothesis to be tested should always be stated **before** the study begins, together with the list of which variables will be analysed. The reason for this is that after the results have been collected, there may be a temptation to search for associations with additional variables. This is known as *'data mining'* or *'data fishing'*. As more and more tests are performed, the likelihood of a type I error increases, i.e. the chance of falsely concluding that an association is significant. This is because, at the $p = 0.05$ level of significance, 1 in 20 tests will appear significant due to chance alone.

To compensate for multiple comparisons, use a method to adjust the p value for the number of tests performed, such as **Bonferroni's correction**.

If testing n outcomes instead of a single outcome, divide α level by n:

i.e. Adjusted $\alpha = \dfrac{\text{Original } \alpha}{n}$

An example is shown in Box 1B.11.1.

Box 1B.11.1

Example: Bonferroni's correction

Suppose we were investigating the association between salt intake and 15 different types of cancer. We would test at:

Adjusted $\alpha = 0.05 \div 15 = 0.0033$ level.

This would ensure that the overall chance of making a type I error is kept at <1 in 20.

1B.12 PARAMETRIC AND NON-PARAMETRIC TESTS

Parametric and non-parametric tests for comparing two or more groups

Samples that are normally distributed can be described by two parameters: the mean and the standard deviation. A difference between samples or between a sample and the population may be measured by examining differences in their means. These differences can be tested using **parametric** tests, such as the ***z*-test** or the ***t*-test**.

Where a population is not normally distributed, its shape cannot be described by two parameters (the mean and standard deviation) alone, and any difference between two groups must be tested using a **non-parametric** test. The **Wilcoxon test** for paired data, and the **Mann–Whitney U-test** for ***u***npaired data are rank non-parametric tests used for **small** samples taken from a **non-normally distributed** population. If samples are large, then the lack of normality is not problematical and the above tests can be used.

The *z*-test and *t*-test are parametric tests used for comparing percentages, proportions and means of groups, with the *z*-test for large samples and the *t*-test for samples <60.

The chi-squared test is used for comparing observed versus expected numbers. If the numbers involved are small, then consider using **Fisher's exact test** instead or applying **Yates' correction**.

t-TEST

The 'unpaired' *t*-test may be used to compare one group to another group. Often one group is the experimental group and the other is the control. An example of an experiment that might utilise the *t*-test is a comparison of the mean blood pressure in 25 patients given a new β-blocker compared with 25 given a standard β-blocker, labetolol.

$$t = \frac{\bar{x}_1 - \bar{x}_0}{\text{se}}, \quad \text{df} = n_1 + n_0 - 2$$

$$t = \frac{\bar{x}_1 - \bar{x}_0}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_0}}}$$

where $\bar{x}$ = sample mean, n = sample size, s = pooled standard deviation, df = degree of freedom.

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_0 - 1)s_0^2}{n_1 + n_0 - 2}}$$

The paired *t*-test typically uses observations in one sample that have been paired with observations in another. This results in a different test statistic, where $\bar{x}$ represents the means of the paired differences. The usual hypothesis in this situation is that the mean of the differences is 0.

$$t = \frac{\bar{x}}{s/\sqrt{n}}, \text{ df} = n - 1$$

ONE-SAMPLE *t*-TEST

This tests the likelihood that a sample came from a population.

1. Mean of the population, μ, and the sample, $\bar{x}$, will both be known
2. Use the formula to calculate a value for the *t*-test $t = \frac{(\quad - \bar{x})\sqrt{n}}{\text{SD}}$

   i.e. $t = \frac{\text{Difference in means}}{\text{Standard error of sample mean}}$

3. Look up the *p* value in a *t*-table, using:

   Degrees of freedom = $n - 1$

An example is shown in Box 1B.12.1.

Box 1B.12.1

Example: One-sample *t*-test

In a population it is known that the mean finger-tapping time is 100 ms per tap

In a study of eight participants with caffeine addiction, tapping times were found to be:

- Mean = 89.4 ms
- Standard deviation = 20 ms.

Does this prove that caffeine addiction is associated with faster tapping speed? The null hypothesis (H_0) is that tapping speed is not affected by caffeine addiction. We should begin by pre-selecting a significance level at which we would be convinced of an effect, in this case 0.05.

$$t = \frac{(\mu - \bar{x})\sqrt{n}}{\text{SD}}$$

$$t = \frac{(100 - 89.4) \times \sqrt{8}}{20} = 1.5$$

Find the *p* value from statistical tables, i.e. the area below $t(1.5)$ using seven degrees of freedom, since df = n − 1.

This value of *p* is 0.07. Since this value is >0.05, there is little evidence to reject the null hypothesis. We therefore do *not* conclude that caffeine addiction is associated with faster tapping speed.

Reproduced with permission from Whiteley (2007).

TWO-SAMPLE (UNPAIRED) *t*-TEST

This tests the likelihood that two samples came from the same population:

1. Calculate the means for both groups (x_1 and x_2)
2. Calculate the pooled variance $\sqrt{\frac{(n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2}}{n_1 + n_2 - 2}}$
3. Calculate the standard error of the difference in means = $s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
4. $t = \frac{\text{Difference in means}}{\text{Standard error of difference in means}} = \frac{(x_1 - x_2) - 0}{\text{SE}(x_1 - x_2)} = \frac{(x_1 - x_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$
5. Look up *t* value (using degrees of freedom = $n_1 + n_2 - 2$) in a *t*-table to find the *p* value.

An example is shown in Box 1B.12.2.

Box 1B.12.2

Example: Unpaired *t*-test

In an RCT, one group of patients is given an inhaled steroid while the other group is given a placebo. The two groups are compared at 6 months with regard to their forced expiratory volume (FEV_1).

Is there a significant difference in FEV_1 at 6 months between the two groups?

Treatment group 1 receive inhaled steroid [$n_1 = 50$; mean FEV_1 (x_1) = 1.64 l; $SD_1 = 0.29$ l]

Treatment group 2 receive placebo [$n_2 = 48$; mean FEV_1 (x_2) = 1.54 l; $SD_2 = 0.25$ l]

An unpaired *t*-test is performed to compare means in the two groups:

- Null hypothesis H_0: mean FEV_1 is same in the two groups, i.e. no difference
- Alternative hypothesis H_A: FEV_1 is not the same in the two groups, i.e. there is a difference.

$$\text{Pooled SD of the two groups} = \sqrt{\frac{(n_1-1)s_1^{\,2}+(n_2-1)s_2^{\,2}}{n_1+n_2-2}} = \sqrt{\frac{(50-1)0.29^2+(48-1)0.25^2}{50+48-2}} = 0.2712$$

$$t = \frac{(x_1-x_2)-0}{\text{SE}(x_1-x_2)} = \frac{(x_1-x_2)}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(1.64-1.54)}{0.2712\sqrt{\frac{1}{50}+\frac{1}{48}}} = 1.8251$$

$$\begin{aligned}\text{Degrees of freedom} &= n_1 + n_2 - 2\\ &= 50 + 48 - 2\\ &= 96\end{aligned}$$

According to the *t*-distribution table for 96 degrees of freedom, $p > 0.05$.

H_0 is therefore accepted at the 5% level and we conclude that there is no difference in the mean FEV_1 between the two groups at 6 months.

Reproduced from Petrie and Sabin (2005).

PAIRED *t*-TEST

Used for matched studies (e.g. case–control trials and matched-pair RCTs):

1. Calculate the mean of the differences
2. Find the standard deviation of the differences
3. Calculate the standard error of the mean = $\frac{\text{SD}}{\sqrt{n}}$
4. t = Mean of the differences ÷ Standard error of mean of the differences
5. Look up t value (using degrees of freedom = number of pairs − 1) in a *t*-table to find the p value.

z-SCORES

An observation from a population may be converted into a **standard normal deviate** (also called a *z*-score or normal score). This is the number of standard deviations that separate the observation from the mean of the population. It is calculated by subtracting the population mean from an individual ('raw') score and then dividing the difference by the population standard deviation.

The *z*-score reveals how many units a case is above or below the mean. The *z*-score allows comparisons to be made between the results of different normal distributions.

$$z = \frac{\bar{x} - \mu}{\sigma}$$

where $\bar{x}$ = raw score to be standardised, μ = population mean, σ = population standard deviation.

The *z*-test is the generalisation of the *t*-test for larger numbers, where $n > 60$. Both *t*-tests and *z*-tests are used to compare two groups. When more than two groups need to be compared, an analysis of variance (**ANOVA**) is used – see below.

In the same way that the *t*-test can be used to examine differences in the means of two smaller populations, the *z*-test may be calculated as follows:

$$z = \frac{\text{Difference in means}}{\text{SE of the difference in the means}} + \frac{\bar{x}_1 - \bar{x}_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_0^2}{n_0}}}$$

The test statistic for the paired situation is:

$$z = \frac{\bar{x}}{s/\sqrt{n}}$$

Z-TEST FOR THE DIFFERENCE BETWEEN TWO PROPORTIONS

When testing for a difference in **proportions**, assume the null hypothesis that both groups come from the same population and have the same proportions (*P*) of the characteristic of interest.

1. Combine the two groups to find the overall proportion, *P*
2. Standard error (difference) $= \sqrt{\frac{P(1-P)}{n_1} + \frac{P(1-P)}{n_2}}$
3. Calculate the difference in proportions between the two groups ($P_1 - P_2$)
4. *z* = Difference in proportions ÷ Standard error of the difference
5. Look up *z* value in the normal distribution table to find the *p* value (note that this is identical to the chi-squared test).

CHI-SQUARED TEST

This test can be used only for **actual** numbers (not for proportions, percentages, etc.). Note that the Greek letter (χ^2) is used for the test and distribution, but that the Latin letter (X^2) is used for the calculated statistic.

1. For each observed number calculate the expected number
2. Subtract the expected number from the observed number ($O - E$)
3. Square the result and divide this by the expected number $(O - E)^2 \div E$
4. X^2 = total of these results for all cells, i.e. sum of (3)
5. Look up X^2 (using degrees of freedom = [rows − 1] × [columns − 1]) to find the *p* value.

An example is shown in Box 1B.12.4.

Box 1B.12.4

Example: Chi-squared test

The paté at a restaurant is implicated in an outbreak of listeriosis.

Here are the **observed** numbers of guests who ate paté, and the numbers who were ill with listeriosis:

		Listeriosis		
		Yes	No	
Ate paté	Yes	12	24	36
	No	42	48	90
	Total	54	72	126

First calculate the **expected** number for each cell assuming that paté was **not** linked to listeriosis, e.g. $36 \times 54 \div 126 = 15.43$

		Listeriosis		
		Yes	Yes	
Ate paté	Yes	15.43	20.57	36
	No	38.57	51.43	90
	Total	54	72	126

Now calculate $\frac{(O-E)^2}{E}$ for each cell

		Listeriosis	
		Yes	Yes
Ate paté	Yes	0.76	0.57
	No	0.31	0.23

Now calculate $\sum \frac{(O-E)^2}{E} = 0.76 + 0.57 + 0.31 + 0.23$

$= 1.87$

Note that statistical tables are not provided in the Part A examination. However, for 2×2 tables (i.e. 1 degree of freedom), the cut-off point for the 95% significance level is 3.84 (i.e. 1.96^2). As 1.87 is <3.84, we can conclude that eating paté was not significantly associated with listeriosis at the $p<0.05$ level.

ANOVA

The same principles as the *t*-test (and *z*-test) are used in the **an**alysis **o**f **va**riance (ANOVA), which is employed when more than two groups need to be compared. In this situation, an *f*-test is used and the null hypothesis is that the means of all the groups of observations are equal.

ANOVA is called one-way analysis of variance. For example, if three groups of animals are treated separately with hormone 1, hormone 2 and saline solution (control) to assess whether there are differences in the resultant growth, then the 'factor' concerned here is hormone and this is a 'one-way' or 'one factor' ANOVA.

Two factors can also be tested with ANOVA. Hormones may affect male animals to a greater or lesser extent than females. By selecting samples of females and samples of males for each treatment, sex then becomes 'factor 2'. Comparisons could then be made between the different hormones and between the different sexes.

A second extension of one-way ANOVA is made when comparing two dependent variables simultaneously across two or more groups. This extension is called multivariate analysis of variance (**MANOVA**). For example, the means of dependent variables (reading, writing, IQ, maths) may be tested across two groups (males, females).

An example is shown in Box 1B.12.5.

Analysis of variance is a special type of regression analysis, and most data sets for which analysis of variance is appropriate can be analysed by regression with the same results. With two groups one-way ANOVA is exactly equivalent to the usual two-sample t-test, and we have $f = t^2$.

WILCOXON'S TEST

This is used for **paired** data with differences, when plotted, that look roughly symmetrical.

1. Find differences between individual pairs
2. Omit zero values
3. Ignoring the signs (+ or −) for the moment, rank the differences in order, placing identical values halfway between the ranks that they would have occupied if they were unique
4. Reapply the signs (+ or −)
5. Find the sum of the positive ranks and the sum of the negative ranks (these are called the 'rank totals')
6. Ignoring the signs again, take the smaller 'rank total' from (5)
7. Look up this 'rank total' in the Wilcoxon table (either the 1% or the 5% section) according to the number of pairs in the sample
8. If the 'rank total' is larger than the number in the table then the result is insignificant at that p value.

MANN–WHITNEY U-TEST

This is used for **unpaired** data: where the data do not appear symmetrical.

1. Rank the results from both groups in a single list (writing each in a different colour)
2. Add up the ranks for the two samples separately (by reading off the different colours): these are the 'rank totals'
3. Use the smaller of the 'rank totals'
4. Look up in the Mann–Whitney table (either the 1% or the 5% section) using n_1 as the number of observations in one group and n_2 as the number of observations in the other
5. If the number from (3) is larger than the number found in (4), then the result is insignificant at this p value.

INTER-RATER RELIABILITY

A common situation in epidemiology arises when two people measure the same phenomenon or object. The **kappa statistic** takes account of the possibility of such chance agreement, and indicates how closely two different measurements (binary or nominal) are aligned. Table 1B.12.1 is a contingency table for the kappa statistic.

Box 1B.12.5

Example: one-way ANOVA

A researcher predicts that students will learn most effectively with a constant background sound, as opposed to an unpredictable sound or no sound at all. Twenty-four students are divided at random into three groups of eight. All students study a passage of text for 30 min. Those in group 1 study with sound at a constant volume in the background. Those in group 2 study with noise that changes volume periodically. Those in group 3 study with no sound at all. After studying, all students take a 10-point multiple choice test on the material.

- Null hypothesis H_0 = no difference between the groups exposed to constant, random or no sound
- Alternative hypothesis H_1 = there is a difference between the groups

Group	Test scores							
Constant sound	7	4	6	8	6	6	2	9
Random sound	5	5	3	4	4	7	2	2
No sound	2	4	7	1	2	1	5	5

1. Calculate the total sum of squares for all the scores, x; and the participants, n:

$$SS_{total} = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

2. Calculate the sum of squares due to the difference between the observations:

$$SS_{between} = \Sigma n_i x_i^2 + \frac{(\Sigma x)^2}{n}$$

3. Calculate the sum of squares f within groups

$$SS_{within} = SS_{total} - SS_{between}$$

$$SS_{total} = (322 + 148 + 125) - \frac{(48 + 32 + 27)^2}{24} = 117.96$$

$$SS_{between} = \left[\frac{2304}{8} + \frac{1024}{8} + \frac{729}{8}\right] - \frac{(48 + 32 + 27)^2}{24} = 30.8$$

4. Calculate (a) the mean square for $SS_{between}$, where df = No. of groups –1

$$MS_{between} = \frac{SS_{between}}{df}$$

(b) The mean square for SS_{within}, where df = No. of participants – No. of groups

$$MS_{within} = \frac{SS_{within}}{df}$$

Source	SS	df	MS	f
Between	30.8	2	15.4	3.68
Within	87.88	21	4.18	

$SS_{within} = 117.96 - 30.8 = 87.88$

Table *contd overleaf*

Box 1B.12.5 *contd*

5. Calculate the *f*-test statistic

$$f = \frac{MS_{between}}{MS_{within}}$$

$f = 3.59$

With df = (2,21), reference tables tell us that *f* must be at least 3.4668 to reach $p < 0.05$, so the observed *f* score is statistically significant.

Adapted with permission from Hall (2007).

Table1B.12.1 Contingency table for kappa statistic

		Observer 1		
		No	Yes	**Total**
Observer 2	No	*a*	*b*	***a+ b***
	Yes	*c*	*d*	***c + d***
Total		***a + c***	***b + d***	***a + b + c + d***

$$I_e = \text{Predicted agreement} = \frac{(a+c)(a+b)+(b+d)(c+d)}{(a+b+c+d)^2}$$

$$I_o = \text{Observed agreement} = \frac{a+d}{a+b+c+d}$$

$$\kappa = \frac{I_o - I_e}{1 - I_e}$$

The agreement due to chance varies according to the proportions of results that are reported as positive or negative. The kappa value is interpreted as shown in Box 1B.12.6.

Box 1B.12.6

Kappa value	Degree of agreement beyond chance
0	None
0–0.2	Slight
0.2–0.4	Fair
0.4–0.6	Moderate
0.6–0.8	Substantial
0.8–1.0	Almost perfect

Kappa has some limitations:

- Using kappa to assess agreement assumes that the raters are independent
- It is useful only when raters are using the same rating scale
- It tells the reader nothing about the reasons for variation.

An example of the use of the kappa statistic is shown in Box 1B.12.7.

1B.13 SAMPLE SIZE AND STATISTICAL POWER

Sample size and power calculations allow the researcher to decide during the design phase of an experiment:

- How large a sample is needed to obtain accurate and reliable results
- How likely it is that the statistical test to be used will detect effects of a given size in a particular situation.

Box 1B.12.7

Example: kappa statistic

Two doctors report results of 29 patients. They agree with each other in 22 cases (75.9%), i.e. 10 + 12. From the table below, we can see that the resultant kappa statistic is 0.542, representing moderate agreement.

		Doctor A		
		No	**Yes**	**Total**
Doctor B	**No**	10 (34.5%)	7 (24.1%)	**17 (58.6%)**
	Yes	0 (0.0%)	12 (41.4%)	**12 (41.4%)**
Total		**10 (34.5%)**	**19 (65.5%)**	**29**

$I_o = 0.759$

$I_e = 0.473$

kappa = 0.542

SAMPLE SIZE

Sample size calculations are a key part of the study design because they ensure that studies have sufficient participants to answer the question posed but not so many so that resources are wasted unnecessarily. In general, there are four factors to consider when calculating the required sample size: see Table 1B.13.1.

Table 1B.13.1 Factors in sample size calculations

Factors	Effect on sample size
Size of difference • Effect size needed for result to be clinically/experimentally meaningful	Smaller effect size requires larger sample
Significance level • p value, i.e. type I error • Usually 0.05 is used	Smaller p requires larger sample
Power • Probability that the study will be able to detect a difference if it really exists: related to type II error ($1 - \beta$) • Usually a power of 80% is used	Higher power requires larger sample
Exposure in baseline population • For case–control study: exposure in controls • For cohort/intervention: disease rate	

POWER

The power of a study to detect a true effect is generally set at 80% or more. If the power of an experiment is low, then there is a high probability that the experiment will be inconclusive. Power can be increased at the expense of significance. In case–control studies, power can be increased for the same number of cases by increasing the ratio of controls to cases, although this too will increase the overall sample size.

Different methods are used for estimating power in the various study designs. In all cases, power should be calculated in advance, and then a sample of appropriate size is recruited. However, for logistical or financial reasons this is not always possible. In the case of very rare conditions, for example, the sample size may be fixed. It can then be used to calculate the power of the study (i.e. to assess the likelihood of detecting a statistically significant effect).

Meta-analyses can be conducted to pool the results from several studies and thereby increase the power of detecting a finding if one exists (see Section 1A.33).

OTHER FACTORS TO CONSIDER

The sample size should be **increased** when:

- Loss to follow-up or a low response rate is expected
- Cluster sampling is used (see Section 1A.21)
- Confounding is expected (i.e. several variables will need to be controlled for)
- Interaction is expected.

The sample size should be **decreased** when:

- **Matched** case controls are used

Note that sample size can reduce only any errors caused by chance: it cannot compensate for bias.

1B.14 REGRESSION AND CORRELATION

Regression and correlation are related statistical techniques. Both examine the relationships between two or more variables, but they do so in different ways.

REGRESSION

Regression is the process of deriving an equation for the best fitting line through the points on the scatter diagram. This line can be found by minimising the squared distances between points and the line (known as the **least squares technique**). The regression equation can be used to determine by how much one variable (y) changes when another variable (x) changes by a certain amount. However, the equation is valid only between the limits of the scatter diagram: to extrapolate beyond this may be invalid. Regression can be used for explanatory or predictive purposes.

MULTIPLE REGRESSION

Multiple regression is used for **continuous** variables:

- Assume that the line can be described by its intercept (c) and slope – called the 'regression coefficient' (m)
- If necessary, divide up the line into multiple components: $y = c + mx \rightarrow y = c + m_1x_1 + m_2x_2 + m_3x_3$
- The coefficients (m_i) are calculated using the least squares technique
- The size of each coefficient (m_i) represents the effect size for that variable.

LOGISTIC REGRESSION

Logistic regression is used for **binary** variables:

- Take natural log of odds (called the 'logit') which will always be between 0 and 1
- ln (odds of disease) = $(y/(1-y)) = c + m_1x_1 + m_2x_2 + m_3x_3$.

Tests for the significance of the slope give identical results to tests for the significance of the correlation coefficient.

CORRELATION

Correlation estimates the **strength** of any linear association between two variables. Unlike regression, correlation is **bidirectional**, i.e. when one variable (x) changes by a given amount, so the other variable (y) changes by the proportional amount.

The strength of the association can be represented by **Pearson's correlation coefficient (r)** using a formula based on the sample size and the values of x and y.

The value of r can vary from −1 through 0 to +1. Box 1B.14.1 explains the interpretation of these values.

Box 1B.14.1

r	Implication
+1	Perfect positive correlation
Positive values	When one value increases, the other variable increases
0	No correlation
Negative values	When one value increases, the other variable decreases
−1	Perfect negative correlation

The association can be tested using a form of the t-test (based on r and n) at degrees of freedom $= n - 2$:

$$t = \sqrt{\frac{n-2}{1-r^2}}$$

Pearson's correlation coefficient is a parametric statistic. Non-parametric correlation methods, such as **Spearman's ρ** (rho), are used when distributions are not normal.

1B.15 REGRESSION TECHNIQUES

Appropriate use, objectives and value of multiple linear regression, multiple logistic regression, principles of life-tables and Cox's regression

Table 1B.15.1 summarises the usage of the different regression techniques.

REGRESSION LINE EQUATIONS

Simple linear regression: $y = a + bx$

Multiple linear regression: $y = a + b_1x_1 + b_2x_2 +$, etc

where a = constant, x = explanatory variable, y = intercept, b = gradient (regression coefficient, i.e the increase in y that corresponds to a unit change in x).

LIFE-TABLES

See Section 3A.6.

Table 1B.15.1 Regression techniques

Univariate regression methods	Appropriate use	Objectives
Linear regression	Describe relationship between two numerical variables Estimated using statistical technique of fitting straight line (line of best fit) to scatter plot of variables using the method of **least squares**	Provides a line of best fit through data points. Calculate relationship between an outcome and an explanatory variable: the **correlation coefficient**, r. Values range from $r = -1$ (perfect negative correlation), through $r = 0$ (no correlation) to $r = 1$ (perfect positive correlation)
Multivariate regression methods	**Appropriate use**	**Objectives**
Multiple linear regression	Analysis of **several explanatory** variables with **one outcome** variable	Adjusts for confounding Understand which variables are associated with an outcome
Multiple logistic regression	Analysis of **several explanatory** variables with a **binary outcome** variable	Adjust for confounding
Cox's regression (see Section 1B.16)	Analysis of the effect of several variables upon the time it takes for a specified event to happen	

1B.16 COMPARISON OF SURVIVAL RATES

Two methods can be used to generate survival functions and survival curves: see Box 1B.16.1.

Box 1B.16.1

Life-tables	Lead to smooth survival curves (see Section 3A.6)
Kaplan–Meier estimates	Produce stepped survival curves Used when the time of an event and follow-up time are known ('censoring')

Differences in survival rates between two groups can be apparent from survival curves, but quantifying the differences requires statistical methods.

LOG-RANK TEST

Comparison of life-tables can be made visually by constructing survival curves, or assessed quantitatively by the log-rank test. This is a special application of the Mantel–Haenszel chi-squared procedure, and is carried out by

constructing a separate 2×2 table for each interval of the life tables to compare the proportions dying during each interval.

The log-rank test is used to test the null hypothesis that there is no difference between the populations in the probability of an event (here a death) at any time point. The analysis is based on the times of events (such as deaths). For each such time, the observed number of deaths in each group is calculated, together with the number that would be expected if there were no difference between the groups.

The log-rank test is most likely to detect a difference between groups when the risk of an event is consistently greater for one group than another. It is unlikely to detect a difference when survival curves cross, as can happen when comparing a medical with a surgical intervention. When analysing survival data, the survival curves should therefore always be plotted.

Because the log-rank test is purely a test of significance, it cannot provide an estimate of the size of the difference between the groups, nor a confidence interval. For these some assumptions must be made about the data. Common methods used are the hazard ratio, including the Cox proportional hazards model.

COX'S REGRESSION

This is also known as **proportional hazards regression** (see also Section 1B.15).

This form of **multivariate regression** is based on the proportional hazards assumption, i.e. that the ratio of hazards (the instantaneous risk of dying at time t) in both groups remains the same. It is a non-parametric technique that makes full use of survival data without making assumptions of survival curve shape. It adjusts for confounders, i.e. it tests the effects of a number of explanatory variables on the hazard. Cox's regression can assess only the effects of one factor on the time to endpoint.

Relative hazard is interpreted in a similar way to a relative risk (i.e. a value of 1 indicates no difference; values >1 indicate a raised hazard; values <1 a decreased hazard).

$$\lambda_i(t) = \lambda_0(t)\exp\{\beta_1 x_1 + \beta_2 x_2 + \ldots \beta_n x_n\}$$

$$\lambda_i(t) = \text{hazard for individual } i \text{ at time } t$$

where $\lambda_0(t)$ = baseline hazard (not usually interested in), β = coefficients, x = covariates associated with individual i.

1B.17 HETEROGENEITY

Heterogeneity refers to differences between observations that would lead an investigator to consider that the observations might have been drawn from different (i.e. heterogeneous) populations. It is often used in a **systematic review** to assess whether the pooling of data is reasonable and how meta-analyses should be analysed (see Section 1A.33). There are three types of heterogeneity: statistical, methodological and clinical. See Table 1B.17.1.

1B.18 FUNNEL PLOTS

The funnel plot (Figure 1B.18.1) is a type of scatter plot that is commonly used to visualise potential **publication bias** in meta-analyses (Egger 1997) (see Section 1A.33). It is also increasingly being used in other situations, such as comparing performance between several organisations.

For a meta-analysis, the estimated **treatment effect** values from the component studies are plotted against a measure of **precision** (study size or standard error). Appropriate confidence intervals are then added. As the sample sizes of the component studies increase, so the precision of their estimated treatment effect increases. This gives

the graph its funnel shape. So, in the absence of bias, the results from small studies should scatter widely at one end of the graph, with the spread narrowing among larger studies.

Table 1B.17.1 Types of heterogeneity

Statistical heterogeneity	This relates to differences in the reported effects between studies, calculated using the chi-squared test for heterogeneity. Statistical heterogeneity can be accounted for by using a random effects model, as opposed to a fixed effects model in the analysis (see Section 1A.33)
Methodological heterogeneity	This is due to differences in study design and presents comparability problems for meta-analysis
Clinical heterogeneity	This relates to differences between studies relating to: • characteristics of the patient population • interventions, or • outcome measures and • length of follow-up As with methodological heterogeneity, clinical heterogeneity may preclude the pooling data from these studies

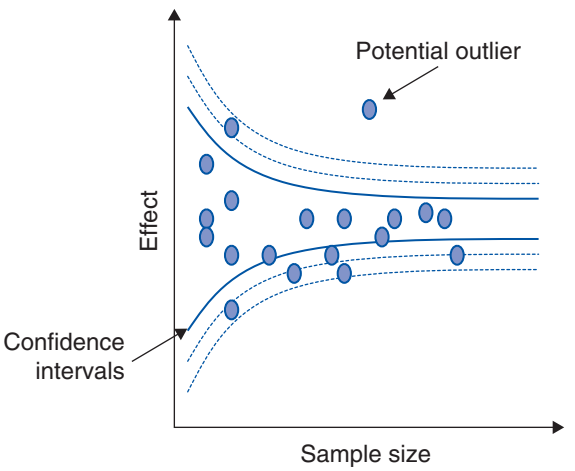


Figure 1B.18.1 Funnel plot

INTERPRETING FUNNEL PLOTS

Funnel plots demonstrate any biases graphically by means of the shape of the plot, although formal statistical tests can also be used.

If the plot of a meta-analysis is symmetrical, then this suggests that there is no publication bias. If the plot is asymmetrical then publication bias is a possibility. However, asymmetry could also be due to small study effects (the tendency for the smaller studies in a meta-analysis to show larger treatment effects).

In a plot of **performance** (e.g. surgical complication rates) those institutions plotted outside the confidence interval envelopes are revealed as outliers who may warrant further investigation.

1B.19 ROLE OF BAYES' THEOREM

Bayesian methods incorporate **prior beliefs** into calculations of probability. In real clinical situations, for example, existing knowledge about a particular patient will affect how much credence clinicians place on a laboratory test performed on that patient. Bayesian methods incorporate this prior knowledge into the probability calculations.

According to Bayes' theorem, the probability of A occurring, given B, depends on three factors. See Table 1B.19.1.

Table 1B.19.1 Bayesian probability

Factor	Alternative name	Symbol	Description
Probability of A	Prior	$P(A)$	Probability of A occurring on its own, irrespective of B
Probability of B	Normalising constant	$P(B)$	Probability of B occurring on its own, regardless of A
Probability of B, given A	Likelihood	$P(B\|A)$	Probability of B occurring given that A occurred

Using these three measures, the probability of A occurring given that B occurred is computed as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Advantages and disadvantages of bayesian statistics compared with classical statistics are listed in Box 1B.19.1.

Box 1B.19.1

Advantages	More flexible Makes use of all available knowledge, therefore possibly more **ethical** Mathematics is **not** controversial
Disadvantages	Different users will obtain **different conclusions** if they choose different priors

DIAGNOSTIC TEST IN A BAYESIAN FRAMEWORK

Posterior odds of disease = prior odds × likelihood ratio of a positive test result

$$\text{Prior odds} = \frac{\text{Prior probability}}{1 - \text{Prior probability}}$$

$$\text{Likelihood ratio} = \frac{\text{Sensitivity}}{1 - \text{Specificity}}$$

$$\text{Posterior probability} = \frac{\text{Posterior odds}}{1 + \text{Posterior odds}}$$

An example of the use of bayesian statistics is shown in Box 1B.19.2.

Box 1B.19.2

Example: bayesian statistics for a diagnostic test

Likelihood ratio for a positive cytomegalovirus (CMV) test = 13.3

Prevalence of CMV infection after bone marrow transplantation ~ 33%

Prior probability of severe disease = 0.33

$$\text{Prior odds} = \frac{\text{Prior probability}}{1 - \text{Prior probability}} = \frac{0.33}{0.67} = 0.493$$

$$\text{Posterior odds} = 0.493 \times \text{likelihood ratio} = 0.493 \times 13.3 = 6.557$$

$$\text{Posterior probability} = \frac{6.557}{1+6.557} = 0.868$$

87% chance of developing severe disease with CMV after transplantation, which is greater than the pre-test probability, indicating usefulness of the test.

Reproduced from Petrie and Sabin (2005).

1B.20 CHOICE OF STATISTICAL TEST

Figures 1B.20.1–1B.20.4 illustrate appropriate tests for particular types of outcome variables. Reproduced with permission from Kirkwood and Sterne (2003).

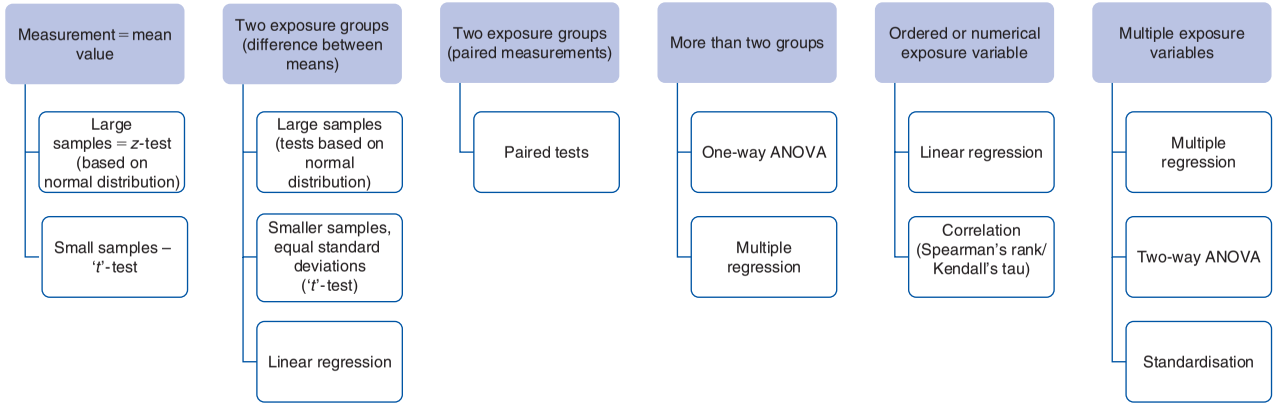


Figure 1B.20.1
Numerical outcome variable

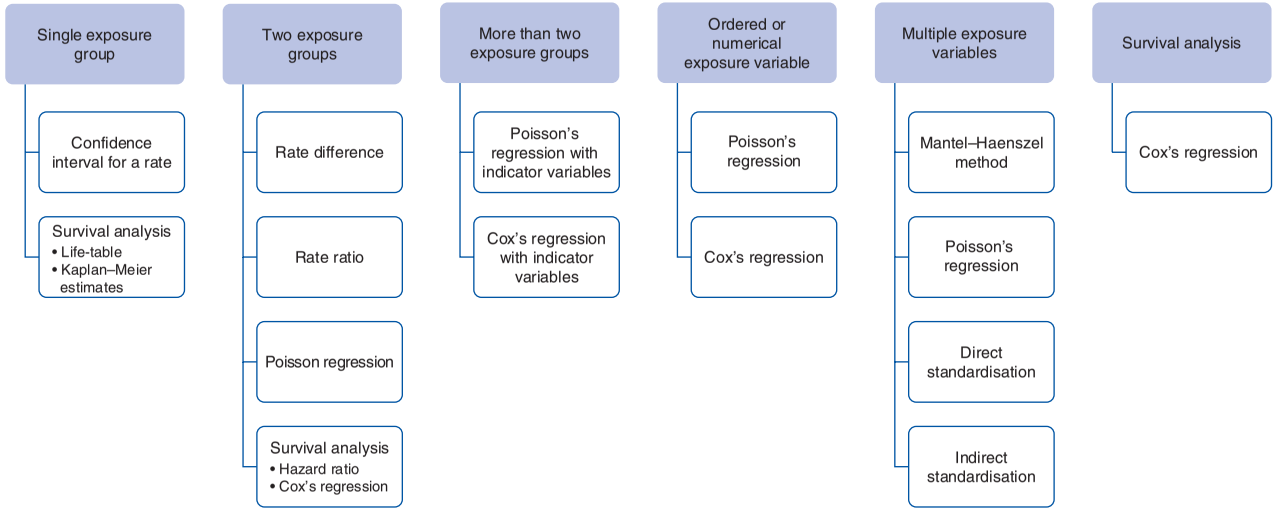


Figure 1B.20.2
Rates and survival times

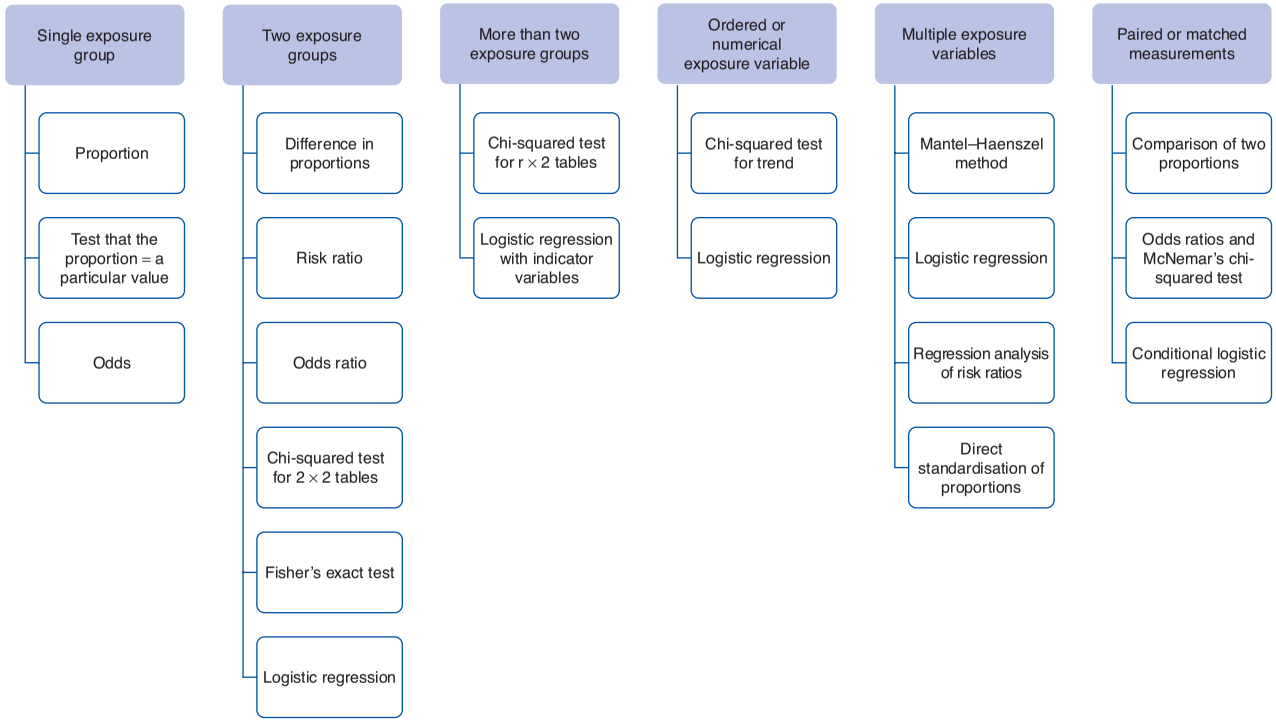


Figure 1B.20.3 Binary outcome variable

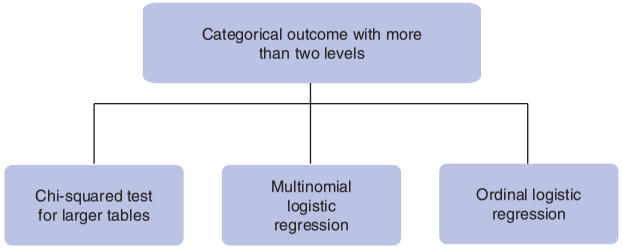


Figure 1B.20.4 Categorical outcome with more than two levels

1C
Assessment and Evaluation

Approaches to the assessment of health-care needs, utilisation and outcomes, and the evaluation of health and health care

Just because a particular health service is provided, it does not necessarily mean that it is effective, or that it is appropriate for the population served. Public health has an important role in assessing the need for a service, and the quality and appropriateness of what is currently provided. Practitioners need both knowledge of research methods and practical skills in order to:

- Measure patients' health and illness, quality of life and living conditions
- Study the nature of health services
- Involve relevant practitioners, patients and communities in making assessments and effecting improvements.

1C.1 NEED FOR HEALTH SERVICES

Uses of epidemiology and other methods in defining health service needs and in policy development

In public health, the generic term 'need' is used to indicate a number of concepts relating either to the illnesses that people experience (need for **health**) or to their treatment (need for **health care**). Bradshaw (1972) described four specific types of need: see Box 1C.1.1.

Box 1C.1.1

Type of need	Description
Normative	Needs as deemed by a clinician
Felt	Requirement of patients to feel better
Expressed	Demand, e.g. visits to a GP
Comparative	Need in one area compared with that in another

Strictly speaking, people are only said to 'need' interventions that can benefit them. Therefore, if an intervention offers no apparent clinical benefit, then there can be no clinical need for it.

A key role of a public health practitioner is to assess the health needs of a given population. When conducted systematically, this process is known as a **health needs assessment** (HNA). The findings of an HNA can be used to guide the allocation of resources in order to improve the health of the population and to reduce health inequalities. It offers an opportunity to:

- **Consult** the population
- Cultivate cross-sectoral **partnerships**
- Develop new **interventions**
- Ensure that health-care provision is **evidence based**.

CORPORATE NEEDS ASSESSMENT

In this type of HNA, the practitioner considers the views of interested parties, and aims to tailor local provision of health care according to these opinions. See Box 1C.1.2.

Box 1C.1.2

Advantages	Disadvantages
• Being an **incremental** process, it allows services to be altered a little at a time, guided by feedback from interested parties • Can be done **quickly** • No need to collect many data • **Responsive** to interested parties	• Can be driven by **power** rather than by need • Can be **disproportionately influenced** by the reaction of participants to certain events, e.g. newspaper headlines • An incremental process may be inappropriate where **large-scale** or **radical reform** is required

COMPARATIVE NEEDS ASSESSMENT

This process uses data from surveys or hospital activity data. It compares findings **observed** locally with those that would be **expected** from a reference population (e.g. regional or national data). For the comparisons to be valid, adjustments to the reference population need to be made in order to reflect the characteristics of the local population (e.g. age, ethnicity, smoking rates, etc.). Note that local prevalence may not be known, e.g. many cases of chlamydia infection will be undetected.

EVIDENCE-BASED NEEDS ASSESSMENT

Evidence from the literature – ideally in the form of guidelines or consensus statements – should be used to shape local provision of health care. Provision should be adjusted to take account of local circumstances such as the

prevalence of a disease in the local population and the findings of patient surveys. Note that the literature may not exist or it may be conflicting.

DATA SOURCES FOR CONDUCTING A NEEDS ASSESSMENT

The data to estimate need may relate to **activity** or **epidemiology** (prevalence/incidence): see Table 1C1.1.

Table 1C.1.1 Data for a needs assessment

Type of data	Problems with data sources
Activity	Data sources such as hospital episode statistics (HES) **underestimate** need since they do not include: • **Unmet need** (i.e. people who do not seek treatment) • People who seek **private** treatment
Epidemiology	Where local epidemiological data exist, they may **overestimate** need if it is assumed that prevalence equals need. This is because some of the people with the condition: • Do **not need** treatment (e.g. hip patients not fit to undergo surgery) • Do **not want** treatment • May have **already had** treatment Some conditions that exist on a spectrum require an arbitrary cut-off (e.g. blood pressure)

STEPS INVOLVED IN A NEEDS ASSESSMENT

The UK's Health Development Agency (now part of the National Institute for Health and Clinical Excellence) outlined a five-step approach to health needs assessment, as shown in Table 1C.1.2.

Table 1C.1.2 Steps involved in a needs assessment

1. Scope	Identify: • Population to be considered • Aims of the HNA (health needs assessment) • Stakeholders that will be involved • Resources needed to conduct the HNA • Risks involved
2. Identification of potential priorities	Gather data to **describe the population** (activity data, epidemiology and opinions from surveys)
3. Selection of a priority	Choose a disease according to its **health burden** Select interventions that are **effective** and **acceptable**
4. Change	Specify the aims of the intervention. Use **change management** techniques (action planning, monitoring, risk management)
5. Review	**Learn** from the project by measuring impact, then disseminate findings and choose the next priority

Adapted from Cavanagh and Chadwick (2005).

1C.2 PARTICIPATORY NEEDS ASSESSMENT

HNAs are traditionally conducted by commissioners of health services or by a third party. As a result, the local population may sometimes be reluctant to accept its findings and conclusions. In order to improve engagement, the local community should participate in the HNA process. For this to occur effectively, the participatory HNA requires:

- Clear **objectives**
- Use of accepted **methods** and data sources
- Support from appropriate **experts**
- Effective **communication** techniques
- Views of **groups whose voices are not normally heard** to be taken into account during the HNA (e.g. ethnic minorities, older people, children, young mothers and transient populations)
- Involvement of the community in **all steps** of the HNA process.

Methods in this type of HNA often produce **qualitative** rather than quantitative data. Such methods of data generation may include:

- Key informant interviews
- Group workshops
- Focus groups
- Visual methods (e.g. mapping or 'transect walks' where groups are given disposable cameras to take pictures that illustrate their needs).

If done well, the participatory process will identify potential needs that service providers would not have recognised without the input of community members.

1C.3 SERVICE UTILISATION AND PERFORMANCE

Formulation and interpretation of measures of utilisation and performance

With such a high proportion of the GDP being spent on health care, it is incumbent on health services to account for how this money is spent (who is accessing a service), and to demonstrate that it is achieving an impact (the extent to which a service is meeting its stated aims).

Health service utilisation statistics are usually used to measure utilisation of health services, but other dimensions of performance include:

- Quality
- Safety
- Effectiveness
- Patient satisfaction
- Waiting times.

WAITING TIMES

Waiting times are indicators of barriers to the use of services. They reflect poor responsiveness of services in meeting demand, and may be due to:

- Insufficient **capacity**
- Poorly designed **care pathways** that fail to manage demand (either referring inappropriately or referring too early), or are unnecessarily long and complex
- Failure to deploy resources in line with **fluctuations** in demand.

In many health-care systems, targets have been introduced to reduce these barriers. Box 1C.3.1 gives examples of waiting time targets in England.

Eng Box 1C.3.1

Service or procedure	Target maximum wait
Emergency	A&E: maximum waits of 4 h Ambulances to respond to 75% of category A calls within 8 min
Cancer	Urgent GP referral to treatment: 1 month Urgent GP referral to first outpatient appointment: maximum wait of 2 weeks
Scheduled operations	Maximum wait of 6 months for inpatients
Outpatient appointments	Maximum wait of 3 months
Primary care	Percentage able to offer GP appointment within 48 h of the request

Reproduced from Department of Health (2004), National Standards, Local Action: Health and social care standards and planning framework 2005/06–2007/08, *available online at www.dh.gov.uk.*

ACTIVITY

Where data are not routinely available but need to be collected specifically, tools can be used to ensure consistent data collection. These tools include those shown in Box 1C.3.2.

Box 1C.3.2

Minimum data sets	A minimum amount of information is collected and collated for each patient using the service (e.g. patient's postcode, age, sex, ethnicity and reason for seeking care)
System prompts	GPs' clinical systems can automatically prompt them to ask patients about relevant health services (e.g. mothers whose young children are due for routine immunisations)

Data sources, types of information collected and organisations in England that use the data are shown in Box 1C.3.3. Table 1C.3.1 gives (roughly) equivalent organisations across the different countries of the UK.

Eng Box 1C.3.3

Data source	Type of information	English organisations that use the data
Hospital episode statistics	Number of procedures (operations, maternity services, psychiatric care) taking place in NHS services, broken down by: • Diagnoses • Health-care resource groups • NHS trusts • Length of stay • Time waited • Admission methods • Patients' age, sex and ethnic group	Department of Health (DH): to shape strategic direction of health services and secure funding Healthcare Commission: to inspect health services Regional public health observatories: provide local information Strategic health authorities: performance management of NHS trusts Primary care trusts (PCTs) and hospital trusts: commissioning, monitoring contracts, performance management

Box *contd overleaf*

Box 1C.3.3 *contd*

Data source	Type of information	English organisations that use the data
PACT, ePACT	Prescription data (cost quantity), mainly from primary care	PCTs: monitoring and performance management of primary care practitioners Prescribing Support Unit at DH and National Prescribing Centre (see Section 3B.5)
Stop-smoking services: numbers attempting to quit	Profile and numbers using services: • General accessibility of service • Groups not accessing services	Strategic health authorities: performance management NHS trusts: shape service delivery
Coverage data: screening, immunisations	Percentage population offered services receiving services Percentage eligible population receiving services	DH/strategic health authorities: performance management NHS trusts: shape service delivery

UK **Table 1C.3.1** UK equivalence table

England	Scotland	Wales	Northern Ireland
Department of Health	Scottish Executive Health Department	Welsh Assembly Government	DH and Social Services
Healthcare Commission	NHS Quality Improvement Scotland	Healthcare Inspectorate for Wales	Second review of public administration (RPA)
Regional Public Health Observatories	Scottish Public Health Observatory	Wales Centre for Health and National Public Health Service Wales	All Ireland Public Health Observatory
Strategic health authorities	NHS boards (responsible for hospitals, community health services and healthy improvement)	Regional Offices of the Welsh Assembly	New single Health and Social Services Agency to replace the existing four health boards
Primary care trusts		Local health boards	Seven new local commissioning groups to be formed
Hospital trusts		Hospital trusts	
Health Protection Agency (HPA)	HPA	HPA	HPA
Local and regional services of the HPA	Health Protection Scotland	National Public Health Service for Wales	Communicable Disease Surveillance Centre, Northern Ireland

UK The following central elements of the HPA support all four countries of the UK:

- Centre for Infections (Colindale)
- Centre for Emergency Preparedness and Response (Porton Down)
- Centre for Radiological, Chemical and Environmental Hazards (Chilton).

In England, the HPA's local and regional services (LaRS) provide surveillance, planning and health protection at

a local level, including local offices called health protection units. In Northern Ireland, Scotland and Wales, the equivalent functions of LaRS are conducted by the Infection and Communicable Disease Service teams of the respective national organisations detailed above. See Section 2G for more information about health protection organisations and roles.

1C.4 MEASURES OF SUPPLY AND DEMAND

See Section 4D.1.

1C.5 STUDY DESIGN

Study design for assessing effectiveness, efficiency and acceptability of services, including measures of structure, process, service quality and outcome of health care

A range of different methods can be used to evaluate the impact and cost-effectiveness of a health service. The technique used will depend on the exact question being asked, the time-frame, level of evidence needed and the budget for the study.

Examples include:

- Health technology assessment (HTA)
- Service delivery and organisation (SDO)
- Experimental study designs (RCTs, qualitative studies, surveys and documentary evidence)
- Comparative studies
- HNA.

Study data can either be routinely collected or else collected especially for the particular study. Examples are described in Box 1C.5.1.

Box 1C.5.1

	Routinely available data	Specifically collected
Structure	Staffing: human resources records Finance: accounts, annual reports, receipts	Equipment audit
Process	Patient care: records, referral letters, treatment plans, information leaflets	Patient diaries
Outputs	Activity: hospital episode statistics	Clinical audit and clinical research reports
Outcomes	Mortality statistics	National confidential enquiries
Health policy	Media (TV programmes, newspaper articles) Official records – parliamentary debates, bills, meeting minutes	Interviews

HEALTH TECHNOLOGY ASSESSMENT

This type of study asks what technology is best for meeting the needs of specific patients or procedures, and what impact the health technology will have on:

- **Outcomes** (morbidity, mortality, quality of life, patient satisfaction)
- **Structure** (resources required to provide service).

Common study designs include:

- Epidemiological study designs (see Section 1A.19)
- Economic evaluations (see Section 1C.14).

NICE Health Technology Appraisals are summarised in Box 1C.5.2.

Box 1C.5.2

NICE Health Technology Appraisals

Types of technologies that may be appraised include:

- Pharmaceuticals
- Medical devices
- Diagnostic techniques
- Surgical procedures
- Other therapeutic technologies
- Health promotion activities

Focus of NICE reviews:

- **Clinical** effectiveness
- **Cost-**effectiveness
- **Acceptability** (to patients and/or professionals)
- **Feasibility**
- **Equity**

Study designs used by NICE:

- **Systematic review**: evidence is rated according to the hierarchy of evidence (see below and Section 1A.37)
- **Economic evaluation**
- **Appraisal**: review of evidence in consultation with individuals, patient/carer groups, manufacturers, health care professionals (e.g. through the medical Royal Colleges)

NICE Health Technology Appraisals make **recommendations** for when and whether the NHS should provide a technology, and the expected impact on resources.

Reproduced from NICE (2004).

SERVICE DELIVERY AND ORGANISATION

Research in this field largely considers issues of health-care process, and how these will impact on outcomes: see Box 1C.5.3.

Box 1C.5.3

Where	Where is health care provided? Is the setting appropriate?
Who	Who provides the health care? Could another professional provide care more cost-effectively? How do different professionals work together?
How	How is health care being delivered? Are there other more effective (or cost-effective) ways of providing it? What factors facilitate or hinder changes to the way in which care is provided?

EXPERIMENTAL STUDY DESIGNS

While RCTs are considered the gold standard for many scientific studies, they have specific limitations when applied to the local study of health service delivery and organisation: see Box 1C.5.4.

Box 1C.5.4

Resources	Resources needed for a well-run RCT are usually unavailable for small-scale studies in health services
Timescales	Impact on health may not be apparent for decades, and the long-term follow-up needed could be prohibitively expensive in an RCT
Changes in policy	Changes in national policy may require changes to be made to local services
Multi-site studies	Customisation of managerial interventions at a local level may make multi-site RCTs impossible

Study designs other than RCTs may therefore be used, such as:

- Qualitative studies (interview, focus group, observation – see Section 1D)
- Surveys
- Documentary studies.

NICE ranks experimental studies according to the scale in Table 1C.5.1 (see also Section 1A.37):

Table 1C.5.1 Type and quality of evidence. The symbols are further indication of the quality of evidence, $^{++}$ is stronger than $^{+}$ and $^{-}$ weaker than both $^{++}$ and $^{+}$.

1^{++}	High-quality meta-analyses, systematic reviews of RCTs or RCTs (including cluster RCTs) with a very low risk of bias
1^{+}	Well-conducted meta-analyses, systematic reviews of RCTs or RCTs (including cluster RCTs) with a low risk of bias
1^{-}	Meta-analyses, systematic reviews of RCTs or RCTs (including cluster RCTs) with a high risk of bias
2^{++}	High-quality systematic reviews of non-RCT studies, or individual non-RCTs, case–control studies, cohort studies, interrupted time series (ITS) and controlled before and after (CBA) studies with a very low risk of confounding, bias or chance, and a high probability that the relationship is causal
2^{+}	Well-conducted non-RCTs, case–control studies, cohort studies, CBA studies, ITS and correlation studies with a low risk of confounding, bias or chance, and a moderate probability that the relationship is causal
2^{-}	Non-RCTs, case–control studies, cohort studies, CBA studies, ITS and correlation studies with a high risk – or chance – of confounding bias, and a significant risk that the relationship is not causal
3	Non-analytical studies, e.g. case reports, case series
4	Expert opinion, formal consensus

Reproduced from NICE (2006).

HEALTH SYSTEM COMPARISONS

Comparative studies consider variations in the structure, process and outcomes of health care and are particularly useful for evaluating:

- **Funding** (see section 4). Examples include: contracting and commissioning systems

- **Performance**. Examples include 'benchmarking' and assessments within health systems (e.g. Healthcare Commission review, see Box 1C.5.5) and between systems (e.g. WHO, The world health report 2000 – Health systems: improving performance 2000).

Eng Box 1C.5.5

Example: Healthcare Commission inspections of the NHS: annual health check

1. Online survey: NHS trusts complete a self-report of their progress in meeting core standards
2. Activities to cross-check organisations' self-reports through:
    3. Documents: supplied by the trust to support statements in the self-assessment
    4. Surveys: patient survey, reports from partners, e.g. local authorities
    5. Observations: visits made to a random selection of trusts
    6. Performance against national targets: assessment uses data submitted to a range of organisations, e.g. DH
    7. Use of resources: local financial audits are used

Reproduced from www.healthcarecommission.org.uk.

HEALTH NEEDS ASSESSMENTS

See Section 1C.1.

NATIONAL DISEASE AUDITS

This study design is helpful for evaluating the effectiveness of health care at a national level, as shown with the example of the National Sentinel Stroke Audit (Box 1C.5.6).

UK Box 1C.5.6

Example: National Sentinel Stroke Audit

Conducted by the Royal College of Physicians of London, this is a biannual audit of stroke unit organisation. All UK hospitals (except those in Scotland) that claim to operate a stroke unit are surveyed according to five core criteria:

1. Consultant physician with responsibility for stroke
2. Formal links with patient and carer organisations
3. Multidisciplinary meetings at least weekly to plan patient care
4. Provision of information to patients about stroke
5. Continuing education programmes for staff

Reproduced from Hoffman et al (2004).

1C.6 STRUCTURE, PROCESS AND OUTCOMES

Measures of structure, process, service quality and outcome of health care

The performance of an activity may be assessed using the framework first described by Donabedian (1966): see Box 1C.6.1. This includes **structure, process** and **output**. An additional category of **outcome** is often added.

Box 1C.6.1

Donabedian's framework
• Structure • Process • Output • Outcome

STRUCTURE

This category consists of all of the **inputs** to the activity, such as:

- Staff
- Budgets
- Buildings
- Beds.

Advantages and disadvantages of considering this criterion are shown in Box 1C.6.2.

Box 1C.6.2

Advantage	Disadvantage
Resource information is relatively easy to measure	Structural data may not be comparable between systems. For example, with regard to staffing, all nurses may not be working at the same level of responsibility or may not have received the same level of training

PROCESS

These are the **activities** that constitute the intervention, and may be considered as follows:

- Strategies, plans and procedures
- Referral patterns
- Prescription practices
- Consultations (duration, number)
- Bed occupancy
- Waiting times.

Advantages and disadvantages are shown in Box 1C.6.3.

Box 1C6.3

Advantages	Disadvantages
Relatively easy to obtain in centralised health-care systems such as the NHS	Processes do not necessarily predict health outcomes
Some processes are directly related to outcomes, e.g. immunisation coverage	

OUTPUTS

These are the **products** of the activity. Examples include:

- Numbers of operations conducted
- Length of stay (an indicator of effectiveness)
- Waiting times (an indicator of access).

OUTCOMES

These are changes in **health status** that are attributable to the activity. Examples include:

- Death (mortality rates)
- Disability (and quality of life)
- Discharge (and complications, e.g. emergency re-admissions).

Advantages and disadvantages are shown in Box 1C.6.4. The challenges of using outcomes to evaluate health services are illustrated with respect to smoking cessation in Box 1C.6.5.

Box 1C.6.4

Advantages	Disadvantages
Ultimately, the aim of health service interventions is to improve outcomes	Not necessarily related to performance: affected by case mix, i.e. patient characteristics
Surrogate endpoints can indicate health outcomes, e.g. CD4 counts in an HIV/AIDS drug trial	Outcomes are relatively long term, so difficult to use in short-term trials
	Often costly to collect or incomplete (apart from mortality data – but these can be difficult to link to interventions)

UK Box 1C.6.5

Example: stop-smoking services	
Structure	Numbers of staff employed as stop-smoking advisors Budget for nicotine replacement therapy (NRT)/bupropion
Process	Number of prescriptions of NRT or bupropion Advisors' training Numbers of clients setting quit dates
Output	Numbers of 4-week quitters
Outcome	Ideally the morbidity and mortality from smoking-related diseases. However, outcomes have to be attributable to the intervention, and the effects of stop-smoking services on lung cancer deaths may not be apparent for 20 years. Surrogate endpoints could be used instead and include length of stay in hospital for ex-smokers undergoing surgery

1C.7 MEASURING HEALTH

Measures of health status, quality of life and health care, population health outcome indicators

The broad definition of health used by public health practitioners means that specific definitions of individual health status are required. At an aggregate level, a variety of health indicators is required to describe a population's health.

HEALTH STATUS

In 1980 the World Health Organization (WHO) published the *International Classification of Functioning, Disability and Health*. This system (which distinguished impairment, disability and handicap) was updated in 2000 and is shown in Table 1C.7.1.

Table 1C.7.1 World Health Organization classification of functioning, disability and health

WHO 1980 term	Impairment	Disability	Handicap
WHO 2000 Term	Body structure/function	Activities	Participation
Description	Pathology and clinical measures	Symptoms and health status	Effect of disability on life
Measurement	Usually observed by clinicians or measured using instruments	Self-reported measure of what the person can do	Usually self-reported. This is the extent to which the condition affects the person's normal life
Examples of measures	Blood pressure, temperature, tumour size as assessed on CT	Beck Depression Inventory Activities of daily living scales	Health-related quality of life
Example 1 (prostate)	Examination: enlarged prostate Test: increased prostate-specific antigen levels in serum	Self-report: trips to the toilet per day	Self-report: person cannot go to cinema because they need the toilet too often
Example 2 (hip)	Clinical examination: reduced joint mobility Investigation: joint degeneration seen on radiograph	Self-report: pain or distance that patient can walk	Self-report: isolation – cannot walk to the park

HEALTH-RELATED QUALITY OF LIFE

Increasing importance is being placed on health-related quality of life (HRQoL) in clinical trials. This is because for many potentially fatal conditions (e.g. lung cancer) or for many long-term conditions (e.g. epilepsy) treatments may not affect life-expectancy but still have a profound impact on the quality of life. Studying HRQoL can also reveal why a particular treatment is acceptable or not to patients. For example, a trial of cancer chemotherapy might reveal that small treatment benefits (a few weeks' extra life-expectancy) were outweighed by intractable nausea and vomiting.

Different theoretical models exist for measuring quality of life, and take into account the factors listed in Box 1C.7.1.

Box 1C.7.1

Expectations	Differences between a person's present experience and his or her hopes and expectations. For example, a professional pianist would be more affected by the loss of the use of a finger than would a cleaner
Needs	Ability or capacity of a person to satisfy basic human functions, e.g. eating, sleeping, enjoying music
Normal living	This is being able to do what the person wants to do (rather than being free of disease or symptoms)

HRQoL scales contain several items that build up a composite picture of health status. Except in exceptional circumstances (e.g. dementia or very young children), HRQoL scales should be completed by the patient rather than by a carer or clinician.

Examples of health-related quality of life scales are listed in Box 1C.7.2.

Box 1C.7.2

Type of measure	Examples
Generic	Short Form 36 (SF36) Nottingham Health Profile (NHP) EQ5D (EuroQoL)
Disease related	Functional Assessment of Cancer Therapy
Specific aspects of quality of life	Hospital Anxiety and Depression Scale; Multidimensional Fatigue Scale

MEASURES OF HEALTH CARE

Ways of assessing different measures of health care are described in Sections 1C.5 and 1C.6.

In addition to quality-of-life measures (above) and population outcome indicators (below), measures can also include:

- Patient satisfaction (see Section 1C.15)
- Adherence to standards, compliance with guidelines and other established criteria (see Section 1C.11)
- Financial accountability (see Section 5).

See also Section 1C.9.

POPULATION HEALTH OUTCOME INDICATORS

These indicators measure the effect of health care on health status at a population level. They often draw on the concept of **avoidable mortality** as an indicator of health-care performance. Assessment of avoidable mortality involves studying the incidence of a collection of diseases with **unnecessary** or **untimely outcomes**.

Examples are listed in Box 1C.7.3.

Box 1C.7.3

Measure	Disease and population	Purpose
Standardised mortality ratios	Diabetes: people aged under 45	Reduce premature deaths from diabetes
	Gastric, duodenal, and peptic ulcers: people aged 25–74	Reduce deaths from gastric, duodenal and peptic ulcers
Procedures	Lower limb amputations in diabetic patients	Monitor and improve management of diabetes
	Termination of pregnancy	Reduce the number of unwanted pregnancies

Indicators reflect:

- Performance of health-care services
- Population characteristics.

Population health outcome indicators can be used to:

- **Prompt the assessment of local health outcomes**, particularly in relation to national targets
- Monitor **variation** in health care (particularly using small area analysis – see Section 1A.18)
- Monitor **trends** in health care, to answer questions such as: 'Is effectiveness of health care improving?' (see Lakhani et al 2005)
- Monitor of **quality of life** as part of a population HNA.

Eng In England, data on population health outcome indicators are published in the *Compendium of Clinical and Health Indicators*, produced by the National Centre for Health Outcomes Development.

1C.8 DEPRIVATION MEASURES

Deprivation measures describe a population according to social and economic disadvantage. They are used to identify people or areas with the highest levels of disadvantage and deprivation, in order to target programmes and resources, and may be measured at the individual or area level.

INDIVIDUAL INDICATORS

Indicators of deprivation at the individual or household level include:

UK NATIONAL STATISTICS SOCIOECONOMIC CLASSIFICATION (NSSEC)

The former system of social class (I–V) based on occupation has been replaced by an arrangement based on occupation and economic group. The Office of National Statistics now uses this system in all official statistics and surveys. A category is assigned according to the current or former occupation of the *Household Reference Person* (i.e. the person responsible for owning or renting the household's accommodation) using the scale in Table 1C.8.1.

UK Table 1C.8.1

NSSEC categories		
1	Higher managerial and professional occupations	
	1.1	Large employers and higher managerial occupations
	1.2	Higher professional occupations
2	Lower managerial and professional occupations	
3	Intermediate occupations	
4	Small employers and own account workers	
5	Lower supervisory and technical occupations	
6	Semi-routine occupations	
7	Routine occupations	
8	Never worked and long-term unemployed	

Reproduced from the Office for National Statistics (2007).

NZ NZIDEP – SOCIOECONOMIC DEPRIVATION FOR INDIVIDUALS

The New Zealand index of socioeconomic deprivation for individuals (NZiDep) does not depend on occupation or education data. Instead, it assigns deprivation levels on a five-point scale to individuals based on responses to eight simple questions, e.g. 'In the last 12 months have you personally been forced to buy cheaper food so that you could pay for other things you needed? [Yes/No].'

OTHER INDIVIDUAL INDICATORS

Alternative markers of socioeconomic status that are sometimes used in public health include:

- Income
- Occupation
- Years of education
- Housing – owner/renter, or occupancy per room
- Ownership of commodities (e.g. car, television).

AREA INDICATORS

Countries have adopted different indicators tailored to circumstances and availability of data.

Eng INDEX OF MULTIPLE DEPRIVATION (IMD)

The Department for Communities and Local Government (CLG) applies the IMD-2007 scale to the English Super-Output Areas (SOA) see Box 3A.1.1 on page 327. Each SOA is scored and ranked for deprivation based on seven weighted measures:

- Income
- Employment
- Health
- Education, skills and training
- Crime
- Housing and services
- Living environment.

Reproduced from Social Disadvantage Research Centre (2007).

Scot SCOTTISH INDEX OF MULTIPLE DEPRIVATION

The Scottish Index of Multiple Deprivation (SIMD) is used across local and national government for directing resources, setting targets, and monitoring social and health inequalities. It is available at data zone level as an overall deprivation index. It includes separate indices for different domains:

- Income
- Employment
- Housing
- Health
- Education
- Geographical access
- Telecommunications.

Reproduced from the Scottish Government's Scottish Index of Multiple Deprivation (2004).

UK JARMAN

The Jarman score is based on factors that affect patients' demand for primary care. It is a measure of **GP practice workload** and is used to calculate additional deprivation payments to practices. Factors are determined from the **census** and include:

- Elderly living alone
- One-parent families
- Children under 5 years old
- Unemployed (as percentage of economically active population)
- Overcrowded households
- Moved house within the last year
- Born in the New Commonwealth or Pakistan.

Reproduced from www.primary-care-db.org.uk/indexmenu/jarman_desc.html.

Eng TOWNSEND

This is another census-based measure of comparative deprivation, and is defined by the proportion of households that:

- Have more than one person per habitable room
- Have no car
- Are not owner occupied
- Include a person who is unemployed.

Wal WELSH TOWNSEND

Welsh Townsend is a direct measure of health needs rather than a measure of deprivation as such. It is included here to avoid the potential for confusion with Townsend. The indicator that is used to allocate financial resources for health care uses Welsh Health Survey data.

Scot CARSTAIRS

This is a measure of relative deprivation used in Scotland. It is based on four census variables:

- Overcrowding
- Male unemployment

- Low social class
- No car.

Reproduced from Carstairs (1989) and Morgan (2006).

NI NORTHERN IRELAND MULTIPLE DEPRIVATION MEASURE (NORTHERN IRELAND STATISTICS AND RESEARCH AGENCY, 2005)

The NIMDM contains seven domains of deprivation, namely.

- Income deprivation
- Employment
- Health deprivation and disability
- Education and skills, training
- Proximity to services
- Living environment
- Crime and disorder.

Ire IRELAND INDICES

Two indices of area level deprivation have been developed in Ireland, based on information from the census. The **Small Area Health Research Unit** (SAHRU) index includes:

- Unemployment rate
- Percentage of households that are local authority owned
- Percentage of the population in social classes 5 and 6
- Percentage of households with no car
- Persons per room.

The **Haase and Pratschke index** includes persons per room and a more detailed classification of social class, unemployment, age, education status, lone parents and change in population size since the previous census. Underlying the index are the three dimensions of rural demographic decline, urban labour market deprivation and social class disadvantage.

NZ NEW ZEALAND DEPRIVATION INDEX

The New Zealand small area index of relative socioeconomic deprivation, NZDep, is based on nine census variables, and is widely used within the country for health resource allocation, needs assessment, research and advocacy. It provides a deprivation score between 1 and 10, with 1 being the least deprived decile of areas, and 10 being the most deprived decile of areas. NZDep is available in four versions according to the census on which it was based (e.g. NZDep2006 was based on the 2006 Census). Scores are available for **meshblocks** (which correspond to a city block) and **census area units** (which correspond to a neighbourhood), as well as larger geographical areas.

NZDep indices are based on the following nine measures generated from census variables:

- No access to a telephone
- Receiving a means-tested benefit, aged 18–59
- Unemployed, aged 18–59
- Household income below a threshold
- No access to a car
- Single-parent family, aged 18–59
- No educational qualifications, aged 18–59
- Living in rental housing
- Living in overcrowded housing.

SA PROVINCIAL INDICES OF MULTIPLE DEPRIVATION

The nine Provincial Indices of Multiple Deprivation (PIMD) for South Africa are the first stage of a larger project to measure multiple deprivation at a small area level for the whole country. Each PIMD was created using the 2001 census: distinct, unidimensional domains of deprivation were combined, with appropriate weighting, into a single measure of multiple deprivation.

COMPOSITE MEASURES

Measures such as the IMD, Townsend and Carstairs are **composite** indicators of deprivation.

Box 1C.8.1

Advantages of composite indicators	Disadvantages of composite indicators
Readily available	Rely on census data, therefore could be up to 10 years out of date
Can shape health service and local authority planning and priorities (e.g. initiatives can be targeted to areas of high deprivation or to specific features of deprivation, such as crime)	Potential for autocorrelation (see IMD above)
Can be used for research into associations between deprivation and health	

1C.9 EVALUATION

Principles of evaluation, including quality assessment and quality assurance

In public health, an evaluation of a service is a **thorough assessment** of whether a health service meets its objectives. This can be performed using routine or specially collected data, and may form part of an **audit** or **research** exercise (Box 1C.9.1).

Box 1C9.1

Research	Results that can be generalised outside the setting of the evaluation
Audit	Results are pertinent only to the setting being evaluated

The perspective of the evaluation must always be stated (e.g. the commissioner, the clinician, the patient or society in general). A rigorous evaluation will address whether the service is **cost-effective** and also whether it is **needed** and **wanted** by the local population.

FRAMEWORKS FOR EVALUATING SERVICES

Two standard frameworks used in public health are those of **Donabedian** and **Maxwell**.

DONABEDIAN

This framework of **structure**, **process**, **output** (and **outcome**) is discussed in Section 1C.6.

MAXWELL

The framework described by Maxwell (1984) consists of six dimensions of quality described in Table 1C.9.1.

Table 1C.9.1 Maxwell's dimensions of quality

Access	Whether patients can see a clinician when they wish Whether there are any tangible or intangible barriers to access. The latter reflect how uninviting or unsuited the service is to particular groups **Tangible barriers** • Geography (services near to home or public transport) • Finance (user charges, costs attached to attending services, e.g. lost income, childminding fees) • Opening hours, waiting times **Intangible barriers** • Languages used • Female doctors available to see female patients
Relevance	Whether services provided are appropriate to patients' needs
Equity	Whether services are provided fairly (not necessarily equally – see Section 1C.10)
Efficiency	Whether the costs of providing interventions are justified by the benefits. Economic evaluations (section 4) are used to address this question
Effectiveness	Whether the outcomes of the interventions are reflected in improvements in health (see Section 1C.6)
Acceptability	Whether the patient (and in some cases the practitioner) is satisfied by the care provided (for patient satisfaction measures, see Section 1C.15)

QUALITY ASSESSMENT AND QUALITY ASSURANCE

In public health, the terms *quality assessment* and *quality assurance* are often used interchangeably with the term *evaluation*. However, in other fields the term *quality assurance* may relate to the process of ensuring that a service meets a defined standard.

1C.10 EQUITY IN HEALTH CARE

Equity and equality are distinct concepts: see Box 1C.10.1.

Box 1C.10.1

	Equity	Equality
Definition	Fairness Distributive justice	Sameness
Issue	Differences in health due to avoidable inequalities (e.g. distance from a general practice surgery)	Avoidable and unavoidable differences (e.g. genetic diseases and road traffic injuries)

As with the concept of fairness, there are different viewpoints as to what constitutes equity.

VERTICAL EQUITY

Vertical equity requires unequal health care to be provided for unequal need. For example, in a system that is vertically equitable, sicker people on a waiting list would be prioritised above others who were less sick but had been waiting longer on the same list. Similarly, a progressive taxation system (where the rich face a disproportionately large tax burden) is vertically equitable.

HORIZONTAL EQUITY

Horizontal equity requires equal health care for equal need. It can be considered in four ways outlined in Table 1C.10.1, although it is recognised that access and use overlap.

Table 1C.10.1 Horizontal equity

Type	Definition	Advantages	Disadvantages
Equal spending for equal need	Budgets are allocated according to health need. For example, resources would be targeted by means of SMRs and area-level deprivation indicators (see 4D3)	Simple to measure because spending can be clearly identified	Health care for a particular condition will cost varying amounts in different settings, e.g. cost of a visit from a district nurse may be greater in rural than urban areas because of a need to travel further
Equal access for equal need	Access can be considered in terms of removing the barriers to service use. The barriers could be geographical (distance to the service), financial (cost to the user) or related to time (whether services are open only during office hours)	Defining principle of services such as the NHS Important for providers to consider whether the supply of services is equitable	Barriers to access may be intangible (e.g. cultural) and therefore difficult to identify
Equal use for equal need	This assumes that people should have the same health-care demands for the same symptoms, so that any differences reflect barriers to access. One example often quoted is that of coronary bypass rates in areas of different SMR for coronary heart disease. Note that many studies that purport to measure access in fact measure health-care use	Overcomes the problem of intangible barriers to access	Assumes that people with the same symptoms have the same demand for health care
Equal health for equal need	This is clearly the ultimate aim for equity in health services	Gold standard	Subject to factors outside the control of health care, which may be unavoidable (e.g. genetics) or avoidable (e.g. poverty)

In a seminal paper in 1971, Julian Tudor Hart described inequities, particularly as result of the introduction of market forces, in the NHS. He coined the phrase, the *inverse care law*, to describe this phenomenon:

'the availability of good medical care tends to vary inversely with the need of the population served'

Tudor Hart's evidence of the inverse care law came from the distribution of GPs in South Wales, but in subsequent years empirical evidence of the existence of the inverse care law has been found in numerous areas of health.

1C.11 CLINICAL AUDIT

A clinical audit is a **systematic review** of care, as measured against **explicit criteria**. It is a central component of **clinical governance**, the latter being a local system for assuring the quality of patient care and for making ongoing improvements.

UK National audit programmes also exist, including:

- Stroke – the National Sentinel Stroke Audit
- Coronary heart disease – Society of Cardiac Surgeons' Cardiac Surgery Audit.

AUDIT CYCLE

In order for lessons to be learnt from an audit exercise, the process should always be seen as an ongoing cycle rather than a one-off event (see Figure 1C.11.1).

The steps involved in conducting an audit are described in Table 1C.11.1.

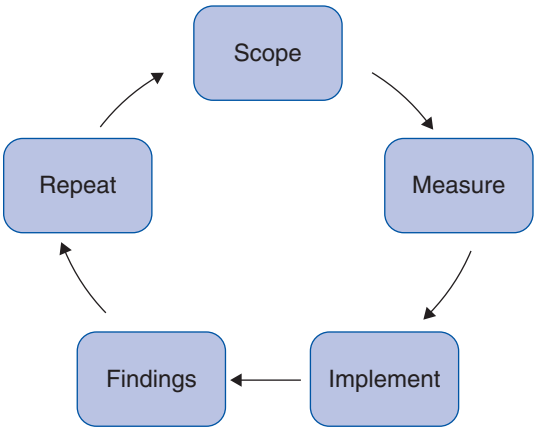


Figure 1C.11.1 The audit cycle

Table 1C.11.1 The audit process

Scope	Begin by selecting criteria against which to audit. Useful sources of criteria include national standards (e.g. National Service Frameworks) and national guidelines (e.g. NICE Health Technology appraisals) The criteria should be: • Explicit statements that define what is being measured • Elements of care that can be measured objectively The scope of the audit can then be determined by deciding: • Which patients should be included • Over what time period they should be audited • Which aspects of care should be considered An appropriate sample should be chosen by using a strategy such as: • Interval sampling (all patients visiting a clinic in a given period of time, e.g. January to March) • Two-stage sampling (a small sample is selected first; if unequivocal conclusions can be drawn, then no more data are collected; if the results are ambiguous, then a larger sample is selected)
Measure	Measure observed performance against the selected criteria. Data sources include: • Registers – to identify patients • Patient records – to examine aspects of care • Electronic systems – for rapid retrieval of information • Teams delivering the care – to find out if information is collected, and the extent to which this information is suitable

Findings	Use the findings to identify areas for change. Ensure that the findings are shared with those who need to know, including: • Teams working in the area of the audit • Area service directors • Boards Produce short summaries of findings that are readily understandable
Implement	Implement the findings of the audit by making improvements to the service Consider adopting change management approaches (see section 5C)
Repeat	Check that these improvements are in place and are making a difference by repeating the audit cycle

SUCCESSFUL AUDIT

Baker and colleagues (1995) identified several factors influencing doctors' participation in audit. These factors and other situations identified by Johnston and colleagues (2000) that may influence the success of an audit are described in Table 1C.11.3.

Table 1C.11.3 Drivers and barriers to success in audit

	Drivers of success	Barriers to success
Time	Protected time	Lack of time main reason for failure or incomplete audit
Strategy	Leadership and structured programmes	Lack of an overall plan for audits
Skills	Training for clinical staff Availability of specialist audit staff	Lack of expertise or advice in project design and analysis
Resources	Modern medical records and IT systems Dedicated audit staff Funding to conduct the audit and implement changes	Lack of these factors
Organisational culture	Dialogue between purchasers and providers Whole team participates in the audit and implementation of change Supportive learning environment with a no-blame culture	Problems between groups and group members

1C.12 CONFIDENTIAL ENQUIRY PROCESSES

Confidential enquiries are national investigations into serious untoward incidents. They are an example of a national investigation of serious concerns. Examples for England are described in Box 1C.12.1. Other examples of national investigations with implications for health include:

- Clinical Standards Advisory Group (1991–2000)
- Bichard Enquiry into the Soham murders (made recommendations regarding child protection procedures following the murder of two young girls).

The purpose of all such investigations is to identify what went wrong and to draw general lessons that can be shared.

Eng Box 1C.12.1

Confidential enquiries in England

- National Confidential Inquiry into Suicide and Homicide in Mental Health Services: www.national-confidential-inquiry.ac.uk
- Confidential Enquiry into Maternal and Child Health – www.cemach.org.uk – collects data on maternal deaths, perinatal, neonatal and infant mortality and causes of stillbirth. Sudden infant deaths are routinely reported to the Coroner's Office
- National Confidential Enquiry into Patient Outcome and Death (NCEPOD): www.ncepod.org.uk

CONDUCT OF ENQUIRIES

A **case–control** study design is often employed, and is used to detect risk factors that predispose to unexpected deaths or other untoward incidents.

Both quantitative and qualitative data may be used, including:

- Routinely collected data (e.g. hospital episode statistics)
- Hospital records
- Confidential surveys of individuals (including hospital staff, HM Coroner, families, GPs, friends).

OUTCOME OF ENQUIRIES

Confidential enquiries produce **reports** that are circulated widely and make recommendations for health services on how to minimise the risk of similar events happening again. The reports sometimes include **self-audit tools** that health services may use to assess their own risk.

1C.13 DELPHI METHODS

This is an iterative technique for **generating consensus** about a particular issue **without face-to-face meetings**. The process works as shown in Box 1C.13.1.

Box 1C.13.1

Step 1	A group of experts is contacted and surveyed
Step 2	Views of the group are shared anonymously among the same group, highlighting areas of disagreement
Step 3	Respondents are asked if they wish to change their views in light of step 2
Step 4	Steps 2–3 are repeated several times until consensus is reached

The Delphi technique was originally used for making technology forecasts, but it has also proved useful in public health, especially in fields where data are lacking, or as a way of engaging people who would not otherwise be involved. Its advantages and limitations are shown in Box 1C.13.2. The application of the Delphi method is described in an example from mental health services in Box 1C.13.3.

Box 1C.13.2

Advantages	Limitations
Anonymity and written process avoids some of the pitfalls of face-to-face discussion, e.g. each contribution is valued equally, with reduced potential for personality to influence responses	Written survey format may be more suitable for some respondents than others
Time-efficient: does not rely on bringing together disparate, busy groups	Open to manipulation from those administering it
Encourages open critique and admission of errors by encouraging contributors to revise earlier judgements	Does not always produce useful results. Sometimes future developments are more accurately predicted by unconventional thinking than by iterative consensus from accepted experts

Box 1C.13.3

Example: Describing service models of community mental health practice

This study used the Delphi method to generate a 'valid and reliable set of categories to describe the clinical work practices of intensive case managers'. Eight case managers took part in the process, which involved a combination of (a) the Delphi method to generate a list of categories to describe their working practices and (b) face-to-face discussions to refine this list. The authors describe the Delphi method as: '… an effective, straightforward, and time-efficient way of obtaining a workable consensus.'

Reproduced from Fiander and Burns (2000).

1C.14 ECONOMIC EVALUATION

See Section 4D.5.

1C.15 APPROPRIATE AND ACCEPTABLE SERVICES

Appropriateness and adequacy of services and their acceptability to consumers and providers

Both the appropriateness (**suitability**) and adequacy (**sufficiency**) of a health service can be seen to take several dimensions, including those shown in Box 1C.15.1.

Box 1C.15.1

Humanity	Degree to which patients are treated with respect
Access	Whether consumers received care when they needed it: • Acceptable waiting times • Information about the service being provided in relevant languages
Environment	Whether the service is offered in an appropriate setting: • Building was clean and well maintained • Adequate privacy was provided • Setting was suitable for its patients (e.g. toys provided for children)
Information	Whether patients are fully informed about their care: • Translation services provided • Lay terminology used to describe the treatment options and procedures

See also Section 1C.5.

The exact requirements for appropriate services will depend upon the factors shown in Box 1C.15.2.

Box 1C.15.2

Individual	Every person has different values, tolerances and ideas of what is acceptable
Patient group	Old or young; long-term or day-case patients have different requirements for appropriate services
Service	Inpatient versus community clinics have different dimensions and requirements for providing humane and accessible care
Priorities	There may be a trade-off between accessibility and humanity, e.g. the need to see a GP promptly may take priority for some over seeing the same GP each visit or each consultation limited to 5–10 min

ASSESSING ACCEPTABILITY TO CONSUMERS AND PROVIDERS

Reasons for assessing acceptability include:

- Feedback should lead to **service improvements.**
- It **guides consumers** in their choice of care. For example, information about patient satisfaction can be provided to patients who are choosing their hospital under the choose and book initiative.
- It should ultimately **guide commissioners** in their decisions of where to award contracts to health-care providers.

Eng INFORMATION USED TO ASSESS ACCEPTABILITY

See Box 1C.15.3.

Eng Box 1C.15.3

Complaints and compliments	• Informal comments/complaints, e.g. through PALS (Patient Advice and Liaison Service), conversations with staff • Formal complaints made directly to services • Ombudsman's reports, complaints reported to a third person
Untoward incidents	• Local analysis of near misses and mistakes in care • National Confidential Enquiries
Visits/enquiries	• Regulatory bodies such as the Healthcare Commission, GMC/NMC • Local scrutiny
Surveys	• National: patient and staff surveys (run by the Healthcare Commission) • General practice: undertaking annual approved patient survey part of the quality requirements of the new GMS contract • Local: bespoke surveys

CONSIDERATIONS FOR ASSESSING ACCEPTABILITY

Considerations for assessing acceptability of services are described in Table 1C.15.1 and illustrated in the example of cardiology services from Northern Ireland in Box 1C.15.5.

Table 1C.15.1 Assessing acceptability of health services

Who is surveyed	Sample selection is important to determine and record. Should people who have already complained be asked, or all patients? Different degrees of satisfaction are seen among different sections of the population: key determinants include age, socioeconomic status and ethnicity
When is the survey taken	During care or after? There is a trade-off between patients' capacity to remember care accurately versus asking someone in the vulnerable position of still receiving care
Where survey is performed	Satisfaction results are generally higher if answered within health-care setting compared with outside, e.g. the patient's home
How consumers are surveyed	Acceptability is highly dependent upon context and individuals. Qualitative studies are good for describing experiences, but are time-consuming to analyse. Quantitative studies explore only what is asked, but are useful when the questions that have to be addressed are known (e.g. after some initial qualitative research to identify key issues). For example, if cleanliness arises as an issue in qualitative research, then quantitative methods could assess how often wards are dirty

NI **Box 1C15.4**

Example: Survey of cardiology patients

A group of patients was contacted immediately after angiography to determine their views on what they believed should influence priorities for cardiac revascularisation. The survey revealed that patients felt that the following factors should be considered:

- Patients' age
- Patients' smoking status
- Whether patients have dependent relatives

Reproduced from Kee et al (1997).

1C.16 EPIDEMIOLOGICAL BASIS FOR PREVENTIVE STRATEGIES

The epidemiology of a disease can guide strategies for preventing it: factors that are found to be associated with the disease, and which meet the criteria for causation, can provide useful targets for preventive strategies (see Section 1A.13).

LEVELS OF PREVENTION

Preventive interventions are considered at three levels: primary, secondary and tertiary: see Table 1C.16.1.

Table 1C.16.1 Levels of prevention

Level of prevention	Definition	Example
Primary prevention	Aims to prevent or delay the onset of disease in those at risk but who do not currently have the disease	Low-dose aspirin given to those at sufficient risk of developing ischaemic heart disease
Secondary prevention	Aims to prevent or delay further progression or complications in people known to have the disease	An exercise programme offered to patients who have had a myocardial infarction
Tertiary prevention	Aims to reduce long-term disability that would otherwise result from the disease	Removing allergens from the environment of a patient with asthma

Note that primary prevention is not the same as screening: the latter determines who in a population is at sufficient risk of the disease to warrant further investigation.

PREVENTIVE STRATEGIES

There are broadly two approaches to preventive medicine: the **population** approach and the **high-risk** approach: see Table 1C.16.2.

Table 1C.16.2 Population and high-risk approaches to prevention

Strategy	Description	Example
Population approach	The whole population receives health promotion and disease prevention interventions to reduce risk throughout because all are at risk	Legislation for wearing seat belts in cars
High-risk approach	Health promotion and disease prevention is targeted at groups identified from epidemiological studies to be most at risk	Chlamydia screening targeted at people aged 15–24 (who have the highest prevalence of the disease among the population)

Epidemiological strategies to identify high-risk groups include:

- Ecological studies to **identify groups** with highest incidence of disease (see Section 1A.17)
- Analytical studies to link **risk factors and diseases** (e.g. blood pressure, diabetes and coronary heart disease)
- Genetic studies to identify families and individuals with **genetic risk factors** (e.g. gene for hypercholesterolaemia – see Section 2D.9)
- Population screening to i**dentify individuals** who have known risk factors (e.g. diabetes for heart disease, age profile for chlamydia infection).

Geoffrey Rose (1992) outlined the rationale for a population-based preventive approach in his seminal work, *The Strategy of Preventive Medicine*. This approach is useful for tackling common conditions that are normally distributed and have a widespread cause – the classic example being deaths from coronary heart disease due to hypertension (see Figure 1C.16.1 and Box 1C.16.1).

However, the **prevention paradox** (see Section 2H.4) significantly limits the effectiveness of the population approach. Current thinking emphasises the need for a combination of both population and targeted approaches to prevention (Anon 2006).

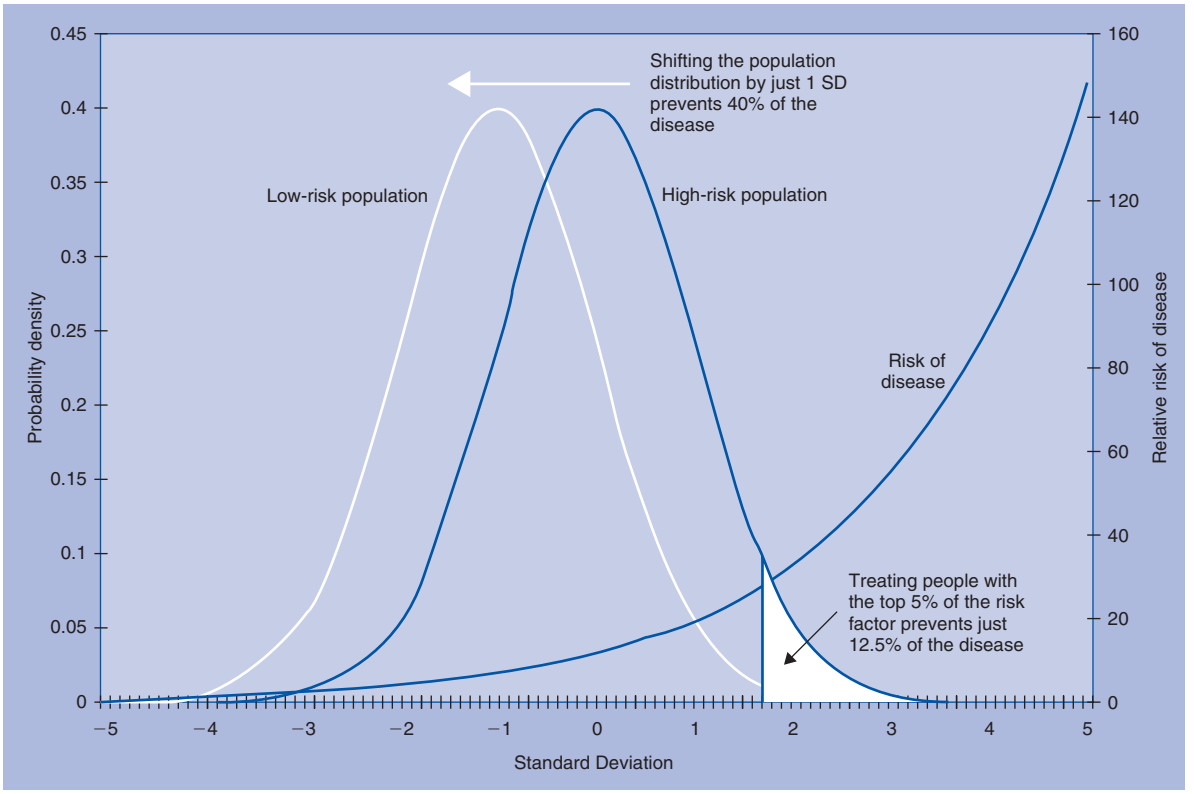


Figure 1C.16.1 Effects of a population-based versus a high-risk-based approach to preventing heart disease

Box 1C.16.1

Example: controlling risk factors for heart disease

The blood pressure of the population is normally distributed, so a minority of people have very high blood pressures. In Figure 1C.16.1, the top 5% of the population with the highest blood pressures are represented by the white triangular area. These people are at high risk of coronary heart disease (CHD).

However, since the numbers of high-risk people are small, their burden of disease on the overall population is modest. In this example, these 5% of people represent 12.5% of the burden of disease. Therefore, many more cases of CHD occur in people with lower blood pressure, i.e. to individuals who are outside the white triangular area. Targeting only the top 5% of the population with very high levels could at best reduce disease by 12.5%.

An approach that slightly reduced blood pressure across the whole population would, however, have a much bigger impact, in this example reducing disease by 40%.

Adapted from Department of Health (2000), Emberson et al (2004), Manuel et al (2006) and Marshall and Rouse (2002).

1C.17 HEALTH AND ENVIRONMENTAL IMPACT ASSESSMENTS

An environmental impact assessment (EIA) is a technique for assessing the potential effects of a proposed development on the environment. All member states of the European Union are legally required to conduct an

EIA for any major civil engineering project. Some EIAs consider health as a criterion against which to assess the proposal, but this is typically covered only superficially.

The WHO **Ottawa Charter** for Health Promotion (1986) proposed an analogous technique of health impact assessments (HIAs) that would concentrate solely on the potential effects of proposed policies or plans on the health of the population. Such policies include those shown in Box 1C.17.1.

Box 1C.17.1

Policy type	Example
Transport strategies	London Heathrow Airport Terminal 5
Civil engineering	Wembley Stadium in North London New landfill site
Urban regeneration	Olympics regeneration in East London
Lobbying	Manchester Airport second runway expansion
Developing countries	Donor aid project appraisals

STEPS IN UNDERTAKING A HEALTH IMPACT ASSESSMENT

There are five key steps to an HIA: see Box 1C.17.2.

Box 1C.17.2

1. Screening	Establish whether there is any health relevance to a policy, programme or other development
2. Scoping	For proposals with health relevance, the questions that the HIA should address need to be decided upon, together with reporting arrangements
3. Appraisal	Assess potential health impacts either through rapid appraisal (over a few days or weeks) or through in-depth assessment (weeks or months)
4. Reporting	Conclusions are made, followed by recommendations that minimise the negative impacts and maximise the positive impacts of the development
5. Monitoring	Actual impacts are monitored to enhance the evidence base for future HIAs

CHALLENGES

Challenges to conducting an HIA may include a dearth of evidence, time or resources.

EVIDENCE

A crucial point to consider is whether there is any evidence on which to base assessments. For new projects, the health impacts may never have been considered, observed or measured before. Quantifiable evidence often proves to be more influential than qualitative evidence, but may be not be feasible to collect.

The evidence for an HIA may originate from several sources, often coming from stakeholders with conflicting perspectives. HIAs may thus require the contradictory information to be synthesised.

QUALITY VERSUS TIME AND RESOURCES

Some HIAs demand rapid results with few resources, in which case it may not be possible to conduct HIAs with the same rigour as longer-term studies. This needs to be explicitly acknowledged at the outset of an HIA.

ACTION

A final, but decisive challenge is to ensure that the findings and recommendations of the HIA are actually implemented. For this to happen, the decision-makers responsible for the programme need to be engaged in the HIA at the reporting stage.

1D
Principles of Qualitative Methods

Scientific studies often provide quantitative data about an issue. For example, they might measure the proportion of people affected by a condition who have a particular risk factor. Sometimes, however, the information required to answer a question pertains more to the character of an issue, rather than to any numerical value. Qualitative methods aim to explore this second type of enquiry, i.e. those that relate to the **kind** or the **quality** of things. These techniques are best used for answering **how?, what?** or **why?** questions. For example, qualitative research questions in health include:

- What influences parents' decisions to vaccinate their children?
- How do different patients view their illnesses?
- Why do some patients decline information about their condition?

In general, qualitative methods are more time-consuming than quantitative methods with regard to data collection and data analysis. However, they can provide information that would be unobtainable by quantitative means.

1D.1 OVERVIEW OF QUALITATIVE METHODS

Overview of qualitative methods, including semi-structured and in-depth interviewing, focus groups, action research, participant observation and their contribution to public health research and policy

The four principal methods for collecting qualitative data, summarised in Table 1D.1.1, are:

- Semi-structured and in-depth interviews
- Focus groups
- Action research
- Participant observation.

Table 1D.1.1 Four principal methods of qualitative data collection

	Semi-structured and in-depth interviewing	Focus groups	Action research	Participant observation
What is done?	Interviewer uses a topic guide, rather than a rigid set of questions. Issues should arise naturally in the conversation. In-depth interviews can be conducted over several hours. Researchers may also conduct several interviews at different stages with the same individuals	Facilitator leads a small number of people (usually 5–10) in a structured conversation. The focus is not just on the responses to the facilitator but also on the reactions to other participants	The people under study (e.g. those using a service) are actively involved as researchers. Action research enables researchers/actors to make changes to projects during the process, not just at the end of the research	Researchers actively take part in a setting as they collect their data, either as a member of the community (e.g. a nurse in a hospital), or a supporter (e.g. hanging around a criminal gang but not taking part in criminal activities)
How is this used?	**Uses:** • Generating new concepts and ideas around an issue • Seminal study: Jocelyn Cornwell – Hard Earned Lives: explored people's accounts of health care over several interviews. Revealed: hierarchies of health settings; what is 'counted' as health; why people do not seek care	**Uses:** • Understand target audiences • Exploring subjects that are difficult to discuss individually, e.g. sexual beliefs and behaviours	**Uses:** • Involving people who would otherwise not normally take part in research • Community development	**Uses:** • Immersing the researcher in the experiences of those under study • Gaining insight into subcultures not usually open to study or observation • Seminal study: Rosenbaum's study of treatment of healthy people in psychiatric institutions when they pretended to experience hallucinations

Qualitative data are analysed in different ways depending, in part, on the method of data collection that was used. There are three principal methods: see Box 1D.1.1.

Box 1D.1.1

Thematic content	The simplest form of analysis, this usually involves reporting interviewees' comments in a structured form. Useful for areas where not much is known and as a starting point for other types of analysis
Grounded theory	Used for developing theories. Data collection takes place in a cyclical process: themes identified from the analysis are fed back into the topics to be discussed during interviews, and new data are used to refine the analysis. The analysis involves organising content into codes and continually refining these codes
Framework	Used for developing practical strategies as a result of research

There are several ways of enhancing the validity and the credibility of qualitative analysis, including:

- Identification of **disconfirming** cases or **deviant data**, then accounting for these, i.e. not solely quoting supportive data

- Developing a **coding structure** with other researchers
- **Presenting back** to interviewees to check that the material accurately represents what they thought
- **Counting responses** in themes to give an indication of how common a particular response was (e.g. 'seven participants said x' or 'most said y')
- Being as **transparent** as possible in the presentation and including raw data if possible (e.g. quotes)
- Describing the **process** of analysis
- **Comparing** and **contrasting** findings with findings from other studies, attitudes at beginning of the study or other data collected in this study
- **Cross-checking** findings using other study methods (e.g. questionnaires) to explore the **effect of the interviewer/facilitator** on findings. For example, a health service employee may hear only positive comments about the health service.

1D.2 APPROPRIATE USE

Each of the principal types of qualitative investigation has a different application (see Table 1D.2.1).

Table 1D.2.1 Applications of qualitative investigation

Semi-structured and in-depth interviewing	Focus groups	Action research	Participant observation
Exploring people's accounts of activities and beliefs	Exploring people's accounts of activities and beliefs Generating large amounts of data in a relatively short time Exploring *how* people come to a consensus or viewpoint	Aims to change practice as well as study it Cyclical research design where data gathering, planning, observing and reflecting all feed into the next planning cycle. Participants are therefore also the researchers Particularly useful for: • Evaluation • Assessing health needs Can use innovative research methods (e.g. drama, charts, creative arts)	Exploring what people actually do, not just what they say they do Capturing information on familiar/routine areas of life that enables views or understandings to become explicit

1D.3 ETHICAL ISSUES

The different qualitative methods each raise potential ethical concerns (see Table 1D3.1).

1D.4 VALIDITY, RELIABILITY AND GENERALISABILITY – DATA COLLECTION

The quality of research is often considered according to three dimensions:

- **Reliability** – the degree to which the collection of data in a study was consistent and **repeatable**
- **Validity** – how well a study's findings represent the '**true**' state of affairs
- **Generalisability** – the extent to which the findings from one setting can be **applied** to another.

Concepts of validity and reliability are underpinned by a **positivist perspective** on research, i.e. that there is one true answer and that a well-designed study would be able to reproduce the same findings if conducted by different

Table 1D.3.1 Ethical issues in qualitative research

Semi-structured and in-depth interviewing	Focus groups	Action research	Participant observation
Confidentiality is a particular problem in studies with small samples. The problem may be mitigated by ensuring that there are several participants with each characteristic of interest. Interviewees should be allowed to choose the degree of anonymity that they require The reporting of negative findings, or findings that are **unfavourable** to the interviewees, may need to be managed In focus groups in particular, it is important to ensure that the group atmosphere is supportive		The nature of informed consent in this context should be considered carefully, especially since there may be a blurring of roles between the researcher and the participants	If the observation is **covert** (e.g. pseudo-patient or mystery shopper), then the study participants will not have consented to the research. This may damage trust among those involved, and reduce cooperation for future studies There should be strong grounds for conducting covert observations (i.e. the study is valuable and the same data would not be obtainable if the research were conducted openly)

people at different times. Much qualitative research comes from a **constructivist** tradition, i.e. research will not yield one 'true' view of an issue, but could yield several equally valid perspectives depending on who is conducting the research and when the research was conducted. However, there are still ways to ensure that qualitative research is as rigorous as possible through considering its reliability and validity.

Reflexivity is an important dimension considered in qualitative research. This recognises that the researcher affects the environment and the people whom he or she is researching. In contrast to a positivist approach that would seek to reduce the effect of the researcher, qualitative research seeks to record and acknowledge this effect of the researcher on study findings.

RELIABILITY

Reliability can be improved by:

- Using the **same data collection tool** (interview schedule, observation or topic guide) for each interview, group or observation
- Using **one researcher** (interviewer, facilitator or observer) for all interviews/groups/observations – or else ensuring that all researchers adhere to the same guidance
- Providing the same **training** and **written guidance** to researchers to keep discussions on topic, but still enabling participants to express their views
- Producing an **accurate record** of the discussion, interview or observations through:
  - **Electronic recording** of interviews (audio or video) rather than note-taking, not only because it ensures an accurate record but also because it enables the researcher to listen fully and take part in the discussion. However, recording may not be appropriate or feasible in all situations, and it may dissuade some subjects from participating. Explicit consent must always be obtained for recording data collection
  - Noting/recording observations **as soon as possible** afterwards to ensure that they are recalled correctly
  - Writing/recording **clearly and consistently**. In participant observation studies, the researcher may be making notes 'on the hoof', and usually develops systems of short-hand or his or her own abbreviations. While this maximises the chance that the notes will be comprehensive and accurate, clear and consistent notes reduce the risk of ambiguity when notes are analysed later.

VALIDITY

Validity is reduced by biases – the nature of which depends on the methods of data collection and analysis (see Table 1D.4.1).

Table 1D.4.1 Effect of bias on validity in qualitative research

Semi-structured and in-depth interviewing	Focus groups	Action research	Participant observation
Potential for bias if the interviewees feel constrained from saying what they actually think and instead say what is socially acceptable, or what they think reflects best on them Interviewee is less likely to be constrained in a comfortable setting Consider: • Location of interview • Personal characteristics of the interviewer (e.g. age, sex, social class) • Conduct of the interviewer	Setting (group conversation) more closely reflects real life than a one-to-one interview. However, there remains the potential for bias since some group members may not say their 'true' feelings because of supposed or voiced opinions from other members of the group. The group can reach **false consensus** through a need to be agreeable with each other	Validity, reliability and generalisability all depend on what methods are used in the action research. These methods may also include quantitative techniques Interaction is not simply between the interviewer and interviewee, but between all group members. This, therefore, offers opportunities for participants to reflect and react to the views and experiences of other group members	One of the most subjective forms of qualitative data collection, findings will be influenced significantly by the researcher and his or her relationship with the person being researched To improve validity, consider: • Relationships and trust gained with community members • Breadth of observations/contacts in the community • Length of time spent in the setting • Understanding of the language and terminology of the community

GENERALISABILITY

This is the extent to which the study sample and setting are **informative** about the population. Generalisability is increased by:

- Using a credible **sample frame**. While this need not be representative (compare quantitative research), it does need to be sufficiently broad. For example, if the entire sample comes from one professional group (e.g. nurses), then the findings of the study may not reflect the opinions of other professionals (e.g. physiotherapists).
- Drawing out **themes** that are transferable to other settings. For example, a generalisable study might consider opinions on asthma rather than on asthma treatment in Cardiff Bay.

1D.5 COMMON ERRORS AND THEIR AVOIDANCE

Typical pitfalls in the conduct of qualitative studies are shown in Box 1D.5.1.

PRACTICAL ISSUES OF DATA COLLECTION

Qualitative research should not be undertaken lightly. Practical matters that should be borne in mind include those shown in Box 1D.5.2.

Box 1D.5.1

Semi-structured and in-depth interviewing and focus groups	Action research	Participant observation
Assumes that what people say accurately reflects what they think and do Research can be sidelined because it is not seen as generalisable	May focus on **changing practice** rather than on adding generalisable knowledge to the field Difficulties in securing sufficient community **participation** may be experienced Professionals may have difficulty in appropriately **devolving power** in research to the community	Research can be sidelined if it is viewed as relevant only to the setting in which it was located.

Box 1D.5.2

Semi-structured and in-depth interviewing	Focus groups	Action research	Participant observation
Relatively **easy to organise**: the interviewer only needs to fit with one other person's diary and can travel to the location preferred by the interviewee (e.g. his or her office)	Can be **difficult to organise**, and there is the danger that some invited participants may not appear Can become **expensive**: may need to provide travel expenses, crèche facilities, refreshments, book rooms, etc.	Can require training and support of those involved in research for the first time to ensure that they have the skills and confidence necessary to take part	**Time intensive** to collect data and to analyse **Expensive**: researcher may need to be in place full time for many months to collect sufficient data

Section 2

DISEASE CAUSATION AND PREVENTION; HEALTH PROMOTION

In order to improve health, practitioners need a broad understanding of what affects health and the evidence base behind approaches to improve it.

Section 2 covers a wide range of subjects, starting with an introduction to different ways of conceptualising the study of disease in **epidemiological paradigms**. The fundamentals of how disease is caused and where threats to health orginate are covered in sections on the **epidemiology** of specific diseases, environment, genetic factors, communicable disease, and **health and social behaviour**.

Finally, theory and practice of protecting and improving health is covered in sections on **diagnosis and screening**, **health promotion** and **disease prevention**.

2A
Epidemiological Paradigms

Scientific method is governed by different frameworks, which set the rules, assumptions and techniques that the scientific community tacitly accepts as valid. These frameworks or models are referred to as **paradigms**. There is a range of different epidemiological paradigms. In this chapter, epidemiology is considered in terms of **life-course** and **academic** paradigms.

Different paradigms in epidemiology offer:

- Alternative ways of viewing the same condition or population group
- Novel approaches to addressing epidemiological problems
- Insight into the effectiveness of certain interventions.

Knowledge of different paradigms enables public health practitioners to conduct insightful critical appraisal of research. By identifying the paradigm of the research method, it is possible to identify and challenge the assumptions that underlie the scientific approach used. In this way, alternative methods – and conclusions – can then be considered.

2A.1 LIFE-COURSE PARADIGMS

The three examples below are means of considering epidemiological influences during life.

PROGRAMMING

This epidemiological paradigm considers the **long-term effects of environmental exposures** during **critical periods** of growth (such as *in utero* development) upon adult diseases. For example, the **Barker hypothesis** states that an effect of *in utero* malnutrition is an increased risk of coronary heart disease in adulthood.

Programming is also known as the **critical period model** because it assumes that there is a specific period during which certain exposures will have lasting effects on the structure or function of organs.

ADULT RISK FACTOR

The adult risk factor approach, in use since the 1950s, considers the impact of behaviours, such as smoking, diet, exercise and alcohol consumption, on the onset and progression of diseases. Their relative contributory effects can be assessed using multiple regression.

LIFE-COURSE

The life-course model combines elements of the **programming** (see above) and the **adult risk factor** (see Section 2A.1) paradigms, by investigating how various biological and social factors affect health and disease in adult life independently, cumulatively and interactively.

A limitation of this paradigm is that few researchers will have access to a sufficiently wide range of biological and social data on cohorts of subjects that extend longitudinally between birth and the outcome of interest.

2A.2 ACADEMIC PARADIGMS

Paradigms may also be classified according to the academic subdisciplines, study tools and frameworks of the science of epidemiology. The conventional subdivisions of epidemiology include:

- Infectious diseases
- Chronic diseases
- Clinical epidemiology
- Psychosocial epidemiology
- Health care
- Genetic epidemiology.

Note that the tools employed in psychosocial epidemiology are rather different from those commonly used in physical and environmental epidemiology.

2B
Epidemiology of Specific Diseases

Epidemiology of specific diseases (and their risk factors) of public health significance

This section covers the occurrence and distribution of certain key diseases within the population. Clearly, it is impossible to have an encyclopaedic knowledge of all diseases. However, an overview of the epidemiology of diseases of particular public health relevance is essential for prioritising, organising and allocating resources.

These key diseases are those that are:

- Major causes of **premature death** (e.g. cardiovascular disease, cancers , trauma)
- Major causes of **morbidity** or high use of health-care resources (e.g. diabetes, asthma and chronic obstructive pulmonary disease, depression, schizophrenia, Alzheimer's disease)
- Sharply **changing** in their **incidence or prevalence** in an area of the world (e.g. obesity in developed countries)
- **Preventable** (e.g. lung cancer).

2B.1 'IMPORTANT' DISEASES

Knowledge of the defining clinical features, distribution, causes, behavioural features and determinants of diseases that currently make a significant impact on the health of local populations, with particular reference to those that are potentially preventable, or require the planned provision of health services at individual, community and structural levels, or are otherwise of particular public concern, e.g. mental health

The World Health Organization's **global burden of disease** project provides an estimate of the relative importance of all communicable and non-communicable diseases, together with intentional harms (e.g. suicide and war). The global burden of disease does not account for the degree to which illnesses are preventable or can be treated, but it does provide a useful guide to which illnesses have the greatest impact globally – and are thus of public health importance.

UK As Figure 2B.1.1 shows, neuropsychiatric conditions, cardiovascular disease and cancers account for most of the UK's burden of disease. Note also that the burden of chronic renal disease is set to increase dramatically in the UK over the coming years. This is due to several factors, including the projected increase in the number of elderly people from susceptible ethnic groups.

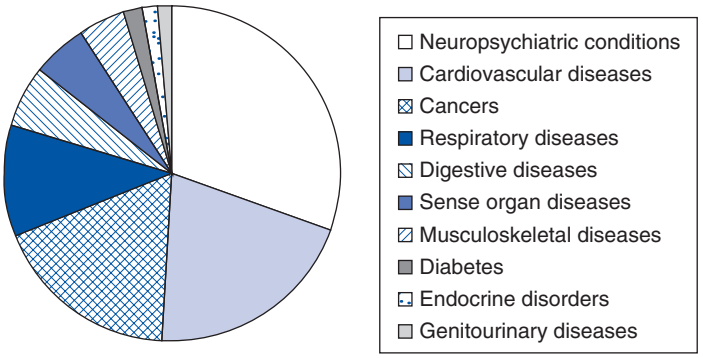


Figure 2B.1.1 Top 10: UK global burden of non-communicable disease. *Reproduced from www.who.int/healthinfo/statistics/bodgbddeathdalyestimates.xls*

The relevance of disease depends very much on context. As Figure 2B.1.2 shows, communicable disease, and maternal, perinatal and nutritional conditions are responsible for most of the mortality in Africa. In contrast, in Europe, non-communicable disease accounts for over 80% of mortality.

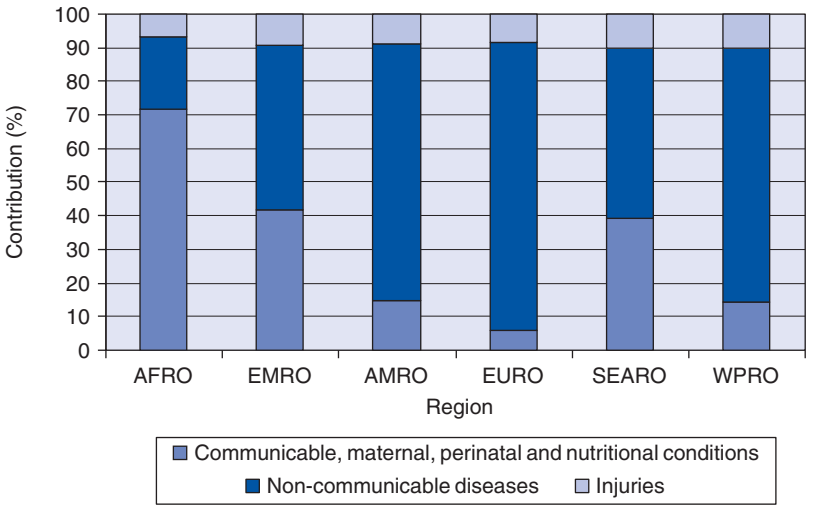


Figure 2B.1.2 Graph of relative impact of injury and disease on mortality in different regions. *Reproduced from www.who.int/healthinfo/statistics/gbdwhoregionmortality2002.xls*

AFRO, African region EMRO, eastern Mediterranean region; AMRO, the Americas; EURO, Europe; SEARO, south-east Asian region; WPRO, western Pacific region.

Similarly, the burden of disease is different across different ages and for different diseases. For major non-communicable conditions, there is some variation in the age at which disease is most common, but the burden is consistently greatest in those aged over 70, as Figure 2B.1.3 demonstrates.

PUBLIC HEALTH DISEASE KNOWLEDGE

The aspects of disease description that are important to a public health practitioner are:

- Incidence and prevalence
- Morbidity and mortality
- Importance (e.g. population-attributable risk fraction [for a risk factor] or the global burden of disease)

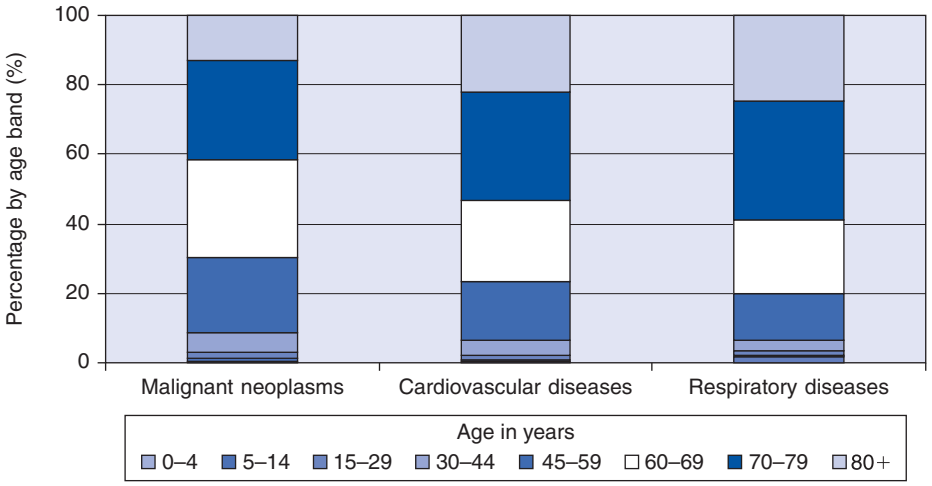


Figure 2B.1.3 Deaths in males from three major noncommunicable diseases

- Time (trends – whether the incidence is rising or falling)
- Place (whether there are areas where the disease is particularly common or rare)
- Person (the type of person who is most at risk, with regard to demographics, lifestyle, health status and workplace)
- Prevention (primary, secondary [early detection through tests/screening] and tertiary).

The aspects of non-communicable diseases that have a notable impact on the burden of disease in developed countries or are otherwise of significant public health importance are covered in Tables 2B.1.1–2B.1.11.

NEUROPSYCHIATRIC CONDITIONS

Table 2B.1.1 Depression

Clinical characteristics	Syndrome (group of symptoms) reflecting sad mood exceeding normal sadness or grief (in intensity and duration) and with functional disabilities. Symptoms include negative thoughts, moods and behaviour, and changes in bodily functions (e.g. eating, sleeping and sexual activity)
Known causes/ aetiological mechanisms	Debated: investigations carried out into neurochemical abnormalities, hormonal risk factors, genetic characteristics and adverse life events/social conditions Can develop as a side effect of some medications, e.g. β-blockers, or following physical illness or injury, e.g. diabetes, cancer, back pain or during/following pregnancy (postnatal or perinatal depression)
Public health relevance	• Most common psychological disorder • Leading cause of disability as measured by years lost due to disability (YLDs) and the fourth leading contributor to the global burden of disease (DALYs) in 2000 • Can mostly be reliably diagnosed and treated in primary care
Prevalence in the UK	Half of all women and a quarter of all men will have depressive episode in their lifetime

Table *contd overleaf*

Table 2B.1.1 *contd*

Time	Increasing incidence worldwide
Place	People living in deprived industrial areas (E&W) are more likely to be treated for depression than people living in other areas
Person	**Gender:** women > men **Age:** increasing incidence with increasing age **Ethnicity:** Western construct ➔ under-detection in some cultures **Socioeconomic status:** unemployed twice as likely as employed **Life events:** depression can follow adverse life events, e.g. divorce, bereavement, job loss **Disease:** can follow chronic or serious disease or occur during or after pregnancy (perinatal depression)
Prevention	**Secondary:** effective treatment to prevent long-term sickness, reduce risk of suicide

Table 2B.1.2 Alzheimer's disease and dementia

Clinical characteristics	Loss of memory, confusion and problems with speech and understanding Alzheimer's disease – progressive, one of the most common forms of dementia. Other common forms include vascular dementia and dementia with Lewy bodies
Known causes/ aetiological mechanisms	Loss of acetylcholine receptors and neurons in the brain occurs in Alzheimer's disease
Public health relevance	Growing in importance (and prevalence) as population ages There are no cures for dementia. Treatments for Alzheimer's disease (acetylcholinesterase inhibitors) are currently recommended by NICE for mild-to-moderate disease Dementia is costly to the NHS/Social Services in terms of drug treatment, personal care, and to society
Prevalence in the UK	One person in 20 aged over 65 years and one person in 5 over 80 years of age will develop dementia
Time	Prevalence steadily rising, as proportion of older people in the population increases
Place	Prevalence greater in countries with an older population (i.e. western Europe greater than South Asia or Africa)
Person	**Gender:** risk slightly greater for women than men **Age:** risk increases with age (see above) **Lifestyle:** smoking, sedentary lifestyle, high-fat/salt diet increases risk of vascular dementia; excessive alcohol linked to development of Korsakoff's syndrome **Disease/disability:** high blood pressure, high cholesterol and obesity; learning disability, Down's syndrome **Family:** genetic risk of early and late-onset forms of Alzheimer's disease with apolipoprotein E4 (*APOE4*) allele

Person (*contd*)	**Other**: memory and cognition exercises are protective (e.g. doing crosswords); head injury can increase risk
Prevention	**Secondary**: people with mild memory loss encouraged to use their memory/cognitive skills to preserve them

Table 2B.1.3 Schizophrenia

Clinical characteristics	A chronic, often lifelong psychotic condition or group of conditions characterised by three types of symptoms: • **Positive**: hallucinations, delusions • **Negative**: flat affect, low mood, withdrawal from social life, lack of motivation • **Cognitive**: memory, concentration problems Symptoms develop gradually – before diagnosis, a patient's behaviours and mood may have been deteriorating for some time (years). Treatment can control many of the symptoms but can often lead to severe side effects
Known causes/ aetiological mechanisms	Not known. Studies have investigated neuro-anatomical and -chemical abnormalities, including larger ventricles, imbalance in serotonin and dopamine transmission Twin studies, family studies indicate genetic element. Injury/infection during pregnancy or early in life has also been investigated
Public health relevance	High burden of disease – long-term chronic illness Higher risk of mortality from physical illness and injury (notably suicide) in people with schizophrenia Diagnosis/definition and treatment controversial and sometimes applied under legal section Cause of stigma
Prevalence in the UK	1% lifetime prevalence
Time	Greater in urban areas than rural
Place	Worldwide incidence fairly similar
Person	**Gender**: men > women **Age**: onset usually late adolescence; in women sometimes around menopause **Ethnicity**: high rates of diagnosis in African–Caribbean but could be artefact, possibly due to institutional racism **Socioeconomic status**: higher in low income groups but note social drift in people with schizophrenia as a result of illness **Lifestyle**: cannabis (but not yet causally related) **Disease**: other forms of psychosis, e.g. post-pregnancy, drug-induced psychosis (note: not schizophrenia) **Family**: higher risk if family members affected **Other**: symptoms may start to appear after stressful life events

Table *contd overleaf*

Table 2B.1.3 *contd*

Prevention	**Secondary:** symptom control with antipsychotics; psychosocial therapies used, e.g. cognitive–behavioural therapy (CBT) **Early intervention:** to promote early detection of psychosis and treatment

Table 2B.1.4 Parkinson's disease

Clinical characteristics	Progressive neurodegenerative condition leading to death of the dopamine-containing cells of the *substantia nigra*. The disease manifests through its effect on movements, such as walking, swallowing and writing
Known causes/aetiological mechanisms	Not known. Genetic studies have found several genes linked to parkinsonism, e.g. *parkin* gene
Public health relevance	Parkinson's disease is a frequent cause of falls, fractures and hospital admission, and is a costly disease, especially in the later stages Diagnosis can be difficult – there are no laboratory tests. Symptoms develop gradually and can be mistaken for normal ageing Covered by National Service Framework for Long Term (Neurological) Conditions (2005)
Prevalence in the UK	Thought to affect 1 in 500 people (though misdiagnosis is relatively common and there are undiagnosed cases)
Time	No marked geographical variation. Prevalence remains stable but mortality for patients under 75 years is decreasing
Person	**Gender:** men slightly more likely to be diagnosed than women **Age:** rising prevalence with age (up to 2% of the population aged 80 and over) Around 1 in 7 cases is diagnosed below the age of 60 years; 10% onset at 40 years **Family:** rare – in most cases sporadic **Other:** • Older antipsychotic drugs can induce symptoms of parkinsonism • Head trauma (e.g. through boxing) can increase risk of Parkinson's disease • Environmental exposures, e.g. herbicides? Possibly linked

CARDIOVASCULAR DISEASES

Table 2B.1.5 Coronary heart disease

Clinical characteristics	Angina (chest pain on exertion or under emotional stress) Myocardial infarction (MI, 'heart attack' – chest/arm pain, sweating, shortness of breath, nausea) Heart failure (ineffective pumping by the heart resulting in breathlessness, oedema, tiredness)

Known causes/aetiological mechanisms	Atherosclerosis (narrowing of the coronary arteries due to a build-up of fat) leads to angina and an MI occurs if the artery becomes occluded. Heart failure can result from subsequent damage to the heart muscle
Public health relevance	Biggest cause of premature death in England Potentially preventable Source of health inequalities: disproportionately affects people in deprived areas National Service Framework for CHD in England (2000) set eight standards for services around preventing and treating heart disease
Incidence in the UK	80–90 cases of MI per 10 000 population diagnosed each year
Time	Mortality rates reducing in UK since 1970s (halved from 1994 to 2004)
Place	Geographical inequalities: Scotland mortality greater than England; deprived areas greater than affluent
Person	**Gender**: men higher than women though risk for women increases after menopause **Age**: risk increases with age **Ethnicity**: mortality greater for south Asian population in UK, low in Chinese, African–Caribbean **Socioeconomic status**: higher mortality (in under-75s) in people on lower incomes **Lifestyle**: smoking, sedentary lifestyle, high-fat/salt diet; heavy drinking (light drinking protective) **Disease**: diabetes, high blood pressure, high cholesterol, obesity **Family**: some rare genetic forms, e.g. familial hypercholesterolaemia
Prevention	**Primary**: lifestyle measures to reduce smoking, encourage physical activity, balanced diet **Secondary**: providing treatment for high-risk people, e.g. those who have already had an MI, those with diabetes or high blood pressure

Table 2B.1.6 Stroke

Clinical characteristics	**Ischaemic stroke** (accounting for most strokes) is caused by occluded blood vessels preventing blood reaching the brain. **Haemorrhagic strokes** are caused by burst blood vessels. **Transient ischaemic attacks** (TIAs) occur when blood supply to the brain is interrupted for a short time Early signs include: numbness, weakness or paralysis on one side of the body; slurred speech or difficulty finding words; loss of sight/blurred vision; confusion or unsteadiness; headache Longer-term effects depend on the affected part of the brain, the severity of the stroke and the health of the person affected

Table *contd overleaf*

Table 2B.1.6 *contd*

Public health relevance	Third biggest killer in the UK and most common cause of disability Potentially preventable with lifestyle changes and medication
Incidence in the UK	Stroke: 174–216 people per 100 000 population per year TIAs: 35 people per 100 000 population per year
Time	Incidence declining in western Europe but numbers increasing due to ageing population
Person	**Gender:** men higher than women though risk for women increases after menopause **Age:** risk increases with age **Ethnicity:** higher risk for south Asian, African, African–Caribbean **Socioeconomic status:** higher mortality (in under-75s) in people on lower incomes **Lifestyle:** smoking, sedentary lifestyle, high-salt/fat diet, heavy drinking (affecting blood pressure) **Disease:** high blood pressure, heart disease, obesity, diabetes, high cholesterol, previous stroke or TIA **Family:** higher risk if family members have had strokes
Prevention	**Primary:** lifestyle modification to increase exercise, eat a healthy diet, reduce alcohol intake; blood pressure checks and control; cholesterol checks and control; treatment to reduce blood clotting **Secondary** (after a stroke): lifestyle modification, treatment to reduce blood clotting, blood pressure checks and control, peer support

COMMON WESTERN FORMS OF CANCER

Table 2B.1.7 Common cancers

Cancer type	Breast	Lung	Colorectal/bowel	Prostate	Cervical
Public health relevance	Major cause of mortality in women, rare in men. Better prognosis with early detection and treatment	High case fatality rate that has not decreased. Large number of cases preventable	Known risk factors. Better prognosis with early detection and treatment	Relatively high incidence in older men. Controversy over benefits of prostate-specific antigen (PSA) testing	Common cause of cancer in women. Early detection improves prognosis. Risk factors linked to social exclusion (deprivation, sexually transmitted infections)
Incidence in UK	41 000 cases/year	37 000 cases/year	34 000 cases/year	30 000 cases/year	3000 cases/year

5-year survival rate (%)	80	Men: 6; women: 7	50	71	64
Time/ place	Survival rate improving in the UK	Survival rate *not* improving	Survival rate improving. Rare in Nigeria	Survival rate improving; more die with it than from it	Second most common cancer in women aged <35
Person	**Gender:** women >> men **Age:** greater proportion of cases in older women **Deprivation:** Higher incidence in more affluent communities **Genetics:** ~5% of cancers due to genetic factors, mainly *BRCA-1* or -2	**Gender:** men >> women but increasing number of women affected **Deprivation:** death rates higher in more deprived communities **Exposure:** asbestos, radon **Lifestyle:** smoking (9/10 cases), passive smoking **Lifestyle:** diet high in fruit and vegetables may lower risk	**Gender:** third most common cancer in men; second most common in women **Age:** 90% occur in over-50s **Health:** polyp in the bowel or previous bowel cancer; obesity **Lifestyle:** diet high in meat and fat, low in fruit and vegetables. Lack of exercise increases risk	**Age:** rare under 50 years **Ethnicity:** most common in men of African descent, least common in Asian **Genetics:** rare but links to *BRCA-2* **Exposure:** radiation	**Deprivation:** higher incidence in deprived than affluent women **Health:** HPV (human papillomavirus) infection strong risk factor. Vaccine introduced in England for girls. Weakened immune system, through smoking, HIV. Poor diet increases risk **Lifestyle:** many sexual partners/sex at an early age increases risk
Tests/ screening (for more detail see 2C.1)	Screening programme via mammogram for women aged 50–70	None available (short lead time)	Screening programme: faecal occult blood test detects tiny amounts of blood in faeces	Screening via PSA but not recommended in England	Screening programme for women aged 20–64 years every 3–5 years via liquid-based cytology

RESPIRATORY DISEASES

Table 2B.1.8 Asthma

Clinical characteristics	A chronic condition marked by spells of shortness of breath, coughing or wheezing caused by inflamed and narrowed airways. Asthma can be triggered by a number of different factors, from environmental pollution to stress or smoking. Triggers for asthma vary between people
Public health relevance	Most common chronic disease in childhood; rising prevalence Emergency admissions for patients with a diagnosis of asthma are included in the Compendium of Public Health Indicators as a measure of health-care performance because, in general, admissions should be avoided through effective management in primary care or at home

Table *contd overleaf*

Table 2B.1.8 *contd*

Prevalence in the UK	Between 10 and 20% of children are diagnosed with asthma
Time **Place**	Increasing prevalence (possibly in part due to changes in diagnosis from bronchitis to asthma) but decreasing severity/disability No clear geographical pattern of asthma occurrence in the UK
Person	**Age:** more common in children than adults **Gender:** more common in boys than girls **Family/genetics:** some genetic component. Higher risk if family members affected **Risk factors** for asthma symptoms depend on the individual but commonly include: **Lifestyle:** smoking can exacerbate asthma symptoms; some people have allergies to certain foods, e.g. dairy products, preservatives. Activities such as sex and exercise can also lead to symptoms **Health:** colds, viral infections, e.g. influenza; stress **Exposures:** air pollution due to ozone, traffic fumes, smoking; house-dust mite faeces; damp environments; pollen
Prevention	Avoiding triggers People with asthma recommended to have the annual influenza vaccine Effectively treated with inhaled steroids and bronchodilators

Table 2B.1.9 Chronic obstructive pulmonary disease

Clinical characteristics	The term chronic obstructive pulmonary disease (COPD) is used to describe a range of conditions where airflow is obstructed, including chronic bronchitis and emphysema. The airflow obstruction is usually progressive, not fully reversible and does not change markedly over several months
Public health relevance	Potentially avoidable disease: most cases (80%) linked to smoking Can remain undiagnosed for years until symptoms are severe Up to 1 in 8 emergency hospital admissions may be due to COPD High burden of mortality (sixth leading cause of death worldwide) and morbidity
Prevalence in the UK	900 000 people in the UK have been diagnosed as having COPD, and half as many are thought to be living with COPD without a diagnosis
Time **Place**	Prevalence stable in men, increasing in women Mortality in the UK from respiratory disease (including COPD) almost double European average

Person	**Gender:** mortality increasing in women, falling in men **Age:** prevalence increases with increasing age **Socioeconomic status:** more prevalent in lower income groups **Lifestyle:** smoking **Other:** occupation, e.g. inhaling airborne particulates
Prevention	**Primary prevention:** stop-smoking initiatives (reducing exposures at work) **Secondary prevention:** medication to control symptoms, physiotherapy to help clear excess sputum, exercise to reduce disability, influenza and anti-pneumococcal vaccinations to reduce the risk from viral infections

OTHER DISORDERS

Table 2B.1.10 Sickle cell disease

Clinical characteristics	Genetic condition leading to abnormally formed haemoglobin that in turn leads to anaemia, severe pain, reduced immunity and sometimes chronic kidney or bone damage. Carriers (people with sickle cell trait who have only one affected haemoglobin gene) do not have symptoms but may have difficulty with activities requiring exertion or high oxygen demand (e.g. scuba diving, mountain climbing)
Known causes/aetiological mechanisms	Genetic: recessively inherited
Public health relevance	Health inequalities – in the UK mainly affects people of African and Caribbean communities Prenatal testing can be carried out by amniocentesis and chorionic villous sampling (CVS) Antenatal and neonatal screening recommended for people at risk (i.e. of ethnic origins most likely to be affected)
Prevalence in the UK	African and Caribbean descent: 1 in 10–40 has sickle cell trait, 1 in 60–200 has sickle cell disease 150–200 babies born in the UK/year with this condition
Place	Most common in people of African and Caribbean descent but also found in people from the Middle East, eastern Mediterranean (e.g. Turkey), Asia.
Person	**Age:** occurs from birth **Ethnicity:** incidence in people originally from Africa, the Caribbean, the eastern Mediterranean, Middle East and Asia **Family:** recessively inherited
Prevention	Identification of affected fetuses through screening programme

Table 2B.1.11 Diabetes

Clinical characteristics	Too much glucose in the blood. Diagnosis according to WHO is: fasting plasma glucose >7 mmol/l on two separate occasions. If not well controlled, leads to range of morbidities (heart disease, kidney failure, blindness and amputation, reduced immunity) and can be fatal. There are two types: • Type 1 – always requires insulin for treatment ('insulin-controlled diabetes') • Type 2 – most common type of diabetes (90% of all diabetes). Can be treated with diet, tablets ('tablet-controlled diabetes') or insulin ('insulin-requiring diabetes')
Known causes/ aetiological mechanisms	High levels of glucose are related to the function or presence of insulin. In type 1 diabetes, there is a shortage or absence of insulin because islet cells in the pancreas (which produce insulin) have been destroyed. It is not known what triggers their destruction In type 2 diabetes, the body produces insulin but in insufficient quantities to act, or there is resistance to insulin, preventing it from acting to reduce glucose levels
Public health relevance	Incidence increasing and younger onset as the incidence of obesity increases National Service Framework for Diabetes set 12 standards for diabetes services. Emergency admissions and deaths included in the Compendium of Public Health Indicators as a measure of health-care performance (in general, admissions and mortality should be avoided through effective management in primary care or at home)
Prevalence in the UK	4.3% in men, 3.4% in women (diabetes diagnosed by a doctor, Health Survey for England) Estimates of regional prevalence using PBS (PHO-Brent-ScHARR*) diabetes population prevalence model Diabetes registers in primary care provide estimate of diagnosed cases in GP-registered population. People have type 2 diabetes for an average of 6–7 years before diagnosis **Public Health Observatory – Brent – School of Health And Related Research*
Time **Place**	Worldwide incidence increasing but not in UK since 1999 Common in some communities, e.g. Pima Indians, Pacific Islanders, south Asian communities in the UK
Person	**Age:** type 1 onset in childhood, declining incidence thereafter; type 2 incidence in adults usually aged 40 or more **Ethnicity:** high in populations of south Asian and African–Caribbean origin (men – Bangladeshi 4x, Pakistani, Indian 3x more prevalent than general population; women: Pakistani 5x, Bangladeshi and African–Caribbean 3x, Indian 2.5x more prevalent than general population) **Socioeconomic status:** more common in those on lower incomes **Disease/health:** obesity, particularly central obesity – strong association with type 2; risk of gestational diabetes in pregnancy **Lifestyle:** inactivity/calorific diet **Family:** particularly for type 1. Also inherited forms of type 1 and 2

Prevention	**Primary:** weight management through promoting a balanced diet, regular exercise can reduce the risk of type 2 **Secondary** (morbidities arising from diabetes): exercise, self-management, medication

2C
Diagnosis and Screening

Diagnosis and screening are linked processes that span from the individual patient through to entire populations. Screening is an example of a practical public health intervention that has the potential to save thousands of lives. As with any clinical intervention, however, before offering any screening programme the first priority is to decide whether the proposal does more good than harm. This a question that is far more complex than might initially seem to be the case.

2C.1 SCREENING FOR DISEASES

Principles, methods, applications and organisation of screening for early detection, prevention, treatment and control of disease

The aim of screening is to reduce the harm caused by a disease or its complications. Screening is a service in which members of a defined population at risk of a disease are asked a question or offered a test to **identify** those **individuals** who are **more likely to be helped than harmed by further tests** or treatment. A screening test does *not* diagnose the disease: this is usually done by a subsequent *diagnostic* test.

Screening programmes vary from country to country in terms of

- Diseases that are screened
- Target age groups
- Frequency of testing
- Tests used, etc.

In addition, GPs and other clinicians may conduct 'opportunistic screening', with or without formal reward, e.g. use of urinalysis as a screening test for diabetes.

PRINCIPLES

In order to be successful, certain prerequisites regarding the **condition**, **treatment** and **programme** must be met (Table 2C.1.1). See Section 2C.9 for full details

Table 2C.1.1 Principles of screening

Condition	Treatment	Screening programme
The condition is an important health problem (i.e. common and serious) Its epidemiology is well understood It has a latent period	The condition can be treated better at an early stage There is a suitable and acceptable diagnostic test	There is robust evidence* that the screening programme reduces mortality and morbidity There are facilities to perform the further diagnostic tests required There is a favourable balance of benefit against harm The opportunity cost is justified by the benefit

**Such as randomised controlled trials.*

It must always be stressed to screened people that a negative test does *not* mean that they are necessarily disease free.

METHODS AND APPLICATIONS

UK UK National Screening Programmes

See Tables 2C.1.2–2C.1.5.

Table 2C.1.2 Antenatal screening programme

Down's syndrome	The **combined test** is available to women who request screening in the first trimester and understand the implications of doing so. It is based on combining an ultrasound measurement of nuchal translucency (NT; thickness of fluid at the nape of the fetal neck as assessed on ultrasound), human chorionic gonadotrophin (hCG), pregnancy-associated plasma protein A (PAPP-A) and the woman's age The **integrated test** provides better performance than the combined test if the woman is prepared to wait until the second trimester for the result. It integrates measurements performed at different times during pregnancy into a single test result. Typically it refers to the integration of NT and PAPP-A in the first trimester with the quadruple test in the second The **quadruple test** is offered to women who attend for screening in the second trimester, and also forms part of the integrated test as described above. It is a second trimester test consisting of α-fetoprotein (AFP), unconjugated oestriol (uE3), hCG, inhibin-A and the woman's age (the **triple test** does not include inhibin-A) If the pregnancy is deemed to be of high risk, then the woman is offered amniocentesis (or chorionic villous sampling [CVS] if <13 weeks' gestation)
Fetal anomaly scan	This is performed at 18–20 weeks' gestation. It is not universally performed in Scotland
Other	Screening for the following conditions is offered to all women: • Anaemia • Bacteriuria • Blood group and rhesus D status • HBV (hepatitis B virus) and HIV • Syphilis • Rubella immunity • Spina bifida • Pre-eclampsia • Sickle cell disease and thalassaemia (not in Scotland) Those women who are at high risk are offered screening for: • Psychiatric illness • Tay–Sachs disease

Table 2C.1.3 Newborn screening programme

Heel-prick test	Used to test for: • Phenylketonuria • Congenital hypothyroidism • Sickle cell disease and thalassaemia • Cystic fibrosis
Hearing screening programme	Oto-acoustic emissions test or attenuated brain-stem audiometry
Physical examination	Offered during first 72 hours of life to detect, among others: • Congenital heart disease • Congenital cataract • Congenital malformation • Cryptorchidism • Developmental dislocation of the hip

Table 2C.1.4 Children's screening programmes

Growth	Height and weight should be measured at the time of school entry
Hearing	Screening for hearing loss is offered to school-aged children
Vision	Screening for visual impairment is offered at age 4–5 years

Table 2C.1.5 Adult screening programmes

Disease	Screening test	Subsequent test	Age group	Frequency
Abdominal aortic aneurysm*	Abdominal ultrasound	Repeat ultrasound	Men aged 65 years	One-off
Breast cancer	Mammography	Fine-needle aspiration	50–70 (first invitation always before age 53)	3-yearly
Cervical cancer	Smear test or liquid-based cytology	Colposcopy	In England: • 25–49 years • 50–64 years In Scotland: 20–60 years	• 3-yearly • 5-yearly • 3-yearly
Bowel cancer	Faecal occult blood	Colonoscopy	60–69	2-yearly
Diabetes*	Questionnaire and body mass index (BMI) measurement	Urinalysis or oral glucose tolerance test	40+ years	To be determined
Genital *Chlamydia trachomatis*	Nucleic acid amplification test	Genitourinary clinic or outreach clinic appointment	Young men and women	Opportunistic
Prostate cancer	Prostate-specific antigen (PSA)	Prostate biopsy	Men aged 50+	Provided only on request
Vascular risk	A Vascular Risk Management Programme is being developed in which the whole population will be offered a risk assessment that could include blood pressure, cholesterol and glucose measurements			

**Being piloted in England.*

PROGRAMMES IN INDIVIDUAL COUNTRIES

Wal Scot There is an all-Wales diabetic retinopathy screening programme, which builds on experience in a programme in the Cardiff area. Scotland also has a national diabetic retinopathy screening programme.

Aus Details of Australian screening programmes are shown in Table 2C.1.6.

Aus **Table 2C.1.6** Australian screening programmes

Australian screening programme	Details
Antenatal and newborn screening	Very similar to UK
Breast cancer	Actively offered, by direct invitation, to all women aged 50–69 years at 2-yearly intervals. Breast screening is also freely available to women from age 40 years
Cervical cancer	Cervical screening is supported by a back-up record system that enables reminders, recall and comparisons with previous histology. There is current debate about changing to 3-yearly and the impact of the new HPV vaccine
Bowel cancer	Use of an immunochemical faecal occult blood test has been successfully pilot tested. Provided that adequate resources for diagnosis and treatment can be demonstrated, may be rolled out over the coming years
Other	Informal screening for melanoma (physical examination) and prostate cancer (PSA testing and digital rectal examination) occurs in the primary care setting – but neither meets the WHO criteria

NZ Screening activities in NZ include both screening programmes and opportunistic screening. The former are distinguished by the fact that all steps along the screening pathway are planned, coordinated and resourced, and are provided with quality assurance and programme monitoring.

Screening programmes in NZ

- Breast cancer screening: '*BreastScreen Aotearoa*' (BSA)
- Cervical screening: National Cervical Screening Programme (NCSP)
- Newborn Baby Metabolic Screening (for phenylketonuria, maple syrup urine disease, galactosaemia, biotinidase deficiency, congenital adrenal hyperplasia, congenital hypothyroidism, cystic fibrosis)
- Adult hepatitis B screening.

Opportunistic screening in NZ

This generally lacks the quality processes that are a feature of screening programmes and usually relies on the population of people who present to the health service for other reasons. Examples of such screening:

- Screening for hearing impairment at school entry
- Antenatal screening: anaemia; rhesus incompatibility (to avoid newborn haemolytic disease); gestational diabetes; serology for syphilis, rubella, hepatitis B; ultrasound screening for anatomical abnormalities, e.g. neural tube defects; risk factors for HIV; chromosomal abnormalities, e.g. Down's syndrome (NT ± maternal serum screening)
- Newborn physical examination, e.g. screen for congenital hip dislocation, undescended testes, cardiac abnormalities
- *Well Child* screening for developmental delays
- Screening for complications of diabetes (retinal, foot and kidney)
- Screening for breast cancer with clinical breast examination
- Mammographic breast screening outside of BSA
- Diabetes screening
- Colorectal cancer screening
- Prostate cancer screening
- Cardiovascular disease risk factor screening (smoking, serum cholesterol, hypertension)
- Screening for alcohol and drug misuse among adolescents and adults

- Osteoporosis risk factor screening (which may include bone mineral density scanning)
- Screening for congenital hearing impairment.

Adapted from National Advisory Committee on Health and Disability (2003).

2C.2 STATISTICAL ASPECTS OF SCREENING

Statistical aspects of screening tests, including knowledge of and ability to calculate, sensitivity, specificity, positive and negative predictive values, and the use of receiver operating characteristic (ROC) curves

The accuracy of a screening test is expressed in four dimensions (sensitivity, specificity, positive predictive value and negative predictive value), calculated as described in Box 2C.2.1 and Table 2C.2.1. Note how the prevalence of the condition affects some – but not all – of the test performance characteristics.

Box 2C.2.1

		Does the person **truly** have the condition?	
		YES	**NO**
Test result	**POSITIVE**	*a* (true positive)	*b* (false positive)
	NEGATIVE	*c* (false negative)	*d* (true negative)

Table 2C.2.1 Screening test dimensions, calculations and effect of high prevalence

Dimension	Definition	Calculation*	Effect of high prevalence
Sensitivity	Proportion of those people who have the disease who are correctly detected by the test	$\frac{a}{a+c}$	No effect
Specificity	Proportion of those people who do not have the disease who are correctly left undetected by the test	$\frac{d}{b+d}$	No effect
Positive predictive value (*'yield'*) (PPV)	Proportion of those testing positive who truly have the disease	$\frac{a}{a+b}$	Increases PPV
Negative predictive value (NPV)	Proportion of those testing negative who are truly disease free	$\frac{d}{c+d}$	Decreases NPV

**Note that in an examination, candidates may need to redraw the 2 × 2 matrix given so that it is in the same format as shown above – otherwise these calculations will of course be incorrect.*

The more sensitive the test, the less likely it is that a negative result will be a true positive – and hence the higher the negative predictive value.

An example is shown in Box 2C.2.2.

Box 2C.2.2

Example: Cervical cancer screening test

88 084 women are screened for cervical cancer but 73 of the 86 569 women who tested negative turned out to have the disease. What is the sensitivity, specificity, positive predictive power (PPP) and negative predictive power (NPP) of the test?

		Cervical cancer		
		Confirmed	Refuted	
Smear test	Positive	267	1248	**1515**
	Negative	73	86 496	**86 569**
		340	**87 744**	

Sensitivity	= 267 ÷ 340	= 0.785 or **78.5%**
Specificity	= 86 496 ÷ 87 744	= 0.986 or **98.6%**
PPP	= 267 ÷ 1515	= 0.176 or **17.6%**
NPP	= 86 496 ÷ 86 569	= 0.999 or **99.9%**

THRESHOLD SETTING

All screening and diagnostic tests require a **threshold** level to be defined (e.g. systolic blood pressure of 130 mmHg for a diagnosis of hypertension). Most patients will be clearly normal or abnormal, but some will remain in the grey area between the two and a line needs to be drawn by the investigator in this grey zone to determine how the test will perform when it is used. The positioning of this line will always constitute a **compromise** between sensitivity (i.e. picking up everyone who has the condition) and specificity (i.e. avoiding healthy people being labelled as positive for the test). See Box 2C.2.3.

Box 2C.2.3

Reasons for setting the line towards high sensitivity	Reasons for setting the line towards high specificity
• Serious disease with definitive treatment • Risk of infectivity to others • Subsequent diagnostic test cheap and low risk	• Unpalatable treatment • Costly, risky subsequent diagnostic test

PARALLEL AND SERIAL TESTING

Two screening tests are often necessary. These may be offered at the same time (**parallel testing**) or sequentially (**serial testing**). See Box 2C.2.4.

Box 2C.2.4

	Parallel testing	Serial testing
Procedure	Two screening tests performed at the same time and the results are subsequently combined	Second screening test is performed only if the result of the first screening test is positive
Effect	Higher sensitivity but lower specificity	Improves specificity at the cost of lower sensitivity

ROC CURVES

A **receiver operating characteristic (ROC)** curve is a plot of (sensitivity) versus (1 – specificity). The name derives from its original use in radar technology (see Figure 2C.2.1). The dotted line shown in the figure represents a useless test that has no discriminatory power.

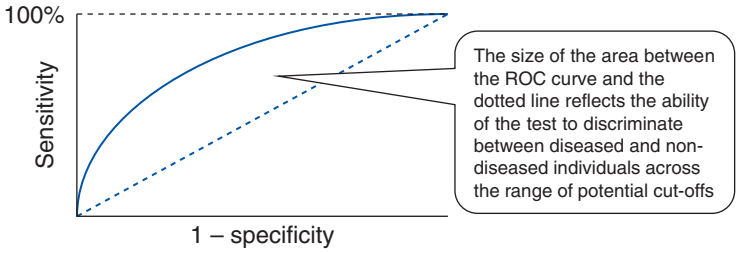


Figure 2C.2.1 ROC curve

Another, perhaps more intuitive, way of displaying this same information is shown in Figure 2C.2.2.

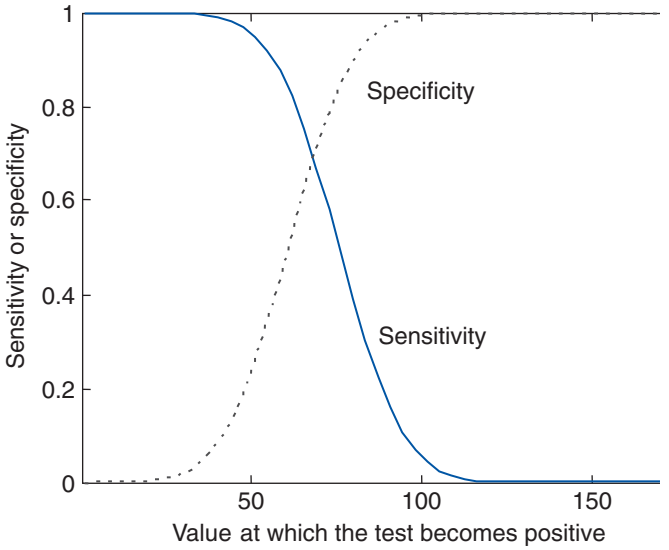


Figure 2C.2.2 Alternative depiction of ROC curve

2C.3 DIFFERENCES BETWEEN SCREENING AND DIAGNOSTIC TESTS

Screening tests are offered to asymptomatic people who may or may not have early disease or disease precursors and are used to guide whether or not a diagnostic test should be offered. **Diagnostic tests** are offered to people who have a specific indication of possible illness (a history, symptom, sign or screening test result) to determine whether or not they have the disease in question.

COMPARISON BETWEEN SCREENING AND DIAGNOSTIC TESTS

See Table 2C.3.1.

Table 2C.3.1 Differences between screening and diagnostic tests

	Screening test	Diagnostic test
Result	The cut-off is set towards high sensitivity. As a result many of the positive results are falsely positive. This is acceptable, particularly if the screening test is not harmful or expensive	The cut-off is set towards high specificity, with more weight given to diagnostic precision and accuracy than to the acceptability of the test to patients
Cost	Since large numbers of people will be screened to identify a very small number of cases, the financial resources needed must be justified carefully	Patients have symptoms that require accurate diagnosis and therefore higher costs are justified
Result of test	The result of the test is an estimate of the level of risk (e.g. risk of Down's syndrome in antenatal screening is based on a combination of maternal age, AFP, nuchal fold, etc.) and determines whether a diagnostic test (e.g. amniocentesis) is justified	The test provides a definitive diagnosis (e.g. a definite diagnosis of Down's syndrome through CVS)
Invasiveness	Often non-invasive	May be invasive
Population offered the test	Those at some risk but without symptoms of disease	Those with symptoms or who are under investigation following a positive screening test

2C.4 CASE FINDING

Case finding is a strategy for targeting resources at individuals or groups who are suspected to be at particular risk of a disease. It involves actively searching systematically for high-risk people, rather than waiting for them to present themselves to medical attention after symptoms or signs of active disease have occurred.

Note the similarities between case finding and screening: both seek to risk stratify the population using a simple and cheap procedure, and assume that better outcomes can be achieved through identifying the early stages of disease and offering prompt treatment.

Advantages and disadvantages of case finding are listed in Box 2C.4.1.

EXAMPLES OF CASE FINDING

Case finding may be used as part of the investigations into an outbreak of a communicable disease (e.g. syphilis) to identify potential sources of the disease. It may also be employed during food-borne outbreaks to identify as many at-risk individuals as possible.

Health data systems can be used to identify 'missed' risk groups (e.g. registered GP patients over 50 years of age with a BMI >30 who may not be on the register of people at risk of coronary heart disease).

Box 2C.4.1

Advantages	Disadvantages
Cheap Low personnel demand Case finding improves the positive predictive value of a diagnostic test by targeting high-risk patients with higher underlying prevalence By targeting preventive care, case-finding tools can help improve care of individuals and reduce costs for the state Cost-effective method for identifying cases of familial conditions such as familial hypercholesterolaemia	Potential to widen health inequalities because some high-risk groups are hard to reach (homeless, refugees, etc.)

ENG The King's Fund *Patients at Risk of Re-Hospitalisation (PARR)* and *Combined Model* case-finding tools use patterns in routinely collected data to forecast which individuals in a population are at high risk of emergency hospital admission in the forthcoming year.

2C.5 LIKELIHOOD RATIOS

The likelihood ratio combines the sensitivity and specificity of a test. It provides a **unified estimate** of the degree to which a test result changes the odds of having a disease:

- The likelihood ratio for a positive result (LR+) indicates by how much the odds of the disease increase when a test is positive
- The likelihood ratio for a negative result (LR−) indicates by how much the odds of the disease decrease when a test is negative.

See Box 2C.5.1.

Box 2C.5.1

Likelihood ratio of a positive test result

$$(\text{LR+}) = \frac{\text{True +ve rate}}{\text{False +ve rate}} = \frac{\text{Sensitivity}}{(1 - \text{Specificity})}$$

Likelihood ratio of a negative test result

$$(\text{LR−}) = \frac{\text{False −ve rate}}{\text{True −ve rate}} = \frac{(1 - \text{Sensitivity})}{\text{Specificity}}$$

Bayes' theorem (see Section 1B.19) states that the value of a diagnostic test is affected by the pre-test odds, i.e. by the likelihood, prior to testing, that the patient has the disease. The likelihood ratio should therefore be considered in conjunction with the following three pieces of information:

1. Prevalence of the disease
2. Characteristics of patient pool
3. Information about particular patient to determine the post-test odds of disease.

INTERPRETATION OF LIKELIHOOD RATIOS

The further away a likelihood ratio (LR) is from 1, the stronger the evidence for the presence or absence of disease.

LR >1 indicates that the test result is associated with the presence of the disease.

LR <0.1 indicates that the test result is associated with the absence of disease.

See Table 2C.5.1.

Table 2C.5.1 Interpretation of likelihood ratios

Value of likelihood ratio	Interpretation
<0.1	Large pre- to post-test changes in probability, i.e. provides strong evidence against the diagnosis in most circumstances
0.1–0.2	Moderate pre- to post-test changes in probability
0.2–0.5	Small pre- to post-test changes in probability but may be useful
0.5–2	No useful change in pre- to post-test probability
2–5	Small pre- to post-test changes in probability but may be useful
5–10	Moderate pre- to post-test changes in probability
>10	Large pre- to post-test changes in probability, i.e. provides strong evidence in favour of the diagnosis in most circumstances

USE OF LIKELIHOOD RATIOS

Note that the calculation of likelihood ratios is not currently tested as a skill in the UK MFPH Part A examination.

The post-test odds (i.e. the likelihood that a particular patient has the disease) are estimated by multiplying the pre-test odds by the likelihood ratio:

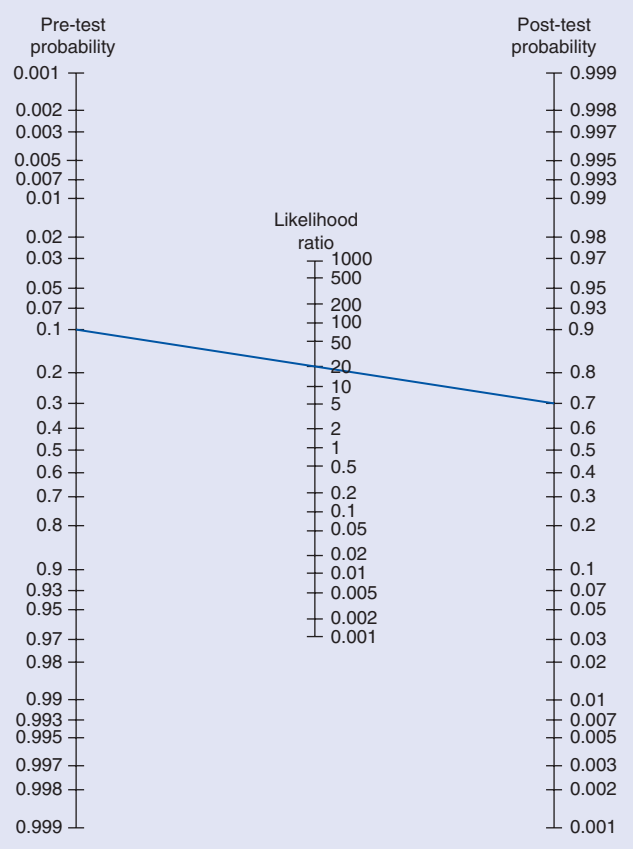
Post-test odds = pre-test odds × likelihood ratio

Factors affecting pre-test and post-test odds are shown in Box 2C.5.2.

Box 2C.5.2

Factors affecting pre-test odds	Factors affecting post-test odds
Prevalence of the disease	Prevalence of the disease in the catchment population Patient-specific patient risk factors (pre-test odds) Diagnostic test itself (the likelihood ratio)

The use of odds, rather than risks, makes the calculation of post-test odds slightly complex – but a nomogram can be used to transform odds into probabilities: see Figure 2C.5.1.



Mark the estimated probability of the disease prior to testing (e.g. 0.1) and mark the likelihood ratio for the diagnostic test (e.g. 20), then draw a line connecting the two. Extend this line until it intersects with the post-test probability to find the new estimate (0.7).

Figure 2C.5.1 Fagan's nomogram for calculating post-test probabilities. *Reproduced from Fagan (1975) with permission from the* New England Journal of Medicine

2C.6 PRE- AND POST-TEST PROBABILITY

See also Sections 2C.5 and 1B.19.

By comparing the pre- and post-test probabilities, it is possible to determine whether surety of diagnosis has risen (i.e. the post-test probability has increased) or fallen (i.e. post-test probability has decreased). In this way, it is possible to provide comprehensive information about a screening test in order to enable informed choice.

PRE-TEST PROBABILITY (OR PREVALENCE)

This is the proportion of people in the population at risk who have the disease at a specific time or time interval, i.e. the **point prevalence** or the **period prevalence** of the disease. In other words, it is the probability – before the diagnostic test is performed – that a patient has the disease. Pre-test probabilities may be estimated from routine data, practice data or clinical judgement.

POST-TEST PROBABILITY

This is the proportion of patients testing positive who truly have the disease. It is the same as the **positive predictive value**.

$$\text{Post-test probability} = \frac{\text{Post-test odds}}{1 + \text{Post-test odds}}$$

2C.7 ETHICAL, ECONOMIC, LEGAL AND SOCIAL ASPECTS OF SCREENING

Screening programmes are inherently attractive to the public and to the media because of their potential to *'nip a problem in the bud'*. This can lead to undue pressure on policy makers to introduce a screening programme without considering the opportunity costs and harms such as those from false results. For this reason, the ethical, economic and social consequences should always be carefully considered prior to the introduction of a new screening programme.

ETHICS

Beauchamp and Childress (2001) set out one of the most widely used frameworks in medical ethics first described in 1979. It consists of four principles, namely **beneficence**, **non-malfeasance**, **justice** and **autonomy** (Table 2C.7.1). Screening programmes have the potential to violate each of these.

Table 2C.7.1 Beauchamp and Childress's ethical principles applied to screening

Justice	Screening programmes should be used only when all other primary preventive measures are in place because primary prevention is likely to be more cost-effective than screening. In certain circumstances (e.g. late-presenting cancers) a screening programme does offer a cost-effective means of improving health. A challenge for those running screening programmes is to ensure that uptake is as good among the deprived populations (who are typically at greatest risk) as it is among more affluent populations Screening programmes are examples of the prevention paradox (see Section 2H.4), which states that the majority of the people receiving a preventive intervention (e.g. screening) will not benefit from that intervention. It could be argued that this is unjust
Autonomy	Communicating risk to patients is notoriously difficult, and it is questionable whether people who partake in a screening programme truly understand the consequences of their participation – especially as the parameters of a screening test are intentionally set at <100% accuracy
Beneficence	All those screened are by definition symptom free but at some risk. Screening programmes have large benefits for those who can be offered early treatment, but most will not benefit in this way
Non-malfeasance	Potential harm from a screening programme includes: • Psychological harm from false positives in the interval before diagnostic testing • Iatrogenic harm from the subsequent diagnostic test (which is often invasive) • Unwarranted reassurance from false negatives (may cause people to belittle symptoms that develop later)

ECONOMICS

A health economist may consider the opportunity costs (i.e. opportunities foregone) from the following perspectives:

- Health sector
- Patients and their families
- Other sectors.

The evaluation of a screening programme would ideally take all costs from each of these perspectives into consideration as part of a randomised controlled trial. Since this is not always possible (especially if the disease in

question manifests itself only much later), other forms of economic analysis are needed (e.g. decision-tree analysis). All evaluations should be subjected to **sensitivity analysis**, and should take account of **discounting** (see Section 4D).

Economic analyses of **adult** screening programmes should specifically address the following points:

- At what age should individuals begin to be offered screening for this disease?
- How often should individuals be screened?

LEGAL ASPECTS

See Table 2C.7.2.

Table 2C.7.2 Legal aspects of screening

Access	Most screening tests are accessed only via an authorised clinician
Accreditation	In order to assure standards, diagnosticians (cytopathologists, radiologists, etc.) must be registered and accredited by a statutory body. To retain their accreditation, they must deal with a minimum number of abnormal cases each year
Vulnerable groups	For certain diagnoses (e.g. genetic screening for Huntington's disease), the right of children *not* to be screened may be protected in law. Likewise, the law offers protection to prisoners and to people with learning difficulties against coercion into screening programmes
Confidentiality	Data from screening programmes would potentially be valuable to companies offering life insurance. For this reason the data are carefully protected

SOCIAL ASPECTS

Social aspects of a screening include those factors that may affect participation, such as those shown in Box 2C.7.1.

Box 2C.7.1

Factors that increase participation in screening	Factors that decrease participation in screening
Perception of disease **severity** Perception of **susceptibility** to the disease	Disease **phobia**

Health beliefs and attitudes are important influences on preventive behaviour, and they tend to vary between subgroups of the population. These should be considered at the planning stage in order to ensure that any new screening programme does not exacerbate health inequalities, e.g. between social classes or between different ethnic groups.

2C.8 INFORMED CHOICE

An informed choice is a decision based on accurate information regarding rights, risks and benefits that has been fully understood. In the context of screening, the decision of interest is whether or not to participate in the programme.

REQUIREMENTS FOR INFORMED CHOICE

UK The General Medical Council advises that patients require the following information in order to make an informed decision with regard to a screening programme:

- **Purpose** of the screening
- **Likelihood** of positive and negative findings
- Possibility of **false results**
- **Risks** attached to the screening process
- Any significant implications of screening for the **particular condition**: medical, social or financial
- **Follow-up** plans, including availability of counselling and support services.

2C.9 PLANNING, OPERATION AND EVALUATION OF SCREENING PROGRAMMES

The introduction of a new screening programme is a major undertaking. The infrastructure, logistics and workforce issues all need to be planned carefully, and an ongoing audit of the programme should form an integral part of the plans.

PLANNING AND OPERATION

The planning, operation and evaluation of a screening programme can all be considered with regard to the **disease**, the screening **test** and the **treatments** offered (Figure 2C.9.1).

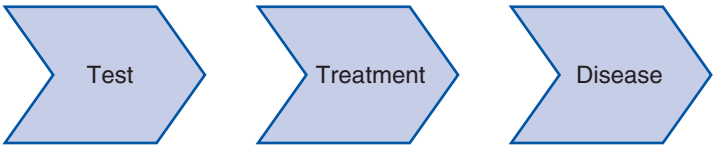


Figure 2C.9.1 Factors to consider in the planning, operation and evaluation of screening programmes

The WHO criteria for assessing the viability, effectiveness and appropriateness of a screening programme are based on those of Wilson and Jungner (1968) and are summarised in Table 2C.9.1.

EVALUATION

Evaluation of a potential screening programme involves consideration of three main issues: feasibility, effectiveness and cost.

FEASIBILITY

Feasibility will depend on:

- Relative ease with which the population can be **organised** to attend for screening
- **Acceptability** of the screening test
- Existence of facilities and resources to perform the necessary diagnostic tests and treatments **post-screening**.

EFFECTIVENESS

This is the extent to which implementing a screening programme affects the subsequent outcomes. This is difficult to measure because of three **biases** that must be borne in mind when evaluating a screening programme: see Table 2C.9.2.

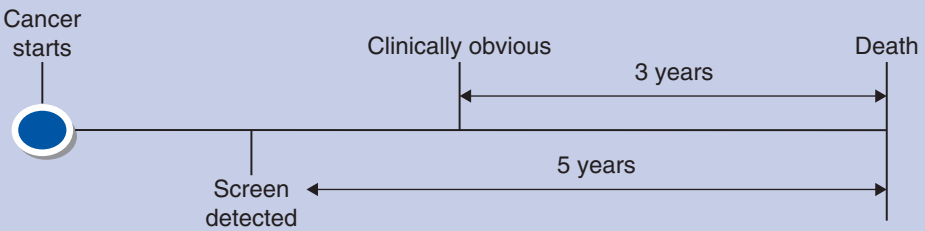
Table 2C.9.1 Wilson and Jungner's (1968) criteria for assessing screening programmes

Disease	**Important** health problem (i.e. common and serious)
	Well-recognised **pre-clinical** stage
	Natural history understood including development from latent to declared disease – detectable risk factor, disease marker
	Long **latent period** (i.e. time between first detectable signs and overt disease)
Diagnostic test	**Valid** (sensitive and specific)
	Simple and **cheap**
	Safe and acceptable
	Reliable
Diagnosis and treatment	**Agreed policy** on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals
	Facilities are adequate
	Evidence-based **policies** covering which individuals should be offered treatment and the appropriate treatment to be offered
	Effective, acceptable and safe **treatment available**
	Cost-effective
	Sustainable
Overall screening programme	**Evidence** from high-quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity
	Evidence that the complete screening programme (test, diagnostic procedures, treatment/intervention) is clinically, socially and ethically **acceptable** to health professionals and the public
	Opportunity cost of the screening programme (including testing, diagnosis and treatment) should be economically balanced in relation to expenditure on medical care as a whole
	Adequate **staffing and facilities** for testing, diagnosis, treatment and programme management should be available prior to the commencement of the screening programme
	Clear **management**, monitoring and quality assurance
	Those offered screening must be able to make **informed choices**

COST

Resources for health care will always be scarce relative to competing demands. The cost-effectiveness of a screening programme compared with other forms of health care (including public health interventions, such as stop-smoking services) should therefore be considered. Costs relate not just to the implementation of the screening programme but also to the further diagnostic tests and the subsequent costs of treatment. Moreover, the costs associated with the absence of screening must be considered too: costs would be incurred in the treatment of patients with more advanced stages of disease.

Table 2C.9.2 Types of biases in screening programmes

Type of bias	Description
Selection bias	Members of the public who participate in screening programmes **differ** systematically from those who do not (e.g. in breast screening, those who present for screening are generally from higher socioeconomic groups and have English as a first language). Performing an intention-to-treat analysis in the evaluation eliminates this form of bias
Lead-time bias	People in the screened group appear to survive longer after diagnosis than those in the unscreened group. However, screened patients will be made **aware of their disease earlier** (the lead time) and this must be subtracted from the total survival time when making comparisons with the unscreened group. This is illustrated in the diagram below in which the apparent survival time from diagnosis is longer in those detected by screening (5 years) than for those detected clinically (3 years) – whereas in fact the date of death is the same in both groups 
Length-time bias	Cases detected through screening will tend to have **less aggressive forms** of the condition. This is because fast-growing tumours will have a shorter pre-clinical stage in which to be detected by screening. Because such fast-growing tumours also have a worse prognosis, it means that survival will appear to be better in cases detected by screening than those detected clinically

2C.10 DEVELOPING SCREENING POLICIES

Evidence basis needed for developing screening policies and implementing screening programmes, including established programmes (e.g. breast and cervix) and those currently in development, being piloted or subject to major research activity (e.g. colon cancer, Chlamydia, antenatal/neonatal screening)

Evidence for screening programmes determines which of the policy options shown in Table 2C.10.1 should be taken regarding established and proposed programmes.

Table 2C.10.1 Policy options for screening programmes

Current programmes	Current programme should **continue unchanged**
	Current programme should **continue but be revised**
	Current programme should be **stopped**
Proposed programmes	Proposed programme **should be introduced**, provided that the resources, both financial and human, are available to ensure adequate quality standards
	Proposed programme **should not be introduced** (e.g. prostate screening not recommended until further evidence shows there to be a reliable test for screening purposes)

EVIDENCE BASE FOR SCREENING POLICIES

UK The National Screening Committee (NSC) advises the UK's four chief medical officers on screening policy. It considers **evidence** about the benefits, risks and costs of screening for conditions currently being investigated by research, and evaluates this evidence according to the WHO screening criteria as set out by Wilson and Jungner (see Section 2C.9). Final policy decisions on screening are based on rigorous assessment of the health technology – often by means of randomised controlled trials.

Before being rolled out nationwide, **pilots** are first established to compare the theoretical benefits of the screening shown in a research setting ('**efficacy**') with those in an ordinary service setting ('**effectiveness**'). This is important because there may be practical issues that were hitherto unapparent, e.g. staff in a service setting may not be as committed and skilled as those in a research team.

UK SCREENING PROGRAMMES

UK UK screening programmes that have been established or are in development are described in Section 2C.1. The evidence base can be accessed via the NSC website (www.nsc.nhs.uk).

NO CURRENT UK SCREENING

UK Some of the conditions in the UK for which routine screening is *not* currently recommended are listed below.

ANTENATAL

- Bacterial vaginosis
- Cystic fibrosis
- Diabetes
- Fragile X syndrome
- Genital herpes
- Hepatitis C
- HTLV-1
- Postnatal depression
- Preterm labour
- B group streptococci
- Thrombocytopenia
- Thrombophilia
- Toxoplasmosis.

NEWBORN

- Amino acid metabolism disorders
- Muscular dystrophy
- Neuroblastoma
- Organic metabolism disorders
- Thrombocytopenia.

ADULT

- Cancers: anal, bladder, lung, oral, ovarian, stomach, testicular
- Diabetes (being piloted in England)
- Hepatitis C
- Hyperlipidaemia
- Osteoporosis
- Renal disease

- Thrombophilia
- Thyroid disease
- Vision.

2C.11 ETHICAL, SOCIAL AND LEGAL IMPLICATIONS OF A GENETIC SCREENING TEST

While all screening programmes have important ethical, social and legal dimensions (see Section 2C.7), these issues are brought to the fore in the context of genetic screening: see Table 2C.11.1.

Table 2C.11.1 Ethical, social and legal implications of a genetic screening test

Ethical	Justice	Equitable **access** to genetic tests needs to be ensured
	Autonomy	Autonomy may be compromised if parents test their children for adult-onset diseases
	Beneficence	Use of genetic information can be valuable in reproductive decision-making
	Non-malfeasance	The potential for long-term psychological harm from the knowledge of a diagnosis needs to be considered
Social	Knowing the genetic profile of an individual has the potential to cause psychological impact including **stigmatisation** Society must deal with the uncertainties associated with genetic susceptibility tests for **complex conditions** (e.g. heart disease) that are linked to multiple genes and gene–environment interactions	
Legal	Certain people are incapable of providing valid **consent**, e.g. people with learning difficulties or prisoners. There are therefore legitimate concerns about the validity of the consent provided by a third party Privacy and **confidentiality** of genetic information Legislation must ensure **fairness** in the use of genetic information by insurers, employers, courts, schools, adoption agencies, the military, etc., and outlaw the potential for discrimination	

WORLD HEALTH ORGANIZATION RECOMMENDATIONS

See Table 2C.11.2.

Table 2C.11.2 WHO recommendations with regard to genetic screening (WHO 2007)

Disease	Need for universal **standard definitions**
Test	Regulatory framework and criteria for test development and use Public information and **education** **Quality assurance** Professional development Partnerships and collaborations Informed **consent** Consent procedures for children and vulnerable individuals in human genetic research
Outcome	Protection from **discrimination** **Data protection:** confidentiality, privacy and autonomy

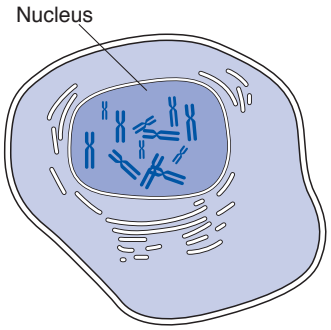
2D
Genetics

Public health genetics is the science of translating genome-based knowledge and technology into benefits at a population level. Since almost all diseases have a genetic component, genetics has a profound influence on almost all branches of public health – an influence that is set to grow rapidly in coming years as knowledge in this field expands. Links between genetic mutations and disease have been identified and can now be used to identify people at risk of certain conditions. They are also increasingly being used to develop revolutionary approaches to treatment.

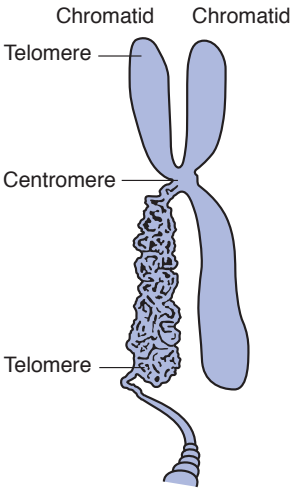
The Part A syllabus requires candidates to understand the basic principles relevant to public health genetics such as **patterns of inheritance** and **gene–environment interactions**. This chapter covers these principles and includes examples of practical relevance to public health.

2D.1 HUMAN GENETICS

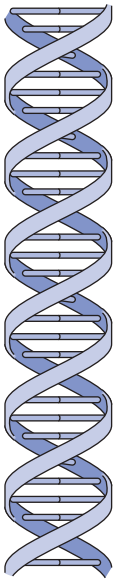
Genetics is the study of how characteristics (including conditions and diseases) are inherited. The related field of genomics considers the interactions of compete sets of genes and their products. Cells in the human body contain genetic material in the form of DNA (deoxyribonucleic acid). DNA is most commonly found in **chromosomes**, located in each cell's nucleus (see Figure 2D.1.1). There are 23 pairs of **chromosomes** per cell, of varying sizes. Twenty-two of these are **autosomal** (i.e. not sex linked) but there is one pair of **sex chromosomes** for each cell: **XX** in females or **XY** in males. Each chromosome carries a number of **genes** (segments of DNA that encode a particular enzyme or protein). It is estimated that there are between 20 000 and 25 000 genes in total in the human **genome**.



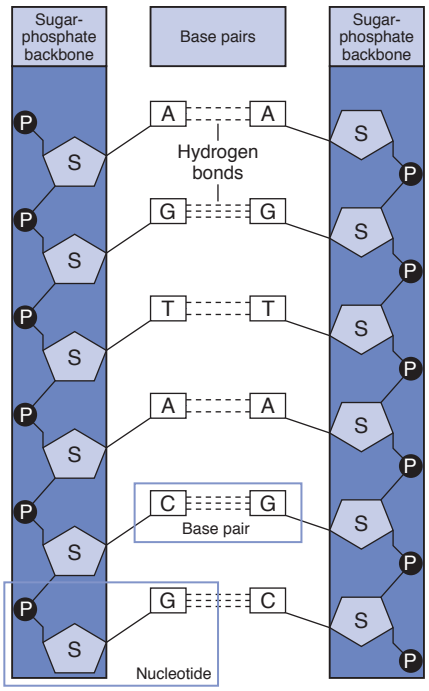
Cell – Most DNA is located in the nucleus, in chromosomes.



Chromosome – There are 46 human chromosomes, with a telomere capping the ends, and proteins called histones controlling the structure.



DNA – Stored as a double helix and is stored in each chromosome.



Each DNA sequence comprises DNA base pairs – there are around 3 million, each comprising nucleotides, either purine (A and G) or pyrimidine (C and G) bases with sugars (S) attached.

Figure 2D.1.1 Storage of genetic material in a human cell

DNA consists of four chemical **bases** (also known as '**nucleotides'**) labelled A, G, C and T. A group of three base-pairs (also known as a 'triplet nucleotide sequence' or a '**codon**') represents instructions, which ultimately produce proteins. DNA acts as a template for **RNA** (ribonucleic acid), produced by **transcription**. This in turn is **translated** into **amino acids**. These are joined together to produce **polypeptides**, which themselves are joined to produce **proteins** (see Figure 2D.1.2).

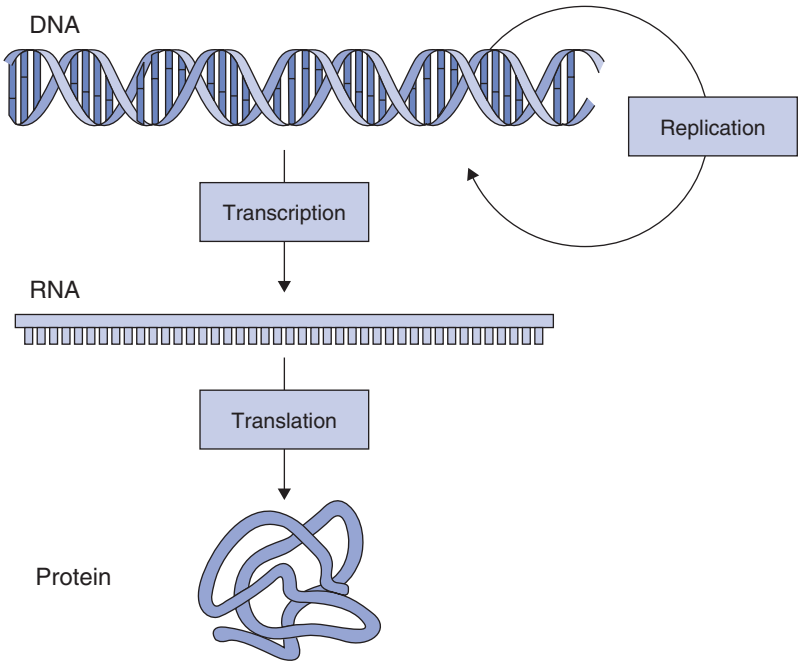


Figure 2D.1.2 Information stored as DNA is transcribed into RNA, which is then translated into proteins

Interspersed between each sequence that codes for a protein (called an **exon**) is a long sequence of DNA (called an **intron)** that does not code for proteins. Exons represent only a small proportion (2%) of the DNA in a chromosome.

VARIATIONS IN GENETIC MATERIAL

The vast bulk (99.9%) of human genetic material is identical, with only 0.1% of DNA differing between individuals (Department of Health 2003). However, the differences in DNA that do exist are responsible for individual characteristics, including propensity to different diseases.

Different forms of a gene are known as **alleles**. These are located at precise points on a chromosome and can lead to inherited variations in characteristics. A genetic **polymorphism** describes a variation in the genetic sequence that is present in at least 1–2% of the population. Most polymorphisms do not lead to disease, but some are clear risk factors or protective factors (e.g. genes coding for haemoglobin affect the severity of malaria symptoms).

Gene mutations are changes in DNA sequence that can lead to disease. They can occur due to chance, age or exposure to environmental hazards. Most genetic mutations are repaired by cells. If mutations occur on sex chromosomes, they can be passed on to an individual's offspring.

An individual carries two copies of each autosomal gene: one on each chromosome. If the person has two identical copies of a particular gene, he or she is said to be **homozygous** for that gene. If the person has different copies of the gene (i.e. different alleles), then he or she is said to be **heterozygous** for that gene.

2D.2 INHERITED DISEASES

Inherited causes of disease in populations

So-called genetic diseases may be considered in three categories summarised in Table 2D.2.1.

Table 2D.2.1 Inherited causes of disease

Single-gene disorders	One altered gene is responsible for illness, e.g. Huntington's disease, cystic fibrosis
Common disorders with rarer inherited forms	For example, 95% of cases of breast cancer occur sporadically, with a lifetime risk for women in England of around 11%. However, for women with *BRCA-1* gene mutations, the lifetime risk of breast cancer is 80%
Known genetic abnormalities or polymorphisms	In these diseases, certain genetic changes alter the risk of disease but do not predict its occurrence absolutely. The concept is akin to biological markers such as cholesterol, where occurrence of disease depends on a range of other factors, including other genetic traits and environmental exposures. Examples include hypertension and schizophrenia

2D.3 PATTERNS OF INHERITANCE

Inheritance is described as being either **mendelian** (i.e. following one of the simple patterns of inheritance that were first described by Gregor Mendel) or **non-mendelian**.

MENDELIAN INHERITANCE

Mendelian inheritance may be autosomal or sex linked, and dominant or recessive. Table 2D.3.1 describes these types of inheritance. The application of mendelian inheritance principles for maximising the effectiveness of treatment is illustrated with respect to determining fetal sex in Box 2D.3.1.

Box 2D.3.1

Example: Non-invasive method of determining fetal sex in early gestation

Fetal sex is of clinical importance in cases where the mother is a carrier of an X-linked recessive disease gene (such as haemophilia or some muscular dystrophies), since a female child will be unaffected. The sex of an unborn child can be diagnosed reliably with ultrasound only from around 12 weeks. Earlier diagnosis has to be made with CVS, which is invasive and carries a small chance of miscarriage. In the UK, polymerase chain reaction (PCR) testing is now becoming routine clinical practice for determining fetal gender safely and accurately

Technology advances: in 1997, Lo and colleagues reported detection of fetal DNA (fDNA) in maternal plasma using PCR. PCR has since been used to detect two genes found on the Y chromosome, the *SRY* gene and the marker sequence DYS14 on the *TSPY* gene. It is highly accurate (95–100% sensitivity) with an extremely low false-positive rate (~100% specificity) from as early as 6 weeks' gestation.

Other applications and ethical considerations: clinical applications of diagnosing fetal sex can extend beyond the risk of X-linked conditions. For example, it can target treatment with steroids of congenital adrenal hyperplasia (CAH), which otherwise leads to the development of ambiguous genitalia in female babies. Early identification of fetal sex enables clinicians to avoid unnecessary steroid use where the fetus is male.

Reproduced from Avent and Chitty (2006).

Table 2D.3.1 Mendelian inheritance

Inheritance pattern	Description	Example
Autosomal dominant	Dominant genetic diseases require only one copy of the gene to be abnormal in order to cause illness. If one parent has the disease, then there is a 50% chance of each child inheriting the disease. Under mendelian inheritance laws, there is no 'carrier' state for dominant conditions, because the disease affects all individuals who have an abnormal copy of the disease gene	Huntington's disease
Autosomal recessive	Both copies of the gene are required to be abnormal in order for the individual to be affected by the disorder. A normal gene is able to 'compensate' for the abnormal gene; therefore, individuals with one copy of the abnormal gene are not affected by the gene mutation, but are described as asymptomatic carriers. Where both parents carry one abnormal copy of the gene, there is a 25% chance of each child inheriting the disease. In addition there is a 50% chance of the child's being a genetic carrier for the disease	Cystic fibrosis
X-linked recessive	Here the mutation is on the X chromosome. Males have only one X chromosome and therefore a mutation in a gene on the X chromosome is more likely to cause disease in men. In contrast, a single recessive mutation on a gene on the X chromosome in women is compensated for by the normal allele on the other X chromosome so that the disease does not occur. Therefore, males are far more likely to be affected with X-linked recessive disorders, and women are more likely to be carriers. In the next generation, all daughters of affected men will be carriers of the condition, while none of their sons will be affected	Haemophilia X-linked colour blindness
X-linked dominant	X-linked dominant conditions are all rare but occur when a single copy of the gene on the X chromosome is mutated	Coffin–Lowry syndrome
Y-linked	Y-linked disorders are extremely rare	Male infertility

NON-MENDELIAN INHERITANCE

Most human diseases do not strictly follow 'mendelian' rules of inheritance. Examples of the ways in which they differ are described in Table 2D.3.2.

Table 2D.3.2 Non-mendelian inheritance

Inheritance pattern	Description	Example
Variable severity	The severity of the disease differs according to the number of abnormal genes inherited. For example, hypercholesterolaemia is more severe in homozygotes than in heterozygotes	Hypercholesterolaemia
Variable penetrance	Some dominantly inherited gene mutations are not always expressed as disease. For example, only 80% of those with *BRCA-1* gene mutations will develop breast cancer even though this is a dominant gene. See 2D.4	*BRCA-1*
Mitochondrial inheritance	As well as being present in chromosomes, DNA is also found in mitochondria. Mitochondrial DNA encodes enzymes that are responsible for energy production in a cell. Mitochondrial DNA mutations can lead to a wide range of symptoms associated with energy deficiency in cells, such as cardiac disorders and exercise intolerance Mitochondrial mutations can only be maternally inherited because mitochondria are rarely passed on in sperm. There is a threshold effect with mitochondrial inheritance whereby a child has to inherit a certain proportion of mutated DNA in order to be symptomatic. As a result, there is no predictable inheritance pattern: mothers with mutated mitochondrial DNA can give birth to both affected and non-affected children	Leigh's disease
Polygenic disorders	See Section 2D.6	Asthma, diabetes
Gene–environment interactions	See Section 2D.7	Atherosclerosis

2D.4 PENETRANCE

Penetrance describes the proportion of those who have a gene for a particular disease who develop the disease. Penetrance varies across different genes, as shown in Box 2D.4.1

Box 2D.4.1

Examples of variable gene penetrance

Huntington's disease – 100% penetrance

Certain alleles associated with Huntington's disease have **full** penetrance. If people carry one *HD* allele with 40 or more CAG triplet repeats, they will definitely develop the disease.

BRCA-1 ***or*** **BRCA-2*: female breast cancer – 80% penetrance***

Hereditary breast cancer is linked to two genes, *BRCA-1* and *BRCA-2*. These have 80% penetrance, and so four of five women carrying either of these genes will develop breast cancer.

2D.5 DIFFERENT GENOTYPES AND PHENOTYPES

The presence of a particular genetic composition (**genotype**) does not necessarily mean that the individual will manifest the particular condition (**phenotype**). Variable phenotypes arise for many reasons, including:

- Incomplete penetrance (see Section 2D.4)
- There is more than one gene relating to one protein or one characteristic (for example, cystic fibrosis; see Box 2D.5.1)
- There is more than one mutation on one gene associated with disease
- Gene–environment interactions mean that the gene is expressed only if certain environmental conditions are satisfied.

Box 2D.5.1

Example: Cystic fibrosis

Cystic fibrosis is a recessively inherited disorder that involves a mutation on the gene that codes for a cell membrane channel, called the cystic fibrosis transmembrane conductance regulator (CFTR) protein. Abnormalities in this protein cause it to malfunction, leading to symptoms including:

- Respiratory symptoms due to thickened respiratory secretions
- Pancreatic insufficiency (caused by an accumulation of thick mucus) causing malabsorption of food
- Elevated levels of chloride in the sweat
- Sterility in men

Hundreds of different mutations have been described on the *CFTR* gene, the most common being DF508 (i.e. a mutation at point 508 on the cystic fibrosis gene). The severity of disease (the phenotype) varies considerably between individuals affected by different mutations – particularly with regard to lung function. Some mutations can be **non-sense** mutations, i.e. the mutation is so severe that none of the protein is produced. Other mutations are **mis-sense**, i.e. some protein is produced but in small amounts or with altered function. Gene–environment interactions can play a part in cystic fibrosis: factors such as smoking or malnutrition are thought to influence the severity of lung disease. In the future, gene therapy may offer an effective treatment for cystic fibrosis by correcting the *CFTR* gene.

Reproduced from National Human Genome Research Institute (2007).

2D.6 POLYGENIC DISORDERS

Family studies have implicated a strong genetic component to many disorders not typically thought of as 'genetic' diseases. As well as rare single-gene causes of common diseases (such as diabetes and Alzheimer's disease), polygenic forms of these diseases are far more common.

Identifying the exact genes that underlie the polygenic forms of these disorders is useful for:

- Improving the understanding of disease **aetiology**
- **Predicting** those at risk of illness
- Targeting disease **prevention** or health promotion
- Targeting **treatments** to individuals with genetic variants most likely to benefit
- Identifying new sites or biological processes to target therapies.

The patterns of heredity for polygenic disorders – where several gene variants are needed for the disorders to develop – are complex. The mechanisms by which the genes interact are, in general, poorly understood. Simple models assume either additive properties or that different subsets of genes cause different types of disease. In reality, the ways that genes interact to cause disease are far more complicated and involve highly convoluted gene–environment interactions.

Genetic studies (see Section 1A.41) have made some progress in identifying genes implicated in polygenic disorders. However, few gene variants have been conclusively linked with particular disorders. Scientific progress has been hampered by factors such as those outlined in Table 2D.6.1. The example of type 2 diabetes (Box 2D.6.1) shows how a greater biological understanding of diabetes can help to explain the findings of genetic studies.

Table 2D.6.1 Challenges to identifying genes in polygenic disorders

Many susceptibility genes	In complex polygenic disorders there are many susceptibility genes, each contributing a small effect to overall disease susceptibility. In order to detect statistically significant gene effects, very **large sample sizes** are needed
Population heterogeneity	Studies conducted on different populations with the same disorder may not identify the same gene variants. This could be because of variations between **different population gene pools** but could also be because these populations are exposed to **different environmental circumstances**, e.g. diet, pollution
Incomplete understanding of disease biology	Candidate gene studies (where particular genes are selected for study because they code for biological functions connected with the disease) have been a fruitful approach to finding genes implicated in disorders. However, the selection of candidate genes requires an understanding of disease pathophysiology that is currently underdeveloped. **Genome-wide scans** are now possible, however, and do not require the same understanding of disease biology

Box 2D.6.1

Example: Type 2 diabetes: identifying susceptibility genes can inform disease aetiology

The incidence of type 2 diabetes has increased in recent years, thereby implicating environmental factors in the aetiology. However family studies indicate that diabetes also has a strong genetic component. Studies have identified monogenic forms of the disorder such as maturity-onset diabetes of the young (MODY types 1 and 2), but these account for <5% of all cases of diabetes. Most cases of diabetes are therefore due to a combination of factors, including intrauterine conditions, adult lifestyle habits and a range of gene–environment interactions.

Candidate gene studies have identified variations in some genes for functions known to be connected with diabetes. For example, the *PPARG* gene encodes a cell receptor involved in adipocyte development, which is the target for thiazolidinedione diabetes drugs. However, incomplete understanding of the biology of diabetes has limited the identification of candidate genes that could be involved in the disorder.

Genome-wide scans have identified other potential regions on several chromosomes, but some of these have not been replicated in other studies. One important gene to emerge from genome studies is *CAPN10*, coding for the calpain 10 protease. When calpain was first identified, there was no known biological link between calpain function and diabetes; since then, biochemical and pharmacological studies have implicated calpain in insulin secretion.

Adapted from McCarthy (2004) and Elbein (2002).

2D.7 GENE–ENVIRONMENT INTERACTIONS

Purely genetic diseases, such as Huntington's disease, are very rare. In contrast, the occurrence of almost all common diseases results from a combination of a range of genetic susceptibility factors that interact with environmental risks. Such diseases are described as **multifactorial**. Environmental factors include:

- Infections
- Chemicals
- Physical hazards
- Nutritional exposures
- Behaviours.

This implies that few people inherit a disease state genetically; rather, they inherit susceptibility to a disease, which is modified by their exposure to environmental factors. Understanding the interaction between genes and the environment will enable the identification of high-risk individuals and thus provide opportunities to target health promotion and disease prevention more effectively.

GENETIC DISEASE PREVENTION

Some genotypes are expressed only under particular environmental conditions, e.g. phenylketonuria (see Box 2D.7.1). If these environmental conditions are avoided, then individuals with this genotype will avoid the disease.

Box 2D.7.1

Example: Classic phenylketonuria (PKU)

PKU is a complete (or near-complete) deficiency of phenylalanine hydroxylase activity. It results in dietary intolerance to the amino acid phenylalanine. The disease is autosomal recessive, i.e. both parents need at least one copy of the faulty gene for a child to have a 25% chance of having the disease and a 50% chance of being a carrier. Babies with PKU are normally detected by the universal screening of newborns, but can also be detected antenatally through genetic testing.

If expressed, PKU can cause severe and irreversible learning disability. However, those with the *PKU* gene who limit their dietary intake of phenylalanine between birth and adolescence do not develop symptoms.

HEALTH PROMOTION AND GENETIC DISEASE

Identification of the genetic factors associated with common diseases, and the interplay between environmental risks, advances our understanding of the aetiology and epidemiology of common diseases. Study of which genes increase the susceptibility to disease and the gene–environment interactions enables practitioners to:

- Identify individuals at high genetic risk of disease
- Understand which environmental factors place individuals or groups (e.g. particular ethnic groups) at greater risk
- Target health promotion messages and disease prevention interventions to those most likely to benefit.

The potential for genetic information to influence heart disease prevention is demonstrated in Box 2D.7.2.

Box 2D.7.2

Example: Preventing heart disease

Health promotion to prevent heart disease usually focuses on universal messages, e.g. advising people to give up smoking, take regular exercise and eat a healthy diet. However, many people can cite individuals who followed this advice and still suffered from ischaemic heart disease in middle age, and other individuals who smoked, rarely exercised but lived well into their 80s. Such examples can undermine public acceptance of health promotion messages.

People who develop ischaemic heart disease often have a family history of the disease, and research implicates several genes in its aetiology. There are rare single-gene mutations (e.g. mutated genes that cause familial hypercholesterolaemia) that increase an individual's risk of dying prematurely from heart disease. Clearly, it is important to lower cholesterol levels in these individuals through diet and drug treatment, in order to reduce their risk of heart disease.

However, in addition to known 'high-risk' genotypes, a range of genetic variations may predispose individuals to heart disease and in turn identify candidates for targeted health promotion approaches. While smoking cessation messages are appropriate for the whole population, one study has suggested that men with a positive family history of coronary heart disease who stop smoking have the potential to decrease coronary heart diease to a greater extent than men without a positive family history.

Adapted from Hunt et al (2003) and McConnachie et al (2001).

2D.8 GENES IN HEALTH AND DISEASE

In the future, gene characteristics may be used to promote health and tackle disease. The field of **pharmacogenomics** uses individuals' genetic characteristics as a basis to understanding the relative effectiveness of pharmaceutical treatments. In particular, people metabolise drugs at different rates, and this may explain why a certain dose may be toxic for some people yet subtherapeutic for others. This may offer the opportunity to target treatment regimens based on individual genotypes. Box 2D.8.1 describes how this technology may be applied to clinical trials to understand non-response or side effects associated with particular drug doses.

Box 2D.8.1

Example: Testing for genetic variants of the cytochrome P450 *CYP* gene

CYP enzymes in the liver are primarily responsible for metabolising compounds such as drugs. Drug interactions and adverse events associated with variation in CYP enzymes are relatively common. Genetic differences in CYP may explain why one person experiences adverse events, while another does not.

In the USA, a *CYP* **genotyping test** is available which provides information on individual variants of the *CYPC19* and *CYP2D6* genes. This may enable clinicians to identify individuals who will rapidly or slowly metabolise a range of treatments and could also be added to pharmaceutical clinical trials to account for the non-response or occurrence of side effects in trial participants.

Reproduced from Davis and Khoury (2006).

GENE THERAPY

Gene therapy is an emerging field where vehicles such as viruses or plasmids are used to insert genetic material into the cells of people with a particular disease. One of the few diseases in which this technology has been applied is severe combined immunodeficiency (see Box 2D.8.2).

Box 2D.8.2

Example: Severe combined immunodeficiency (SCID)

SCID is caused by a single inherited gene abnormality in the adenosine deaminase (ADA) gene, leading to the absence of the ADA enzyme. This leaves children with severely impaired immunity to infections. Gene therapy is being developed to treat this disorder by inserting modified stem cells into the body that generate the missing enzyme.

2D.9 DISEASE IN RELATIVES

Aetiology, distribution and control of disease in relatives

Relatives of people suffering from a disease that has a genetic component may wish to know whether or not they are carrying an abnormal gene. Two types of test are particularly relevant in these circumstances: see Box 2D.9.1.

Box 2D.9.1

Method	Description	Example
Predictive genetic testing	This is of value where an individual has a family history of a disease with high penetrance that develops in adulthood. It can enable them to make informed decisions about having children and, where possible, to institute measures to reduce their risk from disease	Huntington's disease Inherited forms of cancer
Individual carrier testing	This can be offered to asymptomatic individuals who have a relative who is affected by an autosomal or X-linked genetic disorder	Cystic fibrosis

For commoner diseases, population **carrier screening** programmes are administered.

UK In the UK antenatal screening programme, pregnant women in high prevalence areas are offered a blood test to identify whether they are carriers for sickle cell disease or thalassaemia (see Section 2C.11). If the mother is identified as a carrier, then the father is also offered testing.

CONTROL OF DISEASE

In healthy relatives, control measures may include:

- **Pre-implantation genetic diagnosis in assisted conception:** one to two cells from embryos created by in vitro fertilisation (IVF) are tested for genetic diseases. Only genetically disease-free embryos are used for implantation.
- **Antenatal testing and screening**: can enable individuals to make informed decisions about whether to continue with a pregnancy if the fetus has a severe genetic disorder, e.g. Down's syndrome.
- Earlier or more **frequent cancer screening** and **prophylactic surgery**, e.g. for relatives of people with breast cancer who have *BCRA-1* or *-2* mutations or for relatives of people with colorectal cancer (see Box 2D.9.2).
- **Treatment to reduce disease risk**, e.g. cholesterol lowering drugs for those with familial hypercholesterolaemia.
- **Genetic counselling** to help people understand the disease and its potential impact, and to support people in making decisions about the future.

Box 2D.9.2

Example: Colorectal cancer

Around 30% of people with colorectal cancer have a family history of the disease. Two syndromes have been identified, accounting for 1–2.5% of inherited colorectal cancer: familial adenomatous polyposis (FAP, also known as familial polyposis coli) and hereditary non-polyposis colorectal cancer (HNPCC)

	FAP	HNPCC (Lynch syndrome)
Incidence	1 in 7000	1 in 500
Genetics	Dominant mutation in the adenomatous polyposis coli (*APC*) gene, a tumour-suppressor gene	Four genes involved in mismatch repair (*MLH1*, *MSH2*, *MSH6* and *PMS2*), most commonly *MLH1*, *MSH2*
Characteristics	Large numbers of colorectal polyps develop usually by age 30. Most are benign but a high number of polyps means that there is a very high chance that one polyp will develop a mutation in *APC* leading to a malignancy	Very few colorectal polyps, which are often malignant Average age of diagnosis of colorectal cancer is 60 years
Chance of developing cancer	Without surgery, almost all will develop cancer by age 40	70–80% lifetime risk Women also have 30–40% lifetime risk but a 42% chance of endometrial cancer
Control of disease in relatives	Regular colonoscopies to monitor polyps and detect early cancers Genetic testing	Regular colonoscopies to monitor polyps and detect early cancers Genetic testing

Reproduced from Kohlmann and Gruber (2004, updated 2006).

2D.10 MOLECULAR BIOLOGY

Elementary molecular biology as related to genetic epidemiology and microbiology

Molecular biology is the study and manipulation of biological processes and structures at a molecular level. This includes the interrelationship of DNA, RNA and proteins – and it therefore overlaps with the fields of genetics and biochemistry (see Figure 2D.10.1).

MOLECULAR BIOLOGICAL TECHNIQUES

Molecular biological techniques enable scientists to study DNA sequences, RNA, protein synthesis and cell function. Three major processes for manipulating, identifying and visualising DNA are described below.

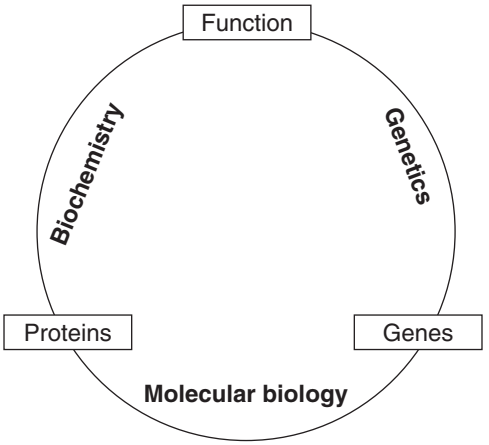


Figure 2D.10.1 Interrelationship of genetics, biochemistry and molecular biology. *Reproduced with permission from en.wikipedia.org/wiki/Image:Schematic_relationship_between_biochemistry%2C_genetics_and_molecular_biology.svg*

PRODUCING FRAGMENTS OF DNA

DNA can be cut into smaller pieces using **restriction enzymes,** which recognise and cleave DNA at particular sequences of nucleotides (see Figure 2D.10.2). The process of cutting DNA into sequences can have several uses:

- Cleaving midway through a gene will interfere with its function, and so, by observing the difference in activity after the restriction enzyme has been added, the function of that particular DNA sequence can be identified
- DNA fragments produced by restriction enzymes can be added to other organisms using vectors such as plasmids, then used to change the host's DNA and its functions.

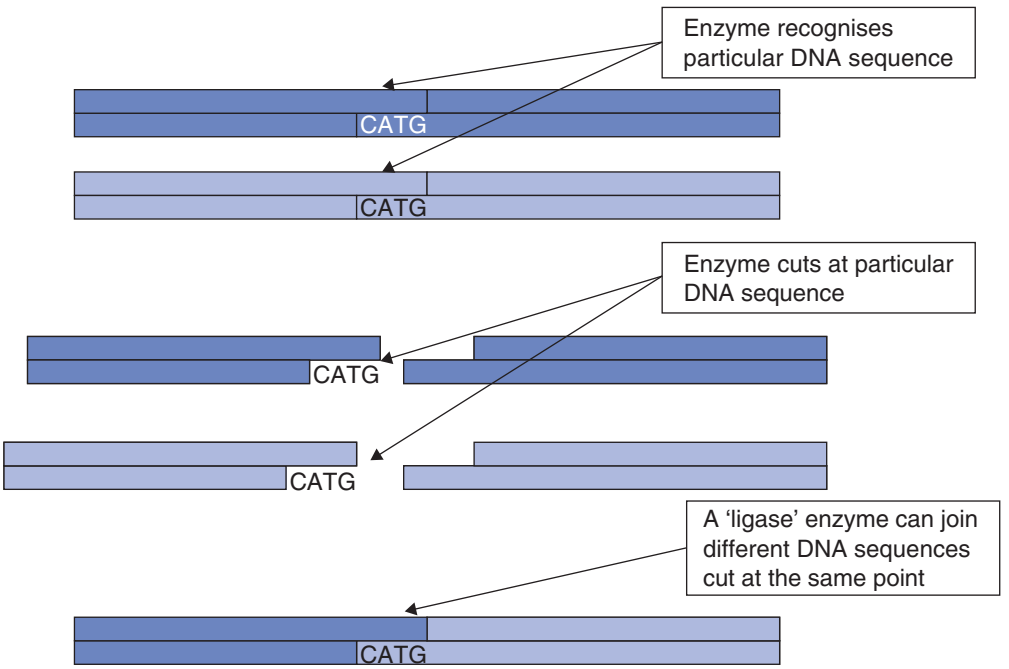


Figure 2D.10.2 Producing DNA fragments

COPYING DNA

The most versatile technique for cloning or making copies of DNA is **PCR** (the **polymerase chain reaction**), as shown in Figure 2D.10.3.

PCR reactions start with a mixture in a buffer solution comprising:

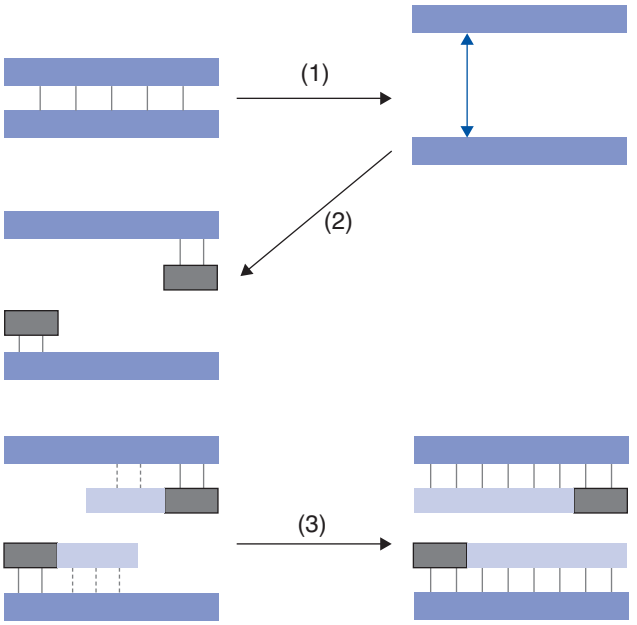
- DNA that is to be copied
- DNA polymerase (an enzyme that replicates DNA)
- Sequence of 'primer' DNA (short sequence of 15–20 nucleotides of DNA that sticks to parts of the DNA strands to be replicated but not to itself or copies of itself)
- Quantity of individual DNA nucleotides.

The process involves three steps, outlined in Table 2D.10.1.

Table 2D.10.1 Steps involved in copying DNA

Step	Description	Temperature (°C)
1. Denaturation	DNA strands are separated	93
2. Annealing	Primers attach to the separate DNA fragments	40–55
3. Extension	Nucleotides make new strands: the polymerase reaction	72

This process is repeated for 30–35 cycles (after 35 cycles more unwanted byproducts are produced than useful DNA).



(1) Denaturing: DNA strands are separated by heat.

(2) Annealing: Primers with specific DNA sequences are added with DNA polymerase to attach to single DNA strands.

(3) Extension: Individual nucleotides bind to strand fragments to make new double strands.

Figure 2D.10.3 Copying DNA: (1) denaturation; (2) annealing; (3) extension

SEPARATING AND VISUALISING DIFFERENT DNA SEQUENCES

Once DNA has been replicated, e.g. by PCR, it can be **sequenced** to identify the order of nucleotides (A, C, G or T): see Figure 2D.10.4.

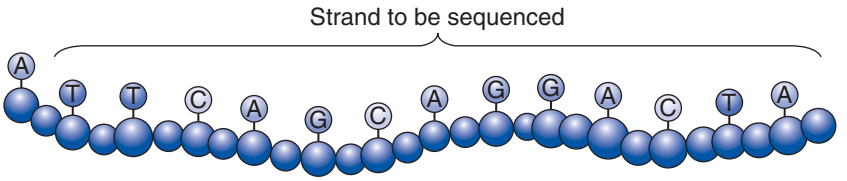


Figure 2D.10.4 Strand to be sequenced

The most common method of sequencing uses **chain termination.** A primer, containing DNA with **tagged** or **terminator** nucleotides, is added to denatured DNA strands (see Figure 2D.10.5).

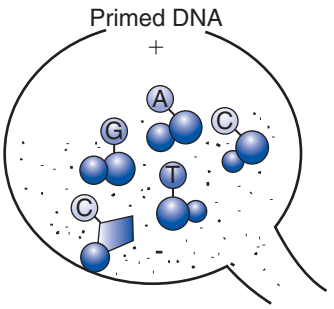


Figure 2D.10.5 DNA primer

The primer bases bind to the relevant bases on the DNA to form base-pairs. However, every time a terminator sequence binds, the strand will end, thereby creating many different lengths of sequences, all ending in the same base-pair (Figure 2D.10.6).

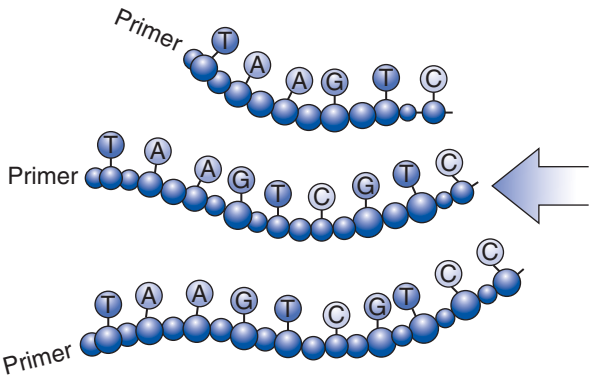


Figure 2D.10.6 DNA sequences of varying lengths

Gel electrophoresis can be used to separate and identify different DNA fragments. DNA fragments are 'run' on an agarose gel under an electric field (DNA is negatively charged so it will run towards a positive electric charge). Fragments of different sizes run along the gel at different speeds. Types of DNA can then be identified according to their final position on the gel (see Figure 2D.10.7). The same process can also be used with RNA or protein fragments.

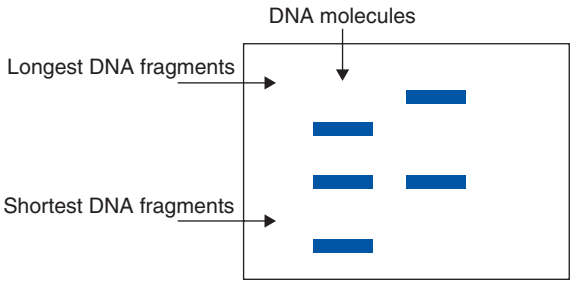


Figure 2D.10.7 Appearance of DNA fragments on gel electrophoresis

Visualising the fragments: fluorescent or radioactively labelled DNA probes are used to bind to the fragments on a gel and can then be visualised using autoradiography (for radioactively labelled probes) or phosphoimaging (for fluorescent probes).

MOLECULAR BIOLOGY APPLIED TO MICROBIOLOGY

Molecular biological techniques are invaluable in the investigation of microbial agents, their epidemiology and communicable disease control. Particular uses have included:

- Identification of different **genetic strains** of particular pathogens that affect their clinical consequences, e.g. disease severity, susceptibility to medication. For example, the Health Protection Agency's report, *Hepatitis C in England* (2005), noted that six genotypes of hepatitis C have been identified. The most common in the UK – genotype 1 – is associated with a lower response to treatment than the other genotypes.
- Improved understanding of disease aetiology through the discovery of **new pathogens**.
- Detection of trends in the occurrence of new strains of disease through the development of **molecular databases** and **surveillance systems** (see Box 2D.10.1).
- Linkage of infectious disease cases to assist the investigation of suspected or actual outbreaks.

Box 2D.10.1

Example: Applying molecular biology to identify tuberculosis strains

UK The Health Protection Agency (HPA) has introduced molecular typing for every culture of tuberculosis (TB) bacteria that tests positive in the laboratory. Positive cultures of TB are sent by local hospitals to the National Mycobacterium Reference Unit, where PCR is used to clone DNA from 15 loci on the TB genome. The analysis of the PCR products is used to:

- **Confirm** the bacterium as TB (or another related bacterium)
- Check for antibiotic **resistance**

The technology, which provides results within a few weeks, enables local HPA teams to **link TB patients** who have the same strain of TB. This can be used to **trigger an outbreak investigation** if linked patients are found to have strains with the same molecular profile. It can also be useful in **detecting false-positive TB results**, which can occur as a result of laboratory cross-contamination, and so prevent unnecessary treatment.

The HPA is developing a national database (the National Microbial Typing Database for *Mycobacterium tuberculosis*) to create a national store of all the results. The database will be useful for linking strains of TB geographically, for identifying clusters of related strains and to start linking molecular characteristics with drug susceptibility.

Reproduced from Health Protection Agency (2006).

2E
Health and Social Behaviour

The impact of different lifestyles on health is well recognised. One of the most important lifestyle choices that a person makes in this regard relates to diet: the health implications of an unhealthy diet are severe and of growing public health importance worldwide.

Influencing behaviour is a complex task that extends well beyond the boundaries of health services into home, occupation and leisure activities. Public health has a key role to play with other agencies – national and local – in interventions that improve health through social behaviour. However, it is not enough for public health practitioners just to know what constitutes a healthy lifestyle. In order to change behaviours, practitioners also need to understand:

- The barriers and motivations to adopting health lifestyles
- The evidence for recommendations related to diet and lifestyle
- How health and lifestyle are measured, particularly the potential strengths and pitfalls of these methods.

2E.1 NUTRITION

Principles of nutrition, its assessment in populations and its short- and long-term effects, and the influence of malnutrition in disease aetiology and in growth and development

Appropriate nutrition is essential for health. Most of human history was spent as hunter–gatherers and therefore evolution is thought to have rendered the body best suited to this type of diet.

ASSESSMENT OF NUTRITION IN POPULATIONS

Several techniques are available to assess a population's food intake, as outlined in Box 2E.1.1.

Box 2E.1.1 Study designs to assess population nutrition

Ecological studies

Dietary comparisons are made between countries (or in the same county over time) with respect to disease outcomes. The potential for confounding makes conclusions difficult to draw, but these studies are a fruitful source of hypotheses.

Retrospective case–control studies

Patients with a given disease are compared with non-diseased controls regarding dietary influences. Like any retrospective case–control study, these have the potential for recall bias and for disease onset to influence diet.

Cohort studies

Data on food intake and other factors are collected from a cohort of initially healthy people. Comparisons are made between individuals who develop disease with those who do not, with adjustment made for confounding factors.

Randomised trials

A nutrient or other food constituent is given to one group only. Disease outcomes in the two groups are compared. As well as being very expensive, the effect sizes seen are usually small – therefore meta-analyses are often necessary.

Dietary surveys

In the UK, the Food Standards Agency conducts food surveys such as the **Low Income Diet and Nutrition Survey**, and the **National Diet and Nutrition Survey**. These involve 24-h food diaries, physical measurements (e.g. BMI and blood pressure), blood analyses, questionnaires and interviews. Food diaries are potentially **unreliable** (participants underestimate consumption) and **biased** (obese people underestimate to a greater degree).

Biomarkers
When assessed as part of a dietary survey, biological parameters provide are more objective measure. Examples of biomarkers include serum cholesterol, serum folate, serum vitamin B_{12}, urinary iodine and urinary sodium.

SHORT-TERM NUTRITIONAL EFFECTS

Acute changes in diet may have the following effects:

- Sugary food → dental caries (the effect is particularly marked in young children)
- Reduction of salt intake → fall in blood pressure within a few weeks
- Lack of carbohydrate → ketosis.

LONG-TERM NUTRITIONAL EFFECTS

In the longer term, nutritional influences may not manifest themselves for several decades:

- Central obesity → type 2 diabetes
- Lack of breastfeeding → type 2 diabetes (immune-modulated response)
- Lack of vegetables and fruit → lung and colon cancer
- Alcohol + carcinogens ingested from smoking → oropharyngeal cancers
- Lack of dietary calcium → osteoporosis.

With regard to cancer causation, dietary factors may have an influence (either beneficial or harmful): see Figure 2E.1.1.

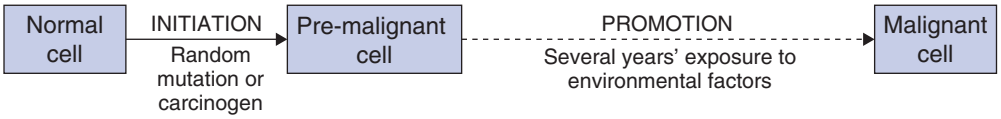


Figure 2E1.1 Influence of environmental factors on development of cancer

GROWTH AND DEVELOPMENT

Pregnancy and early childhood are both critical periods in life for nutrition.

PREGNANCY

The UK Food Standards Agency recommends that pregnant women take a daily intake of 400 µg of folate throughout the first trimester. This is known to reduce the incidence of neural tube defects, including spina bifida.

EARLY CHILDHOOD

Growth references for children are widely used in public health and paediatrics. The WHO has developed child growth standards to assess how well children worldwide are growing. The standards require that children be measured for their height and weight regularly and set healthy ranges for children's weight, height/length, BMI and motor development from birth up to age 5 years. They are based on the WHO's recommendations for:

- Child nutrition (e.g. breastfeeding should be promoted for babies wherever possible)
- Antenatal health (e.g. pregnant women should avoid tobacco, one of the risk factors for low-birthweight babies)
- Infant health care (e.g. children should receive vaccinations and health care sufficient to protect and maintain their health).

2E.2 NUTRITIONAL INTERVENTIONS

Nutrition and food, and the basis for nutritional interventions and assessment of their impact

Protein–energy malnutrition (PEM) accounts for upwards of 5 million deaths per year worldwide. At the other end of the malnutrition scale, **obesity** represents a key public health challenge to developed and transition countries, where cheap food is energy rich.

Deficiencies in iodine, vitamin A and iron are responsible for a high global burden of disease. While classic deficiency syndromes (such as scurvy or pellagra) are uncommon in developed countries, subclinical forms do exist among high-risk groups: see Table 2E.2.1.

Table 2E.2.1

Vulnerable group	Nutritional deficiency
Strict vegetarians/vegans	Vitamin B_{12} deficiency and iron deficiency (because non-haem sources of iron have low bioavailability)
Immigrants	Vitamin D deficiency
People with an alcohol dependency	Vitamin B_{12} and folate deficiency
Very young and very elderly people; pregnant and lactating women	General deficiencies

The purpose of these interventions is to reduce morbidity and mortality through dietary change. These tend to be relatively cheap interventions (low cost per QALY). Examples include:

- **Fluoridation** of water
- Interventions to reduce cardiovascular disease
- **Workplace** campaigns (e.g. for more vegetables in canteen meals)
- **Television** series about healthy lifestyles (e.g. Jamie Oliver's school meals campaign)
- Campaign to promote **home-grown** fruit and vegetables
- Collaboration with supermarkets and with the food industry (low-fat meat products, healthier vegetable-oil spreads, salt reduction in processed foods, improved **labelling** of food products)
- **Free fruit** in schools.

A range of these interventions was employed in the North Karelia Project (see Box 2E.2.1) which reduced rates of heart disease in eastern Finland. Specific interventions may be indicated for geographical deficiencies (e.g. lack of iodine in the Alps; lack of selenium in New Zealand's South Island).

Box 2E.2.1

Example: Reducing deaths from heart disease – the North Karelia Project

The North Karelia Project was established in 1972 in response to excessively high rates of heart disease in eastern Finland. The area was a traditional community, with a major dairy farming industry and a range of socioeconomic problems.

The project included a wide range of interventions focusing on diet, reducing smoking and taking more exercise. It worked at the individual patient level through health practitioners in providing disease prevention and health promotion advice, and at the population level through TV media campaigns and policy changes. Examples of interventions to affect nutrition included:

- Cholesterol-lowering **competitions** between villages and youth and school projects
- **Collaborations with the food industry** to promote low-fat dairy products and sausages and reduced salt foods
- **Encouraging farmers** to grow fruit and vegetables

Surveys conducted every 5 years since 1972 show significant changes in local diet:

1972 dietary habits	1992 dietary habits
90% used butter on bread	10% used butter on bread
Vegetable oil rarely used for cooking	30% used vegetable oil regularly
Most people used full-fat milk	Most used low fat or skimmed milk
Annual consumption of vegetables was about 20 kg per person	Annual consumption of vegetables increased to about 50 kg

Health outcomes included:

- Reduction in **risk factors** (cholesterol, blood pressure levels and percentage smoking)
- Reduction in **mortality** from heart disease among adults aged 35–64 years

Adapted from Puska et al (1983), Henkel (undated) and World Health Organization (2003a).

ASSESSMENT OF IMPACT

The impact of nutritional interventions can be assessed using several modalities:

- **Food sales** (e.g. full-fat milk compared with skimmed or semi-skimmed milk)
- **Clinical markers** (BMI, blood pressure, dental caries)
- **Biological markers** (serum cholesterol, urinary sodium)
- Cardiovascular **endpoints** (myocardial infarctions, mortality).

2E.3 CHOICE OF DIET

Social, behavioural and other determinants of choice of diet

POVERTY

Low-income families often have little disposable income to spend on food. As a result they tend to have poorer nutrient intake and less food variety, and make less healthy food choices. Particular problems include:

- Discount stores often have a smaller range of foods available with less turnover of stock, resulting in older food.
- Tinned and frozen food are cheaper than fresh food.
- Lean cuts of meat are pricier than fatty ones.
- Poorer families often eat food from packets or 'something on toast'.
- People with little money for food cannot afford to experiment with new recipes: a culinary disaster would result in hunger.

Traditionally, poverty has been more closely associated with poor diet in adults than in their children: parents protected their children, and school meals provided wholesome food. During the last 20 years, however, cooking skills and confidence have waned, as has the nutritional content of school meals.

ETHNICITY, CULTURE AND RELIGION

While many black and minority ethnic people in the UK follow Western dietary habits, eating patterns often remain influenced by:

- Traditional cuisine
- Festivals and fasts
- Proscription of certain foods
- Demographic and socioeconomic factors affecting the population as a whole.

As illustrated in Box 2E.3.1, in England recorded ethnic variations in eating habits were fond in the Health Survey for England.

Box 2E.3.1

UK Example: Differences in reported eating habits by ethnic group: the 2004 Health Survey for England

The Health Survey for England asks a sample of the population about their eating habits. The 2004 survey found that:

- A **higher** proportion of all black and minority ethnic groups, except the Irish, consumed the guideline five portions of **fruit and vegetables** than the general population
- All black and minority ethnic groups had a **lower fat intake** than the general population
- All black and minority ethnic groups, except the Irish, were more likely to **use salt** in cooking

Reproduced from Sproston and Mindel (2006).

2E.4 DIETARY RECOMMENDATIONS

Current dietary goals, recommendations and the evidence for them

The media often give conflicting, sensationalist messages about healthy diets; therefore, authoritative recommendations are crucial.

GOALS

UK The Food Standards Agency issues guidance on behalf of the government for England, Northern Ireland, Scotland and Wales.

DIETARY REFERENCE VALUES

Dietary reference values issued by the Food Standards Agency are designed for policy-makers and professionals, rather than the general public:

- Population averages for macro-nutrients and total energy
- Reference nutrient intakes (previously called 'recommended daily amount', RDA) for vitamins and minerals, i.e. 2 standard deviations above the average requirement for each nutrient
- Upper safe level for each nutrient.

RECOMMENDED DIETARY ALLOWANCES

Ire The Food Safety Authority of Ireland (FSAI) published recommended dietary allowances (RDAs) in 1999, providing indications of the level of nutrients required to meet the needs of healthy people. The FSAI adopted the EU population reference intake (PRI) values for most nutrients. However, in the case of folate, iron, calcium and vitamin C, the values for Ireland differ from the EU PRI on the basis of more recent research and consideration of prevailing Irish conditions (e.g. high incidence of childhood anaemia).

ADEQUACY OF NUTRIENT INTAKE

Surveys of nutrient intake are conducted by weighing and recording all food consumption over a 3-day period, and then assessing nutrient intake by using food composition tables. Nutrient intake can then be compared against dietary reference values.

GOOD DIETARY VARIETY

Food usage questionnaires can be used to determine a person's **variety frequency score**. This score is based upon both overall food variety and variety within each food group.

HEALTHY DIETARY PATTERNS

Indicators of healthy dietary patterns include:

- Proportions of energy in the diet that are derived from fat and saturated fat
- Healthy diet score (derived from the Frequency of Food Usage Questionnaire)
- Frequency of eating five or more fruits and vegetables per day.

IMPLEMENTATION OF THE DIETARY GOALS

Scot The Scottish Diet Action Plan was launched in July 1996. It set out 71 recommendations to improve the diets of Scots and established a series of dietary targets for achievement by 2005. The 71 recommendations were the basis on which food and health action in Scotland were shaped in the 10 years following publication of the plan.

In September 2006, Health Scotland reported on the progress, impacts, successes and challenges since the 1996 Scottish Diet Action Plan. The review found:

- A major shift in food production and attitudes to diet and health
- Considerable impact on shopping, buying and eating patterns
- To achieve a better diet, changes are still required across the supply chain and in many policy areas that shape food production and consumption.

Wal In Wales, the use of the dietary goals to reduce health inequality is promoted through the implementation of a nutrition strategy for Wales. The Food and Fitness Implementation Plan (FFIP) seeks to address issues in relation to children, and makes seven key recommendations and actions to achieve them, namely:

1. Extend the Welsh Network of Healthy School Schemes (WNHSS)
2. Improve the food and drink consumed throughout the school day (see Box 2E.4.1)
3. Provide high-quality physical education, health-related exercise and practical cookery skills
4. Provide an environment that will encourage children and young people to take up opportunities for physical activity and healthier foods
5. Develop skills to enable children and young people to take part in physical activity and to prepare healthier foods
6. Develop and deliver training on food and fitness for those working with children and young people
7. Ensure that actions are evidence based, or innovative with evaluation, and that findings are shared.

Wal Box 2E.4.1

Appetite for Life

As one example of action from the FFIP, the **Appetite for Life** report sets more stringent standards for school lunches and minimum standards for all food and drink consumed throughout the school day. It is particularly important in implementing nutritional standards that social aspects of consuming food and drink are considered: mealtimes are not just about nutrition.

Ire Implementation of Irish national recommendations regarding salt intakeis summarised in Box 2E.4.2.

Ire Box 2E.4.2

Implementing nutritional recommendations: salt intake in Ireland

The recommended dietary allowance in Ireland is approximately 4 g salt/day for adults. However, the average daily salt intake is 10 g in adults and exceeds 5 g in children aged 4–6 years, and 6 g in those aged 7–10 years, i.e. all well in excess of physiological requirements. Typical consumption is accounted for as follows:

- 15–20% of total dietary sodium intake from salt added in cooking or at the table
- 15% from naturally occurring sodium in unprocessed foods
- 65–70% from manufactured foods

Meat, fish and bread together account for over 50% of salt intake from foods. The FSAI has worked with bakers to reduce the salt content in bread and the Food Safety Promotions Board has run advertising campaigns to reduce salt consumption.

Reproduced from the Review of the Scientific Evidence and Recommendations for Public Policy in Ireland (2005), available online at: www.fsai.ie.

RECOMMENDATIONS

UK Advice to the general public from the UK Food Standards Agency is shown in Box 2E.4.3.

UK Box 2E.4.3

UK Food Standards Agency: practical advice for the general public

- **Fruit and vegetables** should constitute about a third of the food eaten each day. The ***five-a-day*** target encourages both increased variety and increased amount of fruit and vegetables eaten
- **Starchy foods** should constitute another third of the food eaten. 'Low-carbohydrate' diets are discouraged
- **Fibre** is promoted – both insoluble (wholegrain products, fruit and vegetables) and soluble (oats and legumes). The former bulk the stool and the latter reduce blood cholesterol
- **Fish** is promoted since it contains protein, selenium and iodine. White fish is low in fat; oily fish is rich in omega-3 fatty acids, vitamins A and D. Shellfish are good sources of selenium, zinc, iodine and copper. Game fish should be restricted to one portion a week (they contain mercury) and fish liver oil/fatty fish should not be eaten to excess, as over-consumption over many years is harmful. There is special advice for pregnant and breastfeeding women and for children
- **Meat** should be eaten lean and cooked with little fat. Liver or liver products should be consumed no more than once a week (again due to the risk of vitamin A excess)
- **Low-fat diet** is encouraged, and saturated fat should replace unsaturated fat. Trans-fats (contained in some hydrogenated vegetable oils) may be even worse than saturated fats
- **Sugar** intake should be restricted since it contains calories but few other nutrients and can cause tooth decay
- **Salt** should be restricted to 6 g/day (salt = 2.5 × sodium). The Food Standards Agency is working to improve labelling and wishes 'salt' to be listed rather than 'sodium'
- **Water** is encouraged: approximately 1.2 l of water should be drunk each day (6–8 glasses) – more in hot weather
- **Alcohol** intake should be restricted to 2–3 (women) or 3–4 (men) units a day, spread over the week. Postmenopausal women and men aged >40 should consider drinking 1–2 units of alcohol/day to reduce cardiovascular risk
- **Vitamins, minerals and trace elements** are usually found in sufficient amounts without recourse to supplements. The danger of over-consumption from supplements is highlighted

Ire 'Recommendations for a Food and Nutrition Policy for Ireland' (the Nutritional Advisory Group of the FSAI 1995) provided guidance on promoting healthy eating similar to that of the UK advisory agencies. The **Food Pyramid** has been used for health education, encouraging consumption of bread, cereal and potatoes (6+ portions/day), fruit and vegetables (5 portions/day), milk, cheese and yogurt (3 portions/day), meat, fish, eggs and alternative (2 portions/day) and foods high in fat and/or sugar (choose small amounts).

Aus In Australia, as elsewhere, there is an abundance of dietary advice in magazines – mostly not evidence based – against which even the most authoritative recommendations find it hard to compete. The Commonwealth Scientific and Industrial Research Organisation has recently published *The CSIRO Total Wellbeing Diet* which has become a best-selling book. This offers scientifically proven, practical advice, and contains recipes and 12 weeks of menu plans.

EVIDENCE

The evidence for nutritional recommendations comes from a range of sources. However, there are challenges to:

- Obtaining **undisputed evidence** for nutritional goals for health
- Ensuring that evidence of health risks and benefits is widely **accepted**
- Implementing accepted evidence to **change consumption** in the population.

EVIDENCE FOR HEALTH-RELATED NUTRITIONAL GOALS

Major studies that have generated evidence for the health effects of food substances are listed in Appendix C. Selected studies are also described in more depth as examples in this chapter.

There are significant ethical and practical **obstacles** to conducting randomised controlled trials involving alteration in the diet of widespread food substances such as fats. Many of the outcomes of interest (such as cardiovascular events or deaths) are long term, making trial designs expensive and impractical.

International ecological studies, such as **Intersalt** (see Box 2E.4.4) and **Seven Countries** (see Box 2E.5.1), have capitalised on the differences in dietary habits that exist between different cultures. They have been useful for indicating which factors affect health (see example boxes). However, as with all ecological studies, they are vulnerable to **ecological fallacies** (see Section 1A.7) and, as with all observational studies, there are several potentially confounding factors (e.g. different genetic susceptibility to disease, and other lifestyle variations that could explain the observed observations).

Box 2E.4.4

Example: Evidence that salt intake is linked to blood pressure

Intersalt is a prime example of a large-scale, cross-sectional study. It investigated the relationship between salt intake (measured by sodium excretion) and blood pressure. The study involved 10 079 men and women in 52 different centres around the world. For each individual, sodium excretion was measured over 24 h and was related to the individual's blood pressure.

Regression analyses performed on Intersalt data across the 52 centres and published in 1988 indicated that:

- Urinary sodium excretion was related to blood pressure (at an ecological level)
- BMI and alcohol intake were both independently and strongly associated with raised blood pressure

Intersalt's findings were controversial: commentaries (including those published by the Salt Institute) questioned the study's validity. However, subsequent analysis by the Intersalt group (and by other researchers) found an even stronger association between blood pressure and sodium levels, and further explored the relationship between alcohol and blood pressure.

A range of evidence now clearly supports a reduction in dietary salt intake as a means of reducing blood pressure; a recent Cochrane systematic review of randomised trials to determine the effect of salt reduction on blood pressure found that, *'reduction in salt intake lowers blood pressure both in individuals with elevated blood pressure and in those with normal blood pressure'*.

Adapted from Intersalt Cooperative Research Group (1988), Hanneman (1996) and He and MacGregor (2004).

ENSURING THAT EVIDENCE OF HEALTH RISKS AND BENEFITS IS WIDELY ACCEPTED

Even where evidence is strong, peer-reviewed research may be against **competing commercial interests.** Information about nutrition comes to the public from a range of sources, but rarely from peer-reviewed research sources. Common sources include publicity by manufacturers and retailers of food substances, diet books and magazines. Each of these has a vested interest in maintaining or promoting sales in a product. Advertising for foodstuffs (e.g. showing sports people consuming the product) is often far more powerful an influence on eating habits than published research is.

IMPLEMENTING ACCEPTED EVIDENCE TO CHANGE CONSUMPTION IN THE POPULATION

People's decisions on what they eat and drink are not always based on long-term health goals. Other priorities such as convenience, taste, price, satiating hunger rapidly or achieving weight loss may all come above health in influencing what they consume.

2E.5 LIFESTYLE

Effects on health of different diets (e.g. 'western diet'), physical activity, alcohol, drugs, smoking, sexual behaviour and sun exposure

In general, it is difficult to attribute health effects to a single aspect of lifestyle but Table 2E.5.1 outlines some of the relationships observed between certain factors and health. An example – cholesterol and heart disease – is given in Box 2E.5.1.

Box 2E.5.1

Example: Cholesterol and heart disease – Ancel Keys' Seven Countries Study

The origins of the Seven Countries Study lay in observations, mainly by its lead, Ancel Keys, that heart disease was a greater problem in the USA than in countries such as Italy. Keys posed the hypothesis that *'differences among populations in the frequency of heart attacks and stroke would occur in some orderly relation to physical characteristics and lifestyle, particularly composition of the diet, and especially fats in the diet'.*

Settings and methods: the Seven Countries Study involved populations of men aged between 40 and 59. They were surveyed from 1958 to 1970 on their diet and a range of risk factors for heart disease. Men were selected from 18 areas of 7 countries with contrasting dietary habits and rates of heart disease.

Results: the death rates and health of a cohort of men surveyed since 1958 were followed up. In countries where animal fat was a major component of the diet, notably Finland, deaths from heart disease were higher than in countries where olive oil was the main source, e.g. Crete. In addition to academic publications, Ancel Keys also publicised his findings in books such as *How to Eat Well and Stay Well: The Mediterranean Way* (1975).

Legacy: debate about the links between heart disease and dietary fat has continued since the Seven Countries Study and there are now links reported between other types of fat (trans-fats) and heart disease. However, the Seven Countries Study was notable for demonstrating links between heart disease and dietary fat at an ecological level, and was extremely influential in changing eating patterns across much of the USA and western Europe.

Adapted from Blackburn (1999) and Keys (1980).

2E.6 COMPLEX INTERVENTIONS

Combating complex problems using a wide range of approaches, including health service interventions and broader cultural interventions

Several methods have been adopted to tackle complex issues such as poor diet. They include the so-called **medical**, **behavioural** and **socioenvironmental** approaches.

MEDICAL APPROACH

This approach focuses on disease, with a narrow conception of the causes of disease and the determinants of health. Illness is considered in microbiological or physiological terms. Prevention focuses on known risk factors, e.g. high cholesterol and hypertension as risk factors for cardiovascular disease. Risk reduction focuses on pharmacological interventions (e.g. statins or antihypertensives), and health is equated with the absence of disease and the provision of health services.

Table 2E.5.1 Relationship between lifestyle and health outcomes

Factor		Effect on health
Diet	**Western diet** • High total energy • High saturated fat (butter, red meat) • Low fibre • High salt	Energy → obesity Fat → breast cancer, obesity Low fibre → colorectal cancer Salt → stroke
	Mediterranean diet • High unsaturated fat (olive oil) • High fruit and vegetables • Low red meat • Moderate red wine	Lower cholesterol/heart disease Obesity Lower cancer risk
	South Asian diet (in the UK) • High saturated fat (ghee) • High fruit and vegetables	Obesity Coronary heart disease Stroke Type 2 diabetes
Alcohol	**Moderate alcohol intake**	Probable cardiovascular protection
	Acute excess alcohol	Social problems Accidents Violence Road traffic accidents Peptic ulcer Liver disease Impotence Sexually transmitted infections
	Chronic excess alcohol	Cirrhosis Hepatitis Pancreatitis Suicide
Drugs		Intravenous drug misuse → blood-borne viruses (e.g. HIV/HCV) Psychosocial problems
Smoking		Cardiovascular disease and stroke Lung cancer Chronic respiratory disease At least 50 other conditions linked
Sexual behaviour		Sexually transmitted infections Unwanted pregnancy
Sun exposure		Basal cell carcinoma Squamous cell carcinoma Malignant melanoma

BEHAVIOURAL APPROACH

This approach values education and free choice, rather than legal or fiscal coercion. Disease prevention can be achieved through the provision of information to populations about lifestyle risk factors (e.g. smoking, drinking or diet). In particular, many campaigns now make use of **social marketing** to encourage people to adopt healthier eating habits (see Section 2I.3). Social marketing aims to take into account the priorities and perspectives of particular sectors of a target group for health messages and to tailor messages and health promotion campaigns accordingly.

However, the behavioural approach tends to disregard sociocultural influences on behaviour – such as the way in which dietary choices are influenced by advertising and income. Dietary education may therefore have little impact on poorer families who have no access to affordable fresh fruit and vegetables. This approach has also been criticised for blaming problems on ignorance or personal choices through the attribution of guilt.

SOCIOENVIRONMENTAL APPROACH

This approach seeks to improve health by means of strategies that modify the social, political and economic environment via government and community action. Health education involves recognition of the aspects of home, workplace and community life that are detrimental to health.

UK In recent years, the UK government has accepted this approach. It has published several documents that call for greater efforts with respect to **disease prevention** and **health promotion** through a balance of individual and government actions. An example is the 2004 Public Health White Paper entitled *Choosing Health: Making Healthy Choices Easier*.

2F Environment

There is a growing focus on the impact of human activity on the environment. Climate change, for example, is now widely considered as the most important threat to the planet. However, the impact of the environment – where we live, where we work, our air, our water, our food – is itself of profound importance to health. Public health practitioners therefore need to appreciate the environment from both perspectives:

- The **human impact on the environment**, i.e. the effects of pollution, climate change, policy levers to reduce these impacts
- **Environmental effects on humans**, i.e. the determinants of health, effects of environmental change and the range of policy levers that can be deployed to improve the environment – ranging from housing legislation to water monitoring.

2F.1 ENVIRONMENTAL DETERMINANTS OF DISEASE

As was seen in Section 2E, most diseases can be best thought of as resulting from an interaction between hereditary and environmental causes. The latter may be subdivided into the categories listed in Table 2F.1.1.

POLLUTION

Pollution is contamination of air, water or soil with harmful substances. It may be accidental or deliberate, direct or indirect.

Table 2F.1.1 Environmental determinants of disease

Pollution	Processes or activities generating pollution include energy generation and use, industry, agriculture, transport (see Section 2F.10) and waste disposal
Resources	Quality and availability of resources such as water, land, food and air
Atmosphere	Characteristics of the atmosphere including climate change (see Section 2F.3), air pollution, extreme weather
Behaviour	Human behaviours such as diet, cigarette smoking, coughing, sneezing, etc

ENERGY USE

Rising energy use is a major cause of air pollution and climate change (see Section 2F.3).

AGRICULTURE

Agricultural practices can lead to deleterious environmental effects through:

- Release of farm slurry effluent into waterways. This represents far more of a pollution hazard that does domestic sewage
- Pesticides (chemical agents used to destroy agents that affect the growth of crops) can enter the human body by inhalation, ingestion or through the skin
- Soil pollution through additives to the soil.

WASTE DISPOSAL

The **waste hierarchy** originated in the 1970s in the European Union's waste management framework. It orders options for disposing of waste from the most to the least desirable. As Figure 2F.1.1 indicates, the most desirable option is to **reduce** the amount of waste created in the first place.

There are various ways and various levels whereby the aims of the hierarchy may be achieved. For example, manufacturing industry can adopt low-waste processes, and households can use composting to dispose of wasted organic material. Where reducing use or production is not possible, re-using products should be considered (e.g. domestic re-use of carrier bags). In circumstances where products cannot be re-used, they should be recycled. Only where the other options in the hierarchy are unavailable should incineration or landfill be considered. In the UK, most solid waste is buried or incinerated, with only about 17% of domestic waste recycled. This contrasts starkly with other European countries, such as Germany, where up to half of all waste is recycled.

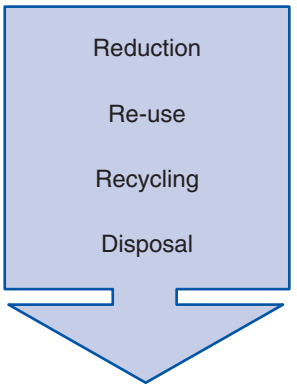


Figure 2F.1.1 The waste hierarchy

INCINERATION

Approximately 14% of UK waste is incinerated. Incineration may be considered more efficient than landfill because the combustion of waste at high temperatures produces fewer harmful chemicals. However, it is questionable whether even incinerators that generate electricity are environmentally sustainable: recycling saves considerably more energy because it entails less use of raw materials. Moreover, the ash from incinerators often contains harmful toxins, some of which (e.g. dioxins) may be carcinogenic.

LANDFILL

Eighty per cent of waste in Britain is disposed of in landfill sites. Landfill presents a number of potentially harmful environmental effects:

- Rotting rubbish emits explosive gases such as methane
- Noxious liquids from dumps risk polluting local waterways
- Additional nuisances include lorry traffic and noise, odours, smoke, dust, airborne litter and pests (e.g. rats and seagulls).

HAZARDOUS WASTE

A number of waste items are at risk of causing harm when they are discarded. These include batteries, paint tins, fertiliser, health-care waste (e.g. body parts, medicines, sharps). Inappropriate disposal of radioactive and chemical waste leads to the generation of contaminated land.

UK REGULATION

In the UK, the Environment Agency (EA) and the Scottish Environment Protection Agency (SEPA) enforce a waste management licensing system to ensure that landfill and incinerator systems for waste disposal do not have an adverse effect on health, the environment or local amenities.

RESOURCES

WATER

Water is a vital resource for agriculture, industry, drinking, food preparation and sanitation. Its quantity and quality strongly influence the risk of disease. Availability of water is an issue not just in arid countries but also in the UK (see Section 2F.5).

Water pollution may damage aquatic animal and plant life, and makes drinking water more difficult to treat. Pollutants in water are summarised in Table 2F.1.2.

Table 2F.1.2 Water pollutants

Fertilisers	**Nitrates** and **phosphates**, mostly from agricultural processes, lead to the **eutrophication** of waterways, i.e. the overgrowth of algae causing oxygen depletion
Metals	**Aluminium** may be naturally occurring, and aluminium sulphate is used in water purification to improve the taste of water **Lead** is usually derived from domestic plumbing systems Other **heavy metals** may be naturally occurring or may have leached from contaminated soils
Slurry	Organic waste includes slurries, silage liquor, surplus crops, sewage sludge and industrial wastes. These may enter the water course if they are poorly stored or are disposed or spread onto land
Sewage	In the UK, acceptable sewage regulations are enforced by the EA or SEPA

WATER MONITORING AND CONTROL

UK In the UK, the Drinking Water Inspectorate (DWI) regulates the quality of water supplied to customers. To ensure that it is safe and acceptable to drink, water undergoes tests for a number of quality characteristics listed in Table 2F.1.3.

Eng Wal The EA promotes water availability through:

- Monitoring **water levels** in ground water, reservoirs, lakes, rivers and coastal waters
- Requiring water companies to produce a long-term **water resource plan**
- Issuing of **abstraction licences** to regulate those who can take water from water bodies.

Table 2F.1.3 Water tests

Physical characteristics	Taste, colour, smell
Chemical composition	The concentration of **nitrogen** is indicative of decomposition of organic matter Chlorine levels and presence of faecal **coliforms** suggest contamination with sewage **Calcium** and magnesium are markers of hard water The absence of oxygen signals stagnant water which can indicate heavy pollution
Bacteriology	Bacteriological examination is performed to detect faecal organisms such as *E. coli*, cryptosporidia and other coliforms

Individual/domestic measures to protect water as a resource include:

- Hosepipe bans
- Short showers not baths
- Short flush toilets or 'hippo' units in cisterns to reduce the amount of water used during flushes
- Fixing dripping taps and broken pipes
- Re-using water (e.g. from washing vegetables to water plants or flush toilets)
- Collecting rainwater using water butts
- Fitting water meters.

FOOD

The quality and availability of food can lead to or prevent disease. Issues concerned with nutritional diseases are covered in Section 2E. Food can also cause disease due to:

- Residues on food from its production or packaging, e.g. pesticides
- Additives deliberately included in food
- Food production methods
- Microbes (see Section 2G).

UK The Food Standards Agency monitors or commissions monitoring on levels of contamination in food.

LAND

There are four major environmental concerns concerning land use and characteristics:

- **Housing**: see Section 2F.5.
- **Contamination**: harmful substances may have been introduced onto the land as part of its current or former uses (e.g. former industrial sites). Contaminated land is a particular problem in urban areas due to increased pressure to build on brown-field sites rather than green-field sites.
- **Radiation**: radon gas leaches naturally from the bedrock in certain regions (e.g. granite in Cornwall).
- **Deforestation**: the clearing of forests for house building or agricultural purposes has profound effects on the climate, soil erosion and water contamination.

ATMOSPHERE

The largest cause of carbon dioxide (CO_2) production is the burning of fossil fuels for electricity, heating and transport. Options to reduce CO_2 production are described in Table 2F.1.4.

Table 2F.1.4 Options to reduce carbon dioxide production

Nuclear	Nuclear energy produces minimal greenhouse gases. However, this option has its own environmental problems, namely the disposal of toxic **waste** produced by generators, the risk of harm from **explosions** or malfunctioning power stations (e.g. Chernobyl) and the risk that power stations could be used as **terrorist** targets
Efficiency	More efficient use of energy sources in homes and industry through increasing insulation (though see indoor air pollution), cleaner fuels, waste minimisation, use of more efficient appliances and engines, reductions in travel (especially by air and in sole-occupancy cars)
Renewables	Develop renewable energy sources such as solar, wind, hydroelectric, geothermic and biomass (where natural materials, such as wood or manure, are burnt or converted into gas to provide energy). These sources generate less greenhouse gas, they are regarded as being safer than nuclear energy and they will not run out – unlike fossil fuels

INDOOR AIR POLLUTION

Pollution from open solid-fuel stoves is a major public health problem in the developing world. Current efforts to improve insulation are tending to decrease ventilation, thereby exacerbating the problem. Passive exposure to cigarette smoke is an enduring problem, although many countries have now banned exposure in the workplace and other enclosed public spaces.

ATMOSPHERIC AIR POLLUTION

Table 2F.1.5 describes a number of factors responsible for atmospheric air pollution.

EU Attempts have been made to quantify the mortality and morbidity associated with vehicle pollution in European cities (see Künzli et al 2000).

MONITORING AND CONTROL

UK Air quality is monitored routinely across the UK. The UK **National Air Quality Information Service** provides information on air pollution nationally and regionally, and issues forecasts with health advice, indicating days when levels are above standard levels. The government's **Air Quality Strategy** for England and Wales sets acceptable levels of the major pollutants. Local authorities are responsible for assessing the air quality in their area, and are required to produce an action plan to address any areas where pollution is higher than health-based standards.

The **UN Economic Commission for Europe** (UNECE) supports countries' efforts to reduce pollution levels and manage their natural resources. It has also negotiated treaties that require signatory countries to reduce pollution (e.g. the Kyoto Protocol for greenhouse gases and the Montreal Protocol for volatile organic compounds).

RADIATION

Radiation – both naturally occurring and fabricated – is a cause of air pollution. There are two types of radiation. **Ionising radiation** has the ability to ionise atoms (i.e. strip them of electrons) and is characterised by high energy. Ionising radiation is emitted as alpha particles, beta particles or gamma rays. **Non-ionising radiation** does not produce sufficient energy to ionise particles. Sources of the latter include sunlight, power-lines and electrical equipment.

IONISING RADIATION

The most common source of naturally occurring ionising radiation is **radon**, a gas arising from uranium in rocks and soils that accounts for about half of UK residents' radiation dose. Naturally occurring levels are generally low and are not associated with harm. Higher radon levels are found in the south west of England due to the granite geology

Table 2F.1.5 Causes of air pollution

Carbon monoxide	Carbon monoxide (CO) is produced when fossil fuels are burnt in insufficient oxygen. Traffic exhaust fumes and cigarette smoke both contain CO. An odourless and invisible gas, CO is highly poisonous and causes many deaths each year from faulty gas appliances or blocked ventilation. Low-cost CO meters are available, but CO poisoning is best avoided by correct installation, maintenance and ventilation of gas appliances
Ozone	Ozone is found in two regions of the atmosphere: • Stratospheric ozone (the '*ozone layer*') is naturally occurring, and by blocking ultraviolet radiation prevents skin damage and skin cancers • Ground-level ozone from anthropogenic sources can cause respiratory symptoms and harms plant life. Ground-level ozone is produced in reactions between nitrogen oxides or volatile organic chemicals and sunlight
Nitrogen dioxide	This is a component of vehicle exhaust gases and power station emissions. Oxides of nitrogen may cause respiratory symptoms, increase ground-level ozone (see above) and contribute to acid rain (which damages plant life). Environmental levels remain relatively high in urban areas due to traffic congestion
Sulphur dioxide	Sulphur dioxide is mostly produced by coal- and oil-fired power stations. Sulphur dioxide causes acid rain, which damages vegetation and can exacerbate respiratory illness. Concentrations have reduced in the UK since the 1960s
Lead	Lead is produced both through lead extraction and as a by-product of other industrial processes. Lead was formerly added to petrol, but now only unleaded and low-lead fuels are available. Lead toxicity can cause cognitive impairment, renal disease and abdominal pains. Levels in the air have decreased markedly as a result of unleaded petrol
Fine particles	Particles 1000th of a millimetre in diameter are generated naturally, by industrial processes, and by traffic emissions – particularly diesel. They can exacerbate respiratory and cardiovascular disease. In the UK, levels frequently exceed health-based standards
Volatile organic compounds	These include benzene and 1,3-butadiene, which are components of traffic emissions. These chemicals are carcinogenic and they contribute to ground-level ozone and smog
Radon	See page 215

in the area and have been linked to an increased risk of lung cancer (especially in miners). Householders in areas with high radon levels should monitor and reduce the levels of radon in the home. The Health Protection Agency recommends that radon levels should not exceed 200 Bq/m^3. Actions to reduce indoor radon concentrations usually involve improving ventilation to expel radon into the atmosphere.

Naturally occurring ionising radiation accounts for 85% of the radiation that people are exposed to. The remainder is generated from three sources: see Box 2F.1.1.

Box 2F.1.1

Health care	Health care (radiotherapy and diagnostic radiology)	**14%**
Industry	Mainly from measurement purposes and for producing electricity	**<1%**
Nuclear weapons	Fallout from previous nuclear weapon explosions and other accidents and incidents worldwide	**<1%**

NON-IONISING RADIATION

Sunlight is the main source of **ultraviolet radiation** and consists of ultraviolet A and B radiation: Box 2F.1.2.

Box 2F.1.2

UVA	Skin ageing
UVB	Sunburn

Exposure to both types increases the risk of developing cataracts and skin cancer (melanoma, squamous cell skin cancer and basal cell cancer).

There are suggestions that some electromagnetic radiation may be associated with the development of cancer. Most research evidence to date suggests that any risks, if they exist at all, are probably very small. Two major areas of research (and public concern) have been mobile telephones and power lines: Box 2F.1.3.

Box 2F.1.3

Mobile telephones	The increasing use of mobile phones has been accompanied by concerns about possible harmful effects on health arising not only from exposure to the radio waves that are produced by the phones held close to the head but also from the base stations that serve the telephones
Power-lines	High voltage power-lines emit extremely low-frequency (ELF) (50–60 Hz) magnetic fields. ELF radiation may possibly increase the risk of leukaemia in children in homes with high levels of exposure

MEASURING RADIATION

As described by the US Centers for Disease Control (2003), radiation is quantified using different methods and different units, according to the amount of radiation released, the absorbed dose or the risk to health. The main methods are shown in Table 2F.1.6

Table 2F.1.6 Measuring radiation

Unit	Symbol	Measure	Illustrative examples
Becquerel	Bq	**Amount** of radioactivity in a material (unit relates to disintegrations per second)	**10 Bq/cm^2:** threshold level for contamination of surfaces requiring remediation on health grounds
Gray	Gy	**Absorbed dose**: energy deposited in each gram of tissue, indicating acute radiation damage to organs	**10 Gy:** destroys bone marrow
Sievert	Sv	**Risk of exposure** or **effective dose:** risk of cancer from chronic or low doses of radiation. Adjusts for the fact that the same absorbed dose from different types of radiation has different capacities to cause cancer	**0.1–2 mSv:** 1 chest X-ray **2 mSv:** average exposure to natural background radiation in 1 year **20 mSv:** annual dose limit for a radiation worker **1 Sv:** increases lifetime risk of cancer by approximately 5%

MONITORING AND CONTROL

UK The **Radiation Protection Division** of the Centre for Radiation, Chemical and Environmental Hazards at the Health Protection Agency monitors and researches the effects of radiation in England; the **Environment Agency** is a major source of guidance and regulation for a wide range of environmental issues in England and Wales; and **DEFRA** (the Department for the Environment, Food and Rural Affairs) is responsible for implementing legislation and regulation of chemicals and environmental threats to health in the UK. However, current UK policy is largely set by **EU directives**.

HUMAN BEHAVIOUR

Human behaviour with respect to individual health is described in Section 2E. With respect to the environment, it can be tempting to view individual human behaviour as insignificant compared with industrial impacts and global threats. However, human behaviour is of significance in three ways: see Table 2F.1.7.

Table 2F.1.7 The relationship between human behaviour and the environment

Individuals as polluters	Domestic consumption of energy has an impact on climate change. Individuals can reduce their own consumption of fossil fuels
Individual vulnerability to environmental hazards	Individuals are exposed to environmental hazards such as air pollution. In many instances, it is possible to control exposure to hazards, e.g. using sun block to limit UV exposure
Individuals as lobbyists and influencers	Individuals can lobby governments and policy-makers on policy change to reduce industrial and global pollution and consumption of natural resources. Industries also rely on individuals as employees and consumers. Individuals can also make decisions about the goods that they consume based on their environmental impact

2F.2 HAZARD AND RISK

Risk communication is an important aspect of public health, and is arguably the most important area of communication for public health practitioners. Successful risk communication relies on having an understanding of the different concepts of risk, and the skills and integrity to communicate risk information appropriately and honestly. Poor risk communication not only affects the issue under discussion but may also irreparably damage the trust of the public.

If all other factors are held constant then, technically: **Risk = Hazard x Exposure.**

However, **Sandman** (1987) proposes an alternative concept of risk that takes into account the public's response to a hazard. He defines risk as being, 'What people *feel* to be the likelihood of an event' rather than the epidemiological 'likelihood of an event'. He divides the 'risk' that people are worried about into two components:

- The technical (epidemiological) aspect of the risk. This relates to the magnitude and probability of the undesirable outcomes (e.g. an increase in the cancer rate, a catastrophic accident, number of dead fish in the river or even a decline in property values). He calls this the '**hazard**'. Thus a hazard can influence how society perceives risks.
- The non-technical aspects of the risk. These are the perceived negative features of the situation itself (as opposed to those of the outcomes). Such features are listed below and together they constitute the '**outrage**'.

Sandman's concept is the sum of these two aspects: **Risk = Hazard + Outrage.**

In his opinion, risk communication depends on securing an appropriate degree of 'outrage' in the public, so that they are neither frightened unnecessarily nor apathetic about real problems. Sandman (1987) lists nine factors that can increase or decrease the level of outrage: see Table 2F.2.1.

Table 2F.2.1 Sandman's factors affecting public outrage and hazard perception

Voluntariness	A voluntary risk is much more acceptable to people than a coerced risk, because it generates no outrage. Consider the difference between being pushed into a swimming pool and diving into a swimming pool
Control	Most people feel safer driving a car than sitting in the passenger seat. When prevention and mitigation are in the individual's hands, the risk (though not the hazard) is seen as being lower than when they are in the hands of a government agency
Fairness	When people feel that they are subjected to greater risks than their neighbours, without access to greater benefits, they are naturally outraged – especially if the rationale for the extra burden is perceived to be due to political, rather than scientific, reasons
Process	Outrage is affected by public perception of the authority or government body, i.e. whether it is perceived as being trustworthy or dishonest, concerned or arrogant. Perceptions can be improved by proactively informing the public and by responding to the concerns of the community
Morality	Where a risk has acquired a moral dimension (e.g. childhood cancer), discussion of cost–risk trade-offs is unacceptably callous. Imagine a public health specialist arguing that an occasional child death is an 'acceptable risk'
Familiarity	Novel or exotic risks provoke more outrage than do familiar risks
Memorability	Memorable incidents (e.g. Chernobyl, Bhopal, Three Mile Island) make a risk easier to imagine, and thus more risky (as defined above)
Dread	Some illnesses are more dreaded than others. Compare cancer with, say, heart failure. The long latency of most cancers and the surreptitiousness of most carcinogens add to the dread
Diffusion in time and space	If hazard A kills 50 anonymous people a year across the country, and hazard B has one chance in 10 of wiping out an entire town of 50 000 people sometime in the next century, then a risk assessment tells us that the two have the same expected annual mortality of 50 deaths per year. However, an 'outrage assessment' would show that hazard A is probably acceptable but hazard B is certainly unacceptable

Reproduced from Sandman (1987).

Practitioners can deal with public communication of risk more effectively by **listening** to the concerns expressed by the public, including those of pressure groups, in order to:

- Take into account the factors that influence risk perception
- Understand the strength of feeling and the points of view
- Use appropriate media and language to communicate relevant information in a meaningful form.

2F.3 CLIMATE CHANGE

Effects of climate change and global warming

The main causes of human-induced climate change are increased levels of greenhouse gases, especially CO_2 and methane. Most of the heat coming from the sun is reflected off the earth back into space. But so-called 'greenhouse gases' trap the warmth. As the concentration of greenhouse gases increases, so the planet's temperature is predicted to rise. Increased atmospheric levels of CO_2 result from deforestation (trees would normally convert CO_2 to oxygen through photosynthesis) and from increased combustion of fossil fuels.

GLOBAL WARMING

Average global temperatures have risen over the past century due to both natural and human factors. Climatologists predict that extreme weather (e.g. floods, hurricanes and droughts) will become more frequent because of climate change. In the UK, the 1990s were the warmest decade in the last 100 years, and by 2080, the average annual temperature is predicted to rise by between 2 and 3.5°C.

EFFECTS OF CLIMATE CHANGE ON HEALTH

Effects may be direct or indirect: see Box 2F.3.1.

Box 2F.3.1

Direct effects	Indirect effects
• Extreme **heat waves** have led to excess mortality, particularly among old people or those in vulnerable sectors of society • Changing epidemiology of **infectious diseases,** e.g. spread of malaria as animals/insects migrate to different geographical areas • Fewer deaths related to **hypothermia** (although paradoxically the temperature in the British Isles may fall due to the effects on the Gulf Stream)	• More **floods**, promoting the spread of water-borne diseases • More asthma and respiratory illnesses as a result of increased **atmospheric pollution** • Increasing numbers of **refugees** as some areas are submerged under rising sea levels • **Droughts** and changing food patterns lead to starvation, famine and increasing numbers of refugees

MONITORING AND CONTROL

Efforts to restrict the burning of fossil fuels and stop deforestation are among many measures to limit climate change contained in the 1997 **Kyoto Protocol**. The protocol is a legally binding agreement for countries to cut emissions of the six main greenhouse gases. It came into force in 2005, but while 141 countries have ratified the protocol, two of the biggest producers of greenhouse gases – the USA and Australia – rejected the treaty. Many countries that signed up to the treaty are experiencing difficulties in meeting its targets.

2F.4 SUSTAINABILITY

Principles of sustainability

Environmental sustainability encompasses the concern to balance the needs of the present generation with those of the future. It involves long-term considerations about resource use, set in the context of the current high rate of resource use and production of pollutants.

UK The **Sustainable Development Commission** adopts a broad view of sustainability, considering not just environmental, but also social and economic developments. It outlines five principles of sustainable development: see Table 2F.4.1

Table 2F.4.1 Principles of sustainable development

Environmental limits	Respecting the limits of the planet's environment, resources and biodiversity to improve the environment and ensure that all the natural resources needed for life remain for future generations
Healthy and just society	Meeting the diverse needs of all people in existing and future communities; promoting wellbeing, social cohesion and inclusion; and creating equal opportunity for all
Good governance	Actively promoting effective, participative systems of governance in all levels of society – engaging people's creativity, energy and diversity
Responsible use of science	Ensuring that policy is developed and implemented on the basis of strong scientific evidence while taking into account scientific uncertainty through the precautionary principle (see Section 2F.7) as well as public attitudes and values
Sustainable economy	Building a strong, stable and sustainable economy that provides prosperity and opportunities for all, an economy in which environmental and social costs fall on those who impose them (polluter pays, described in Section 2F.7) and in which incentives are in place to promote efficient resource use

2F.5 HOUSING

Health problems associated with poor housing and home conditions, inadequate water supplies and sanitation

Poor housing and home conditions can be associated with health problems. Specific features of poor housing and home conditions are summarised in Table 2F.5.1.

Table 2F.5.1 Housing conditions affecting health

Pollutants	Examples include asbestos, carbon monoxide, radon, lead and volatile organic chemicals
Damp	Causes moulds and spores
Temperature	Cold and heat extremes
Design	Poor housing design or layout
Noise	Intrusive or excessively loud noises or noise made at antisocial hours
Social and behavioural environment	Overcrowding, sleep deprivation
Housing allocation	Homelessness, temporary accommodation, housing tenure, housing investment

Eng The amenities available to a neighbourhood (such as access to transport, shops and open spaces) have a profound impact on health. These are mediated though their effects on safety, social cohesion and crime. In England, government policy explicitly recognises the benefits of open spaces for enabling local people to be active and maintain social links. Councils have a duty to consult locally on plans for maintaining open spaces, sports facilities and buildings.

HEALTH EFFECTS OF POOR HOUSING

The relationship between housing and health is complex: individual exposures may not relate directly to particular morbidities. People living in poor housing conditions typically also experience other forms of deprivation (poor education, unemployment, ill health, social isolation, etc.), and this makes it difficult to assess, isolate or modify the overall health effects. In general, housing has both physical and mental effects on residents' health.

PHYSICAL HEALTH PROBLEMS

Poor housing is associated with asthma, skin allergies and respiratory diseases. It is also linked to physical accidents and injuries. Exposure to specific toxins has particular effects, e.g. radon is associated with cancer, CO causes asphyxiation and asbestos causes asbestosis and malignancy.

MENTAL HEALTH PROBLEMS

Poor housing can lead to depression, isolation, anxiety or aggression. Noise-related stress, from poor sound insulation, is associated with lack of sleep, mental distress and depression: for an example see Box 2F.5.1.

Box 2F.5.1

Example: WHO LARES Survey

The WHO's LARES (Large Analysis and Review of European housing and health Status) was a survey across eight countries involving 8500 people, which studied links between housing conditions and physical and mental ill health. It found, for example, that people were more depressed and anxious when their housing did not protect against noise, or they were subjected to overcrowding or to social isolation. Dampness and visible mould growth were related to asthma, nasal allergies and eczema.

Reproduced from Bonnefoy et al (2004).

HEALTH PROBLEMS ASSOCIATED WITH INADEQUATE WATER SUPPLIES AND SANITATION

Inadequate water supplies remain a major problem for around 15–20% of the world's population, largely in Asia and sub-Saharan Africa. Inadequate water supplies and sanitation problems can result from natural disasters (flood, earthquake or drought) or from pollution. Human interventions to reduce water shortages can themselves exacerbate health problems (e.g. malaria associated with irrigation channels).

Inadequate water supplies can lead to a range of **infectious diseases**. These may be due either to microbial agents in the water or to water-related vectors. WHO (2007b) outlines the most significant diseases associated with inadequate water supply and sanitation (see Table 2F.5.2).

Table 2F.5.2 Infectious diseases associated with water

Diarrhoeal illness	88% of diarrhoeal illness is due to unsafe water supply, inadequate sanitation and poor hygiene • Improved water **supply** reduces morbidity by 6–25% • Improved **sanitation** reduces morbidity by 32%
Malaria	There are 396 million malaria episodes per year worldwide, with most of the disease burden in sub-Saharan Africa. Intensified irrigation, dams and other water-related projects contribute to disease burden Better management of water resources reduces transmission of malaria and other vector-borne diseases
Chemical pollutants	Pollutants can also cause disease, e.g.: • **Arsenic:** contamination of ground water is found in many countries, including Argentina, Bangladesh and the USA. Arsenic leads, among other things, to skin lesions • **Fluorosis:** naturally elevated fluoride levels in drinking water usually occur due to the geology of the area. It can lead to skeletal and dental fluorosis • **Aluminium:** aluminium sulphate contamination of the Camelford reservoir in Cornwall in the 1980s may have caused acute symptoms, and possibly long-term neural damage

The UN's **Millennium Development Goals** include aims to reduce water shortages and improve sanitation. Reducing the proportion of people without access to safe drinking water and basic sanitation improves wellbeing directly. It will also have an indirect effect on goals associated with reducing morbidity and mortality, by affecting the risk of illnesses associated with water supply, e.g. diarrhoea.

2F.6 MONITORING AND CONTROL OF ENVIRONMENTAL HAZARDS

Methods for monitoring and control of environmental hazards (including food and water safety, atmospheric pollution, and other toxic hazards, noise, and ionising and electromagnetic radiation)

See Sections 2F.1 and 2F.3.

2F.7 USE OF LEGISLATION IN ENVIRONMENTAL CONTROL

The control of environmental pollution and waste production by businesses often requires legislation to be effective. This is because there is commonly a perceived or real conflict between economic priorities and the environment. The issue is particularly pertinent for small businesses, and also in developing countries where the information and expertise to implement environmentally friendly measures may be lacking.

Many environmental hazards are not limited to geographical boundaries. For example, acid rain caused by UK power plants in the 1970s and 1980s mainly fell in continental Europe. Therefore, legislation and strategies to tackle pollution and waste production should be set at an international level. European legislation may be **primary** (e.g. treaties and directives) or **secondary** (e.g. regulations and statutory instruments).

EU EUROPEAN UNION LEGISLATION

EU European environmental legislation is based on two guiding principles: 'polluter pays' and the 'precautionary principle': Box 2F.7.1.

Box 2F.7.1

Polluter pays	This is the principle that the party responsible for creating pollution should provide the investment needed to meet higher environmental standards or to create a system to recover, recycle or dispose of products after use. The payment may also be a tax on business or consumers for using an environmentally unfriendly product, such as some types of packaging
Precautionary principle	Credible but unproved health hazards should be treated as real threats until shown otherwise. Decisions affecting the environment should not be delayed while scientific data are collected. Instead, policy-makers should err on the side of protecting the environment

European legislation aims to restrict levels of pollution, and to ensure the safe disposal of waste through measures outlined in Table 2F.7.1.

Table 2F.7.1 Measures in European legislation to restrict pollution

Permits or licences	Issued to those using natural resources, disposing of waste or producing pollution (e.g. sewage companies, water companies and manufacturing industries). Conditions attached to the permit typically set limits on the amount of pollution that may be produced, or stipulate the use of specific procedures and **best available techniques** (BATs)
Taxes or levies	By reflecting the environmental costs of production, these aim to discourage production of emissions. An example is the climate change levy
Trading schemes	Companies participate in a scheme that is committed to reducing emissions. Firms may 'sell' or 'buy' excess emission allowances and thus compete on the basis of an overall reduction in emissions
Product labelling schemes	*'Eco-friendly'* or energy efficiency ratings are displayed on household and office appliances

Compliance with the terms of permits and regulations can be enforced through:

- Inspections, warning letters or formal cautions
- Enforcement or prohibition notices
- Suspension, revocation or modification of permits
- Fines and imprisonment.

REGULATORY AGENCIES

UK In the UK, the **Environment Agency** regulates large and complex industrial processes, while local authorities regulate smaller-scale industries. Other relevant regulatory agencies include the **Health and Safety Executive, Trading Standards** and the **Food Standards Agency**.

Environmental health officers from the local authority enforce the regulations on the Clean Air Act, and prevent the sale of contaminated foods.

EFFECTIVENESS OF LEGISLATION

Compliance with environmental legislation is variable. There are many reasons for this, including:

- Business are not always aware of legislative requirements
- Monitoring in some countries is more rigorous than others

- Penalties for non-compliance are comparatively weak, e.g. compared with health and safety offences
- Compliance is not always perceived as central to a firm's survival.

2F.8 HEALTH AND SAFETY AT WORK

Appreciation of factors affecting health and safety at work (including the control of substances hazardous to health)

Factors affecting health at work may be associated with the work **environment**, occupational **equipment** or other **employees**. They can be either beneficial or harmful, and be mediated through either direct or indirect effects (see Section 2H.6).

Workplace exposures may be considered as physical, substance related and psychological.

PHYSICAL EFFECTS

Back disorders are the commonest cause of workplace ill health, and musculoskeletal problems in general are a major cause of morbidity. They are particularly associated with:

- Sedentary occupations
- Lifting and manual handling of heavy objects
- Driving for long periods
- Repetitive tasks.

SUBSTANCES

UK **The Control of Substances Hazardous to Health (COSHH) (2002) Regulations** cover the following types of hazardous substance:

- **Substances used** directly in work activities (e.g. adhesives, paints, cleaning agents)
- **Substances generated** during work activities (e.g. fumes from soldering and welding)
- **Naturally occurring** substances (e.g. grain dust)
- **Biological agents** such as bacteria and other microorganisms (e.g. leptospirosis in sewage treatment workers).

The effects of these substances may be atopic, infectious or carcinogenic, and they may manifest in a range of different timeframes: see Box 2F.8.2.

Box 2F.8.2

Acute effects	Losing consciousness as a result of being overcome by toxic fumes
Early effects	Skin irritation or dermatitis as a result of skin contact (particularly common among beauticians and hairdressers)
Late illnesses	Cancer may develop many years after the exposure to the carcinogen (e.g. asbestos)

PSYCHOLOGICAL FACTORS AFFECTING HEALTH

Work is the single biggest cause of stress in Britain. Employment-related factors that cause stress and mental ill health include:

- Bullying
- Harassment
- Discrimination
- Lack of control
- Lack of reward
- Long working hours.

People at particularly high risk are:

- Temporary workers
- Women (a higher proportion are in lower status jobs)
- Those working illegally (e.g. asylum seekers)
- Lone workers
- Staff in small businesses (small businesses can have the most difficulty complying with workplace regulations)
- Pregnant women.

Depression and anxiety are the most common causes of long-term incapacity. Stress is also linked to premature death from coronary heart disease. Workplace exposures can include:

- Stress
- Inadequate team support
- Inadequate reward
- Lack of control over roles and responsibilities in the workplace.

In certain circumstances, employees can be a risk to the health of other people, e.g.:

- **Health-care workers** with infections. Safeguards include restriction of exposure-prone procedures for those with blood-borne infections, e.g. hepatitis B
- **Food handlers**, carers of at-risk groups with gastrointestinal illnesses. Safeguards include removing from work during illness and until the infection has cleared
- **Drivers**, e.g. public transport, goods: safeguards include strict regulations and regular monitoring over substance (drugs, alcohol) intake.

HEALTH AND SAFETY AT WORK

Effective management of health and safety issues at work not only maximises the benefits to health but also makes economic sense for companies and society. This is because it reduces absenteeism, increases productivity and avoids compensation claims.

UK Key agencies involved in occupational health are listed in Table 2F.8.1.

Table 2F.8.1 UK agencies involved with occupational health

Agency	Purpose
Health and Safety Executive	• Sets regulations • Monitors workplaces • Investigates incidents • Enforces workplace regulations
Trade unions	• Campaign for workers' rights to healthy working conditions • Survey employees' working conditions • Provide advice on legal issues concerning health and safety • Represent employees with employers on health and safety issues
Local authorities	• Local inspection, enforcement of regulations
Employers	• Implement regulations • Provide safe working conditions
Occupational health departments	• Undertake pre-employment health checks • Advise employers on workplace hazards and their control • Advise employees on health and safety issues • Monitor sickness absence • Offer immunisations if appropriate, e.g. in health and social care jobs

OCCUPATIONAL HEALTH REGULATIONS

UK Pertinent regulations include:

- Control of Substances Hazardous to Health Regulations (COSHH) (2002)
- Reporting of Injuries, Diseases and Dangerous Occurrences Regulations 1 (RIDDOR) (1995): required for all serious work-related incidents
- WHO occupational limits for pollutants
- European Working Time Directive (among other provisions, this restricts the working week to 48 h).

SAFETY AT WORK AUDITS

Procedures to monitor and control health and safety factors at work can be considered in several sequential steps: see Box 2F.8.3.

Box 2F.8.3

Assessment	Assessment of hazards in the environment
Identification	Identification of those at most risk (e.g. through pre-employment health checks)
Surveillance	Surveillance of hazards, incidents and health-related morbidity (through systems of incident reporting, feedback and analysis)
Mitigation	Action to reduce exposure to hazards (through training, provision of protective equipment to exposure and emergency planning)
Monitoring	Monitoring of the effect of interventions to reduce effects of hazards (e.g. through incidence, sickness and absence records)

2F.9 OCCUPATION AND HEALTH

The effects of occupation on health vary according to the:

- Type of occupation
- Personal risk factors
- Levels of social support.

In general, there are substantial health benefits to being in work, but certain occupations can expose employees to particular risks (see Section 2F.8).

UNEMPLOYMENT

Being out of work is associated with physical, mental and social effects. These are related partly to the length of time spent unemployed. Although being out of work may cause morbidity, the causal relationship also works the other way around: those who are physically or mentally unwell may be more likely to leave jobs or have difficulty working. Studies of factory closures indicate that ill health and health-care use are also associated with job insecurity, and particularly with the **anticipation** of job losses.

There are consequences from unemployment on the **individual**, on **families** and on **society** at large: see Box 2F.9.1.

Box 2F.9.1

Individual	Suicide rates among the unemployed increase within a year of job loss Cardiovascular mortality rises over 2 or 3 years, and continues for the next 10–15 years
Family	Spouses of people out of work can also experience poor health. For those whose partner has been forced to give up work due to illness, they may be burdened with caring responsibilities. In families where individuals are seeking work, the effects of poverty appear most severe in the short term. In the longer term, families may adapt to unemployment
Society	Increased unemployment places a greater burden on society through fewer people making tax contributions, and due to the increased social security and health service needs of those without work

2F.10 TRANSPORT AND THE ENVIRONMENT

Transport policies and health impact assessment for environmental pollution

Transport policies need to acknowledge the potential **benefits** from increased transport and accessibility. These improve the prospects of employment, availability of goods and greater choice of social activities. It is also important to recognise the health benefits associated with walking or cycling, as well as the risks associated with other forms of transport – particularly via road and air, for example; the effects of transport on the health of Londoners has been the subject of a health impact assessment (see Box 2F.10.1). The main problems associated with transport are:

- Air pollution and global warming
- Noise pollution
- Accidents
- Effects on social cohesion (e.g. from a busy dual carriage-way bisecting a community).

Road transport is one of the greatest contributors to air pollution, particularly in towns and cities (see Section 2F.1). However, while emissions from most sources are decreasing, the impact of air travel is increasing sharply. According to the European Commission, EU emissions from international aviation increased by almost 70% between 1990 and 2002.

HEALTH IMPACT ASSESSMENTS

See also Section 1C.17.

A health impact assessment is an important tool for informing policy-makers about the health effects of potential transport policy decisions. However, there are many challenges in conducting this type of research. In general, the evidence regarding the quantitative effects on health is variable. For example, while the mortality data from traffic accidents are uncontested, there are **methodological disputes** about how to measure the effects of noise disturbance, the health benefits of being physically active, transport effects on respiratory illness, etc.

SUSTAINABLE TRANSPORT

Sustainable transport policies focus on the following elements.

ENVIRONMENTALLY FRIENDLY TRANSPORTATION

Alternatives to cars can be encouraged by sustainable transport policies: see Box 2F.10.2.

Eng Box 2F.10.1

Example: *On the move*: Health Impact Assessment of Transport in London (2000)

In London, traffic is responsible for the production of 99% of carbon monoxide, 76% of nitrogen oxides and 90% of hydrocarbons. As part of London's health strategy, the NHS commissioned a health impact assessment of transport in London.

Scope: the assessment focused on:

- Traffic accidents
- Air pollution
- Noise
- Health benefits from physical activity
- Community severance, mental health and inequality effects

The positive effects of transport from increased access to goods and services and ability to move around London were outside the scope of the assessment.

Methods: researchers assessed and quantified health effects by:

- Collating and interpreting routine data
- Reviewing evidence from published and grey literature
- Consulting stakeholders

Figures of accidents indicated that road traffic accidents led to 226 fatalities in the previous year. In comparison, the authors estimated that pollution would lead to 380 premature deaths and 350 hospital admissions due to respiratory disease.

Findings: this health impact assessment indicated that pollution is as important as (and perhaps more important than) accidents with respect to the health of Londoners.

Reproduced from Watkiss et al (2000).

Box 2F.10.2

Transport option	Promotion
Walking	Pedestrian zones
Bicycles	Bicycle lanes
Public transport	Subsidised public transport systems

TAXATION

Taxation can be used to encourage industry and consumers to adopt more fuel-efficient vehicles and thereby reduce vehicle emissions.

REDUCING CAR USAGE

This can be achieved through:

- Levies or taxes for road use (e.g. London Congestion Charge)
- Parking restrictions and charges
- Subsidising the cost of public transport
- Encouraging car sharing through introducing share-only lanes and high-volume-occupancy lanes
- Reducing the transport of freight on the roads and increasing rail transport.

REDUCING THE ENVIRONMENTAL IMPACT OF TRANSPORT

Carbon-offsetting initiatives seek to compensate the environment for emissions of carbon dioxide. Processes that emit carbon dioxide are matched with a proportionate investment in projects that either reduce the emission of carbon dioxide or remove an equivalent amount of the gas from the air. For example, air passengers can make a donation to a non-governmental organisation that will fund projects designed to reduce emissions.

Wilkinson (2006) describes how researchers are exploring the option of carbon capture as an alternative approach to reducing the impact of carbon dioxide production. The damage to the environment results from carbon dioxide in the atmosphere. Carbon capture relies on finding locations to store carbon dioxide so that it does not reach the atmosphere. Depleted oil and gas fields could be ideal carbon capture locations: they have increasing storage space, and the insertion of carbon dioxide could make the continued extraction of gas and oil more straightforward. Note that carbon capture would not result in a reduction in the use of hydrocarbons, or in the production of carbon dioxide. It may be best regarded as a relatively inexpensive short-term solution the value of which lies in buying time to develop alternative fuel sources. An example is air transport (see Box 2F.10.3).

EU Box 2F.10.3

Example: Air transport: an international perspective to transport policy

In Europe in the last 20 years, air transport policies have focused on expanding the air travel market, rather than on reducing or containing its environmental impact. One of these policies is the exemption of aircraft fuel for international flights from all taxes. While individual countries in the EU have the option to charge fuel tax for domestic flights, only the Netherlands currently does so.

As a result, air transport has increased rapidly. Although planes are becoming more fuel efficient, this is more than offset by air traffic. The European Commission recommends that the aviation industry should become part of the EU **Emissions Trading Scheme**. Under this scheme, the European Commission places an overall 'cap' on emissions per year. Within these limits, companies are given an emissions allowance. If they are in danger of exceeding their allowance, they have two options: either to reduce their emissions (e.g. by investing in new technology, changing production) or to 'buy' the allowances of their competitors on the open market.

Adapted from www.europa.eu.int/comm/environment/climat/pdf/ia_aviation.pdf, www.europa.eu.int/comm/environment/climat/aviation_en.htm.

2F.11 CHEMICAL INCIDENT MANAGEMENT

Damage from chemical incidents can arise from **acute events** (such as a chemical spill or an explosion) or due to **chronic processes** (such as leeching into the soil or water supply). The detection and management are different between the two, but some of the principles and possible consequences are similar. Examples of chemical incident management are described in Boxes 2F.11.1 and 2F.11.2 on pages 233 and 234.

The most common environmental effects of a chemical incident are:

- Contaminated **land** (particularly from leaks, slow accumulation of chemicals)
- Contaminated **water** supplies
- **Air** pollution, e.g. from industrial fires.

UK The Health Protection Agency provides several checklists that outline the management of different types of chemical incident. While the details differ depending on the nature of the incident, the general steps involved are outlined below (the order of the steps will depend on the nature of the incident).

INCIDENT ASSESSMENT

This involves confirming which chemical or chemicals are involved, the extent of contamination and the potential for further contamination. The pathway through which chemicals are leaking or leeching must be determined. Samples of soil, water and air from the local environment are taken.

HEALTH ASSESSMENT

Appropriate epidemiological study designs to assess health effects are generally either a **case–control** study (requires a case definition and for cases to have been identified) or a **cohort study** (if there is a defined population that was potentially affected).

Both study designs could involve biological sampling of those exposed and with symptoms (e.g. blood or urine analysis), and a survey of all exposed people to quantify physical and psychological symptoms. After an incident it may also be appropriate to consider whether to instigate long-term epidemiological surveillance, or to follow up a selected cohort (measuring the nature and extent of any long-term health effects).

RISK MINIMISATION

After an incident, these steps usually take the form of making arrangements for moving people or materials and are designed to:

- Provide **advice** to exposed and potentially exposed populations
- Ensure appropriate **shelter**: this may involve simply advising people to stay inside but could also require arranging alternative accommodation if homes or businesses are damaged or unsafe
- Manage **evacuation**: as well as identifying a suitable place for people to be evacuated to, it is also important to consider how to inform, prepare and move the public
- **Clear up** contamination where feasible.

COMMUNICATION AND COORDINATION

In a chemical incident, a range of agencies will be involved in minimising the damage and clearing up the incident. It is therefore important for all organisations involved to understand their respective roles and responsibilities. Clear lines of communication are required in order to ensure that the most current information is available to all – particularly with regard to the potential hazards faced. Different agencies will be involved depending on the scale and nature of the incident, but Tables 2F.11.1 and 2F.11.2 indicate agencies likely to be involved in the UK.

Aus There is a **National Disaster Organisation**, but state police control most incidents. The police work in association with fire, health and environmental protection authorities. Other organisations (e.g. water utilities) are involved as appropriate. Officers from all of these agencies have well-defined, statutory duties in such situations.

PUBLIC INFORMATION

Those affected by the incident need to be kept well informed. The general public may have been directly or indirectly exposed to the effects of the incident, and many will be concerned that they could be affected. Useful information includes:

- Actions to take to minimise personal risk
- Minimising risk to others
- Where to get help and advice.

Table 2F.11.1 UK national agencies involved in reducing the impact of environmental incidents

Eng National agencies	Wal National agencies	Scot National agencies	UK National agencies
Environment Agency DEFRA	Environment Agency DEFRA	Scottish Environment Protection Agency (SEPA) Scottish Executive Environment and Rural Affairs Department (SEERAD)	Food Standards Agency (if potential for contaminated food stuffs) Health and Safety Executive National newspapers, TV and radio Internet Specialist press
Health Protection Agency	National Public Health Service for Wales	Health Protection Scotland	

Table 2F.11.2 UK local agencies involved in reducing the impact of environmental incidents

UK Local agencies
• Emergency services (police, fire, coastguard) • Local authority • Environmental health department • Health and safety departments • Pollution control • Health protection units, departments of public health • Water companies (if potential for contaminated water supplies) and local Environment Agency office • Health and Safety Executive (local branch) • Eng NHS: A&E, acute hospital trusts; PCTs; NHS Direct • Wal NHS: A&E, local health boards; NHS Direct • Scot NHS: A&E, local health boards; NHS 24 • Media: local newspapers, regional TV and radio

Useful media for inform people include:

- Local radio
- Helplines
- Special mailings
- Websites
- Press releases to local, specialist and national media.

LEARNING

Once the process has been completed, all the agencies involved should identify lessons learnt through the incident. Based on those experiences, plans should be modified to prepare for managing future incidents and to prevent similar incidents from occurring.

Boxes 2F.11.1 and 2F.11.2 give examples from the UK and New Zealand.

UK Box 2F.11.1

Example: Management of the Buncefield Oil Depot Fire

The incident: On 11 December 2005, Buncefield Oil Depot near Hemel Hempstead (south-east England) was the site of 'possibly the largest explosion in peacetime Europe'. A fog of petrol fumes caught alight, leading to explosions for the next 24–48 h. Over the next few days, London and significant parts of south-east England were covered by the smoke plume, which spread to parts of northern France and Portugal. The site housed hundreds of businesses, employing several thousand staff. Nearby, there was a traveller community and a substantial residential community. Over 2000 people were evacuated from their homes and sections of the nearby motorway were closed to prevent exposure to the fumes produced by the explosions.

Agencies: the police led the response to this incident, with input from a range of agencies, including Dacorum & Watford and Three Rivers Primary Care Trusts, the Health Protection Agency, the Environment Agency, the Health and Safety Executive and the fire service.

Environmental effects: the fire actually generated few pollutants – its high temperature meant that organic chemicals in the fuel were destroyed. The Met Office provided information on the plume direction and spread using observations, satellite images and computer modelling. This indicated that smoke from the fire rose high in the atmosphere.

Health and social effects: data collected from NHS Direct and A&E in the hours and days after the incident indicated that Buncefield was not a major incident for health; there were no deaths and few injuries.

A survey was also used to find out residents' experience of the fire, acute exposure and longer-term health effects.

Longer term: risks under investigation include:

- The risk of contamination of water supplies from foam
- Longer-term health effects – both physical and psychological
- Employment and housing in the Buncefield area

Adapted from HPA conference, 7 March news.bbc.co.uk, 12 December 2005: Massive blaze rages at fuel depot, The public health impact of the Buncefield oil fire, 2006: www.hpa.org.uk/publications/2006/buncefield/buncefield.pdf, DEFRA, Initial review of air quality aspects of the Buncefield oil depot explosion: www.defra.gov.uk/environment/airquality/buncefield/index.htm.

NZ Box 2F.11.2

Example: New Zealand management of hazardous substances (Hazsub) events

Hazsub are mainly accidental releases from transportation-related events or from fixed facilities such as factories and storage plants. Rarely, they may be intentional releases from criminal or terrorist activity or from natural hazards such as a volcanic eruption. Responding to such incidents involves the coordinated actions of multiple agencies.

District health boards (DHBs) (covering hospitals, public health services and contracted primary care providers)	Plan the response to such incidents, with Hazsub events included in the DHB Major Incident and Emergency Plan. Public health services have a key role in preparedness planning
Fire service	Scene management, containment of released hazardous substance and decontamination of individuals at the scene
Ambulance service	Assessment, triage, initial treatment and transport of people injured during a Hazsub release
Police	Securing and managing the scene and investigating the incident if the release results from criminal or terrorist act, potentially with support from the New Zealand Defence Force
Hazardous substance technical liaison committees	Advise and support the fire service when dealing with hazardous chemical incidents
National Poisons Centre	Maintains a database of chemicals, health effects and recommended treatment
National Radiation Laboratory	Expert advice and technical support to the emergency and health services if ionising radiation involved
Civil defence and emergency management groups	Consortia of local authorities, emergency services, major utilities and others to ensure that emergency management principles are applied at the local level
Ministry of Health	Produces the National Health Emergency Plan (NHEP) Provides advice to DHBs, public health services and other agencies to prepare their own action plans
Ministry of Civil Defence and Emergency Management (MCDEM)	Responsible for the administration of the Civil Defence Emergency Management Act 2002. Uses national standard: the Coordinated Incident Management System (CIMS)

2G
Communicable Diseases

Infectious diseases represent the largest cause of childhood and adolescent deaths worldwide. They account for over 13 million deaths a year and are the cause of over half of all deaths in developing countries. In developed countries infectious diseases are of importance for the following reasons:

- Re-emergence of old scourges (e.g. TB)
- Novel infections (e.g. SARS)
- Threat of pandemic influenza
- Hospital-acquired infections (e.g. MRSA)
- Burden of long-term conditions (e.g. HIV/AIDS)
- Recognition that certain cancers are caused by viruses (e.g. cervical cancer).

2G.1 STATES IN THE DEVELOPMENT OF INFECTIOUS DISEASES

Definitions (incubation, communicability, latent period, susceptibility, immunity and herd immunity)

See Table 2G.1.1.

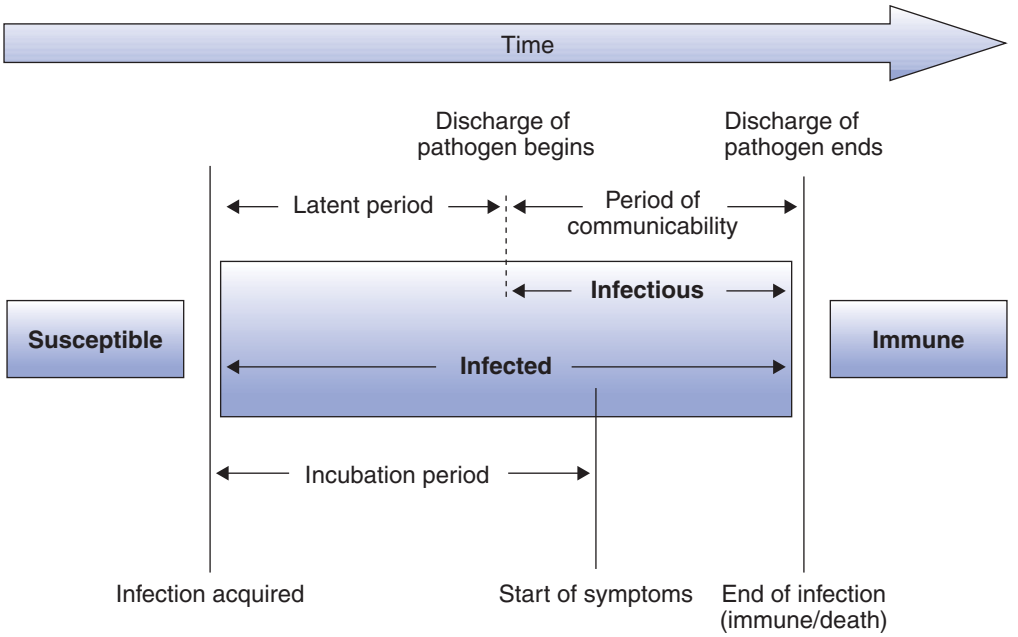


Figure 2G.1.1 Temporal measures in infectious disease

Table 2G.1.1 States in the development of infectious diseases

Susceptibility	A susceptible person has insufficient resistance against the pathogen to avoid infection
Incubation period	Also known as the 'subclinical period', this is the time between infection and the onset of clinical symptoms. Its duration may be affected by the infecting dose Establishing the incubation period enables the following to be determined: • When infection occurred • Who might be a contact and how long they should be quarantined for
Latent period	Time between initial infection and the start of infectiousness
Infectious period	Time during which the infected person is capable of transmitting the infective agent (often begins before the onset of symptoms)
Degree of infectivity	This will usually vary during the infectious period. It determines: • The extent to which the disease will be transmitted in the population • For how long patients should be isolated or quarantined
Index case	The first case of the disease to come to the attention of the investigator
Primary case	Case that occurred due to infection from the same source as the index case
Secondary case	Case that occurred due to infection from a primary case
Carrier	Person who becomes infected and infectious but not ill
Serial interval	Time between the onset of symptoms in a primary case to the onset of symptoms in a secondary case

Secondary attack rate	Probability of disease among known (or presumed) susceptible people following contact with a known **primary** case
Basic reproductive number (R_0)	Average number of new infectious cases in a completely susceptible population produced by a single case during its entire period of infectiousness

IMMUNITY

Immunity is the state of having sufficient biological defences to avoid infection. It can be considered as passive or active: see Table 2G.1.2.

Table 2G.1.2 Active and passive immunity

	Definition	Natural acquisition	Artificial acquisition	Duration
Passive immunity	Passive immunity is acquired by transfer of antibodies (or T lymphocytes) from another individual	Transplacental transfer from the mother	**Inoculation** of specific protective antibodies (from immunised animals, or from human convalescent serum)	**Short:** days to months
Active immunity	Active immunity is acquired through an encounter with an antigen	**Infection** (with or without clinical manifestations)	**Inoculation** (of either the agent itself in killed, modified or variant form, or else of fractions or products of the agent)	**Long:** usually several years

HERD IMMUNITY

If most people in a community are immunised against an infection, the spread of that infection is significantly reduced and even unvaccinated people are at much lower risk of catching the illness. For childhood infections such as measles, about 95% of children must be vaccinated in order to achieve this level of protection.

2G.2 SURVEILLANCE – NATIONAL AND INTERNATIONAL: ITS EVALUATION AND USE

Surveillance is the ongoing systematic collection, collation, analysis and interpretation of data, and the dissemination of information (to those who need to know) in order that action may be taken: in summary it provides information for action. See Table 2G.2.1.

Table 2G.2.1 Principles of surveillance

Case definition	Using clinical or microbiological criteria
Cases identified through variety of sources	Clinicians or laboratories Primary, secondary or tertiary care
Data sources	Electronic database, e.g. *CoSurv* Special forms, e.g. legionella questionnaire
Systematic collection of data for cases that satisfy the case definition	
Analysis of data and summary statistics	
Feedback to data providers and distribution to those who require it for action	Disseminate information – newsletter/bulletin/articles/ reports (e.g. CDR weekly/Health Protection Report)
Continuing surveillance for evaluation of interventions	

NATIONAL SURVEILLANCE

UK The Health Protection Agency's (HPA) Centre for Infections (CfI) is responsible for coordinating communicable disease surveillance across England and Wales. In Scotland, this role is performed by Health Protection Scotland (HPS). Certain key infectious diseases are kept under constant surveillance in order to:

- Detect trends
- Evaluate prevention and control measures
- Alert appropriate professionals and organisations to infectious disease threats.

See Table 2G.2.2.

Table 2G.2.2 The main sources of information for national surveillance

Notifiable diseases	By UK law, certain diseases are deemed 'notifiable', meaning that doctors have a statutory duty to report them. These reports should be made on the basis of clinical suspicion, rather than waiting for laboratory confirmation The HPA produces a **NOIDS Weekly Report** (**N**otifications **o**f **I**nfectious **D**iseases) and **WARNER Reports** (**W**eekly **A**nalysis **R**eport of **N**otifications above **E**xpected **R**ates) See Section 2G.8 for a list of the UK notifiable diseases
Reports/special forms/returns	Laboratory Case report – rare disease Incident report Royal College of General Practitioners weekly returns (e.g. for influenza, whooping cough, asthma) Genitourinary/sexual health clinic disease statistics (called KC60 returns in England and Wales) Vaccine coverage: COVER (Cover Of Vaccination Evaluated Rapidly); ISD (Information Services Data, Scotland) statistics Death registration Hospital episode statistics NHS Direct/NHS 24 Surveillance systems: • Sentinel surveillance, e.g. HPA/Nottingham University Primary Care Surveillance Scheme (Q Flu, Q Research) • Enhanced TB Surveillance (TB register in London)/Enhanced Surveillance of Mycobacterium Infection (ESMI) • Enhanced Meningococcal Surveillance
Surveys	Survey of Prevalent HIV Infections Diagnosed (SOPHID) Seroprevalence surveys Unlinked anonymous surveys, e.g. Blood-Borne Viruses (BBV) in injecting drug users, BBV in antenatal women

See Section 1C.3 for details of disease surveillance specific to the four UK countries.

Enhanced surveillance schemes use multiple data sources to gather additional detailed information on worrisome diseases (e.g. meningococcal disease, TB). The additional information is used to:

- Advise policy-makers whether action is required to address an emergent threat such as pandemic influenza
- Detect new or imported infections, such as SARS.

Ire In Ireland, the Health Protection Surveillance Centre (HPSC) is a unit within the Population Health Directorate of the Health Service Executive. The HPSC operates the Computerised Infectious Disease Reporting (CIDR) system to manage the surveillance and control of infectious diseases in Ireland and to monitor antimicrobial resistance. Health personnel electronically return notification, outbreak and enhanced surveillance forms. The data provide the basis for the weekly infectious diseases and outbreak reports published by the HPSC.

An example of enhanced surveillance (for pandemic influenza) is shown in Box 2G.2.1.

Box 2G.2.1

Example: Enhanced surveillance for pandemic influenza

The WHO runs a network of over 100 National Influenza Centres that monitor influenza incidence and analyse virus isolates. Any unusual activity and/or isolates are reported immediately through the WHO Global Influenza Programme. The current early warning system has potential weaknesses in that some countries lack the ability to test for H5N1, cases tend to occur in rural areas and other illnesses with similar symptoms are common.

The WHO plans to provide infrastructure to improve case detection, actively search for human cases where an outbreak in animals is identified, support epidemiological investigation and promote collaboration between Asian countries.

Interpandemic period	Phase 1	No new influenza virus subtypes have been detected in humans. An influenza virus subtype that has caused human infection may be present in animals. If present in animals, the risk of human infection or disease is considered to be low
	Phase 2	No new influenza virus subtypes have been detected in humans. However, a circulating animal influenza virus subtype poses a substantial risk of human disease
Pandemic alert period	Phase 3	Human infection(s) with a new subtype but no human-to-human spread, or at most rare instances of spread to a close contact
	Phase 4	Small cluster(s) with limited human-to-human transmission but spread is highly localised, suggesting that the virus is not well adapted to humans
	Phase 5	Larger cluster(s) but human-to-human spread still localised, suggesting that the virus is becoming increasingly better adapted to humans but may not yet be fully transmissible (substantial pandemic risk)
Pandemic period	Phase 6	Pandemic: increased and sustained transmission in general population

Adapted from www.who.int/csr/disease/influenza/pandemic/en/, www.who.int/csr/resources/publications/influenza/WHO_CDS_CSR_GIP_05_8-EN.pdf, www.cdc.gov/flu/pandemic/phases.htm.

INTERNATIONAL SURVEILLANCE

The value of global surveillance and of information sharing (such as emergence of resistance in key pathogens) lies in:

- Guiding infection control policies
- Engaging and prompting dialogue with policy-makers
- Developing advocacy and educational programmes
- Stimulating research.

Global surveillance serves as an **early warning system** for epidemics and provides the rationale for public health intervention. Early detection of communicable diseases and immediate public health intervention can curtail the numbers of communicable illnesses and deaths, and reduce the negative effects on international travel and trade.

Global **surveillance** depends on strong national **surveillance** systems. The UK's HPA houses the WHO reference laboratories for various infections (e.g. influenza, TB, SARS) and it is the European Co-ordinating Centre for the Global Programme on Drug Resistant Tuberculosis.

Box 2G.2.2 shows an example of a global programme.

Box 2G.2.2

Example: The WHO Information for Action – global programme for vaccines and immunisation

The programme monitors and assesses the impact of strategies and activities for reducing morbidity and mortality of vaccine-preventable diseases. Its global goals by 2005 included:

- Polio eradication
- Measles mortality reduction
- Maternal and neonatal tetanus elimination (MNTE)

Reproduced from WHO, Immunization surveillance, assessment and monitoring, available online at: www.who.int/immunization_monitoring/en.

EVALUATING A SURVEILLANCE SYSTEM

An evaluation of a surveillance system should describe:

- The public health importance of the disease under surveillance
- Details of the surveillance system: its objectives, components and case definition (which should be clear and consistent)
- Details of the analysis and reporting system
- Conclusions and recommendations to improve the system.

It should consider each of the attributes shown in Box 2G.2.3.

USEFULNESS OF SURVEILLANCE

One way to assess the impact of a surveillance system is to list the actions that have been taken as a result of its findings. These might include:

- Early detection of outbreaks
- Monitoring of trends
- Early warning of changes of incidence
- Guidance to public health programmes (e.g. identification of high-risk groups to receive neonatal BCG to prevent childhood TB)
- Collection of cases for further studies.

Box 2G.2.3

Qualitative	Simplicity
	Flexibility
	Acceptability
	Representativeness
	Completeness
Quantitative	Sensitivity
	Positive predictive value
	Timeliness
	Resource use (direct costs)

2G.3 METHODS OF CONTROL

Communicable disease control consists of activities aimed at preventing the spread of infection. In a health-care setting it is termed 'infection control' and aims to avoid cross-contamination between patients, from health-care workers to patients, and from patients to health-care workers. See Table 2G.3.1.

Table 2G.3.1 Methods of communicable disease control

Universal precautions	Hand washing (e.g. hand hygiene campaigns)
	Infection control standard, contact, droplet and airborne precautions, e.g. facemasks, personal protective equipment
Isolation	Single room isolation, e.g. MRSA
	Negative pressure room for source isolation, e.g. TB
	Positive pressure room for protective isolation, e.g. for immunocompromised patients such as those receiving a bone marrow transplant
Decontamination	Decontamination of persons
	Disinfection of equipment and the environment
Quarantine	Quarantine of contacts, e.g. plague
Immunisation	Vaccine prophylaxis of exposed individuals
Chemoprophylaxis	Antibiotics offered to people who have come into contact with an infectious disease (e.g. meningococcal meningitis or anthrax)
Source removal	Biociding the water in a cooling tower to control an outbreak of legionnaires' disease
	Product recall
	Closure of a restaurant

2G.4 DESIGN, EVALUATION AND MANAGEMENT OF IMMUNISATION PROGRAMMES

Note that the terms **vaccination** and **immunisation**, while used interchangeably in common parlance, are in fact slightly different: see Box 2G.4.1.

Box 2G.4.1

Vaccination	Administration of a vaccine
Immunisation	Administration of a vaccine plus the development of an immune response by the body

DEVELOPMENT OF A VACCINATION PROGRAMME

In developing a new vaccination programme, the following areas of work must be considered:

- Scientific evidence
- Programme strategy
- Administration
- Finance
- Vaccine purchase and distribution
- Communication
- Informatics.

UK In the UK, the Department of Health (DH) has overall responsibility for immunisation policy. It is supported by the HPA, which undertakes vaccine research, epidemiology and surveillance. The Joint Committee on Vaccination and Immunisation (JCVI) is an independent expert advisory committee that advises the government on matters relating to communicable diseases. This committee receives papers and hears presentations and must make its recommendations in light of a cost–benefit analysis.

Ire In Ireland, the Health Service Executive (HSE) National Immunisation Office website has up-to-date information for parents and professionals on the childhood immunisation schedule and on other topics such as immunisation for travel (www.immunisation.ie/en).

The DH and the HSE take guidance from the Immunisation Advisory Committee of the Royal College of Physicians of Ireland. That committee has representatives from the Department, the HSE, the RCPI, the DH and Social Services, Northern Ireland and other relevant organisations. The Committee's guidance on immunisation is available on the Health Protection Surveillance Centre website (www.ndsc.ie/hpsc).

VACCINATION STRATEGIES

WHO recommendations are shown in Table 2G.4.1.

Table 2G.4.1 WHO recommendations for immunisations as part of national strategies:

All countries	Diphtheria Hepatitis B Measles Pertussis Poliomyelitis Tetanus
High-risk countries	Hib vaccine BCG vaccine Yellow fever

Immunisation schedules vary by country, depending on local epidemiology, funding, a consideration of the risk and efficacy of vaccines at different ages. The annual WHO/UNICEF joint reporting form lists the vaccination schedules of all countries (www.who.int/immunization_monitoring/data/schedule_data.xls).

In addition, targeted vaccination strategies may be implemented in certain circumstances: see Box 2G.4.2.

Box 2G.4.2

Circumstances		Example
Travel		Hepatitis A
Occupational		Hepatitis B
Outbreak		Meningitis C
Mass infection	**Eradicate**	Smallpox
	Eliminate	Polio
	Contain	Deliberate release of smallpox

IMPLEMENTATION OF A VACCINATION POLICY

Funds are normally secured centrally for implementation of the vaccine programme locally. Communication of the new policy is achieved through:

- Letters to all registered doctors, nurses and pharmacists from the Chief Medical Officer/Chief Nursing Officer/ Chief Pharmacist
- Website publicity
- 'Green Book' updates
- DH manages a network of 'immunisation coordinators'
- Public promotion (see below).

After resources have been secured, vaccine manufacturers are invited to submit bids through a competitive procedure. Criteria for successful bidding are safety, efficacy, availability, price and record of the company against previous contracts. Wherever possible, more than one supplier is chosen to minimise the likelihood of vaccine shortages.

PUBLIC RELATIONS

A new promotion programme is developed for the public, which may include the following: leaflets; fact sheets; press, television and radio advertisements; videos; and internet materials (including 'Question & Answer' formats, frequently asked questions (FAQs) and mail-in facilities for internet questions).

All new materials are pre-tested with the appropriate target audience and amended in light of consumers' comments, and their impact is monitored.

OTHER IMPLEMENTATION ACTIVITIES

- New surveillance arrangements to measure the impact of a new policy through laboratory-based data or disease notification data
- Sero-epidemiological surveillance for population impacts
- Market testing is undertaken to evaluate the impacts of any advertising accompanying the introduction of new vaccines
- New data collection arrangements are set up in advance so that the vaccine coverage of a new initiative can be monitored

- In the case of recent new vaccine policies, expert groups have been set up in advance to monitor and investigate reports of serious adverse events
- Special studies are also prepared that link hospital or primary care records of clinical events with immunisation data so that risks of adverse outcomes can be assessed.

LOCAL IMPLEMENTATION

- Will need local implementation group, usually lead by immunisation coordinators but may include consultant in communicable disease control (CCDC), pharmacists, health visitors, child health, community paediatricians, PCT primary care leads
- A local training programme for health professionals.

2G.5 CHOICES IN DEVELOPING AN IMMUNISATION POLICY (SALISBURY 2005)

Issues for policy-makers to consider when developing immunisation policy include:

- Outbreak response (including whether to create a vaccine stockpile)
- Surveillance
- Containment
- Investment in future research.
- Choice of selective versus universal or mass vaccination.

Examples of policy decisions are given in Table 2G.5.1.

Table 2G.5.1 Examples of recent UK policy development decisions

Pertussis	The change from whole-cell pertussis vaccine to acellular *pertussis* vaccine was not made until 2004 – later than in other industrialised countries. This was a deliberate decision since protection from the whole-cell vaccine was excellent, and reactions when the vaccine was given at 2, 3 and 4 months of age are fewer than when the vaccine is given later (e.g. as in other countries). In 2004, a 'five-component' acellular *pertussis* vaccine became available with efficacy shown to at least equal to that of the previously used whole-cell vaccine
Varicella	Varicella vaccine is not currently recommended for the routine UK childhood programme, although it is a routine vaccination in the USA. This policy will not be changed until there is confirmation that its introduction in childhood will not increase the burden of varicella-zoster in older age groups
Pneumococcus	Pneumococcal conjugate vaccine (PCV), which is in routine use in certain countries such as the USA, has been introduced into the UK childhood programme

2G.6 OUTBREAK INVESTIGATIONS

Outline the steps in outbreak investigation, including the use of relevant epidemiological methods

An outbreak is a localised epidemic (the latter being the occurrence of more cases of disease than expected in a given area, or among a specific group of people, over a particular period of time).

The objectives in controlling an outbreak will be to:

- Minimise the number of **primary cases** of illness through prompt recognition of the outbreak, and through the identification and control of the source of the infection or contamination
- Minimise the number of **secondary cases** of infection by identifying cases and taking appropriate action to prevent any spread
- Prevent further episodes of illness by identifying **continuing hazards** and eliminating them or minimising the risk that they pose

- Introduce measures to **prevent future outbreaks**.

OUTBREAK CONTROL PLANS

Outbreak control plans should include:

- Description of the **roles** and **responsibilities** of each of the organisations and individuals
- Arrangements for **informing** and **consulting** the key personnel (e.g. directors of public health, the regional epidemiologist, relevant reference laboratories, senior managers from the health service and health protection agency, and the DH)
- Arrangement for **liaison** with local government, hospitals and health authorities
- **Facilities** required to manage an incident (e.g. an incident room equipped with telephones, fax machines and other efficient electronic communication systems) – including arrangements for outside normal working hours.

OUTBREAK CONTROL GROUP

An outbreak control group should generally be convened when an outbreak occurs and any of the following apply:

- The disease poses an **immediate health hazard** to the local population
- There are a **large number of cases**
- Unexpected cases appear in **several districts**
- The disease is **unusual** and **severe**.

INVESTIGATION OF AN OUTBREAK

The management of an outbreak consists of five tasks that should be conducted concurrently – one of which is an epidemiological sequence that should be conducted serially (Figure 2G.6.1).

2G.7 IMPORTANT INFECTIOUS DISEASES

Knowledge of natural history, clinical presentation, methods of diagnosis and control of infections of local and international public health importance (including emerging diseases and those with consequences for effective control)

Tables 2G.7.1–2G.7.5 group diseases as follows:

- Food-borne diseases
- Viral hepatitides
- Vaccine-preventable diseases
- Nosocomial infections
- Sexually transmitted infections.

UK Note that the Health Protection Agency identifies four groups of people as posing an increased risk of spreading gastrointestinal infection: see Box 2G.7.1.

Box 2G.7.1

Group A	Any person of **doubtful personal hygiene** or with unsatisfactory toilet, hand-washing or hand-drying facilities at home, work or school
Group B	**Children** who attend pre-school groups or nurseries
Group C	People whose work involves **preparing or serving unwrapped foods** not subjected to further heating
Group D	**Clinical and social care staff** who have direct contact with highly susceptible patients or persons in whom a gastrointestinal infection would have particularly serious consequences

Reproduced from PHLS Advisory Committee on Gastrointestinal Infections (2004).

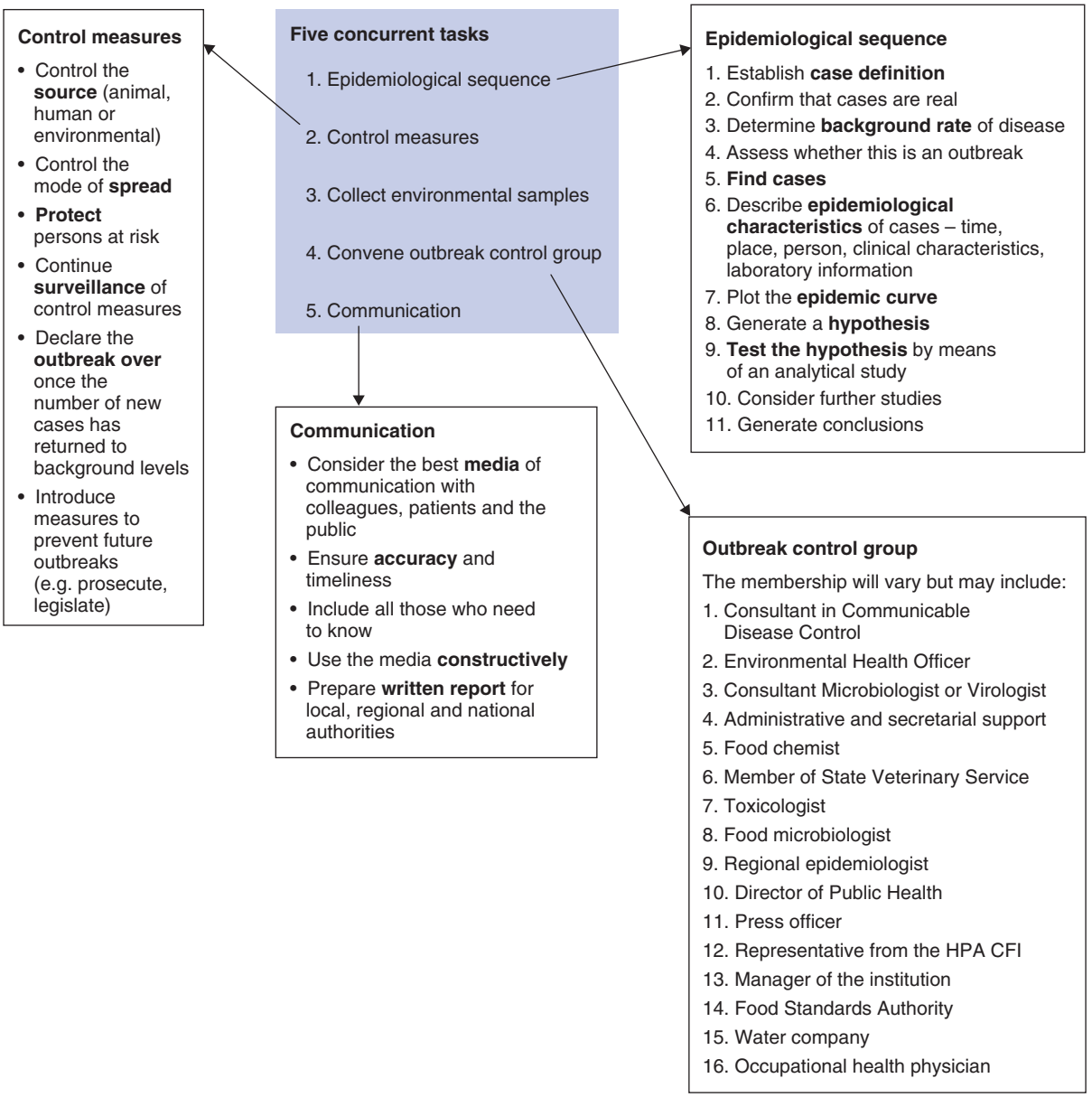


Figure 2G.6.1 Investigation of an outbreak

UK Note also that references to notifiable disease status in Tables 2G.7.1–2G.7.5 pertain to England and Wales (see Section 2G.8).

Table 2G.7.1 Food-borne gastrointestinal diseases

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Campylobacter	Ranges from Asymptomatic to severe diarrhoea (~50% cases bloody stools)	Stool culture (high sensitivity with same-day sample) Microscopy (sensitivity lower than culture)	Gastrointestinal tract of birds (particularly poultry) and animals, cattle and domestic pets	Animal to person (water or food contaminated with faeces) Person to person (direct contact with faeces of index case, e.g. person changing soiled nappies) Raw or undercooked meat (especially poultry), non-pasteurised milk	**Preventive measures:** Chlorination of drinking water supplies Milk pasteurisation Adequate hygiene, domestic and commercial Kitchens Adequate cooking of poultry Hand hygiene Advice to travellers abroad Control of patient, contacts and immediate environment **Exclusion** of symptomatic cases **Food hygiene** and **hand hygiene** Avoid **non-pasteurised milk** and untreated water

Table *contd overleaf*

Table 2G.7.1 *contd*

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Cholera **(toxin-producing *Vibrio cholerae* 01)**	Watery diarrhoea Vomiting 50% case fatality – severe untreated cases	Determine if toxin producing Stool culture Direct microscopy of stools PCR	Untreated/ polluted water	Consumption of untreated water, contaminated shellfish and foods eaten raw or washed in contaminated water. Person-to-person spread (by faecal–oral route) is likely to be a threat only when hygiene is very poor and sanitary facilities are inadequate Cholera rarely occurs in the UK – most cases are imported; therefore, ask about travel history in week before onset	Notifiable disease Secondary spread is rare if hygiene is good Cases should be excluded until 48 h after first normal stool Fatalities rare in the UK: ensure oral rehydration and appropriate antibiotics Food advice to travellers **Prevention:** Predominantly education, and adequate hygiene and sanitation, especially for travellers Advice for travellers to countries with epidemic cholera: A simple rule of thumb is, *'Boil it, cook it, peel it or forget it'* Vaccination: killed whole-cell vaccine leads to poor short-lasting cover and is of little value. Not available in UK and is no longer a requirement for travel to any country **Control:** Safe drinking water supplies

Cryptosporidiosis (*Cryptosporidium parvum*)	Healthy individuals – self-limiting Immunocompromised – severe illness may lead to death Diarrhoea – may be bloody	Stool microscopy Intestinal biopsy Serology Genotyping techniques	Gastrointestinal tract of humans and animals (particularly farm and domesticated); water contaminated with faeces	Person to person Animal to person Swimming pool outbreaks	**Prevention:** Hand hygiene Adequate water treatment Monitoring water quality Immunocompromised – avoid contact with farm animals, drink boiled water, avoid contact with infected cases Disinfectants, e.g. hydrogen peroxide **Control:** Exclusion of cases until 48 h after first normal stool Cases should avoid swimming for 2 weeks Contact tracing History of raw water consumption Good practice guideline – nursery/farm/swimming pool/hospitals

Table *contd overleaf*

Table 2G.7.1 *contd*

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Shigellae **(*Shigella sonnei* – common in the UK and mild)** **(S *dysentriae*, S. *flexneri*, S. *boydii* – imported and severe)** **(*S. dysenteriae* type 1– produce exotoxin – severe illness)**	*S. sonnei* – mild diarrhoea Others: Watery diarrhoea, vomiting 50% bloody stools *S. dysenteriae* 1 – toxic megacolon/ haemolytic – uraemic disease and death	Isolation of organism from stools Serotyping Phage typing	Humans	Person to person Contaminated food/ water Faecal–oral route	**Prevention:** Hand hygiene Adequate cleaning of toilet area Treatment of drinking and swimming water Advice to travellers **Control:** *S. sonnei:* exclusion of case for 48 h after first normal stool Other *Shigella* species: exclusion of case for 48 h after first normal stool (unless the case is in a risk group – when clearance is needed, i.e. two negative stools taken at least 48 h apart; contacts in risk groups also need microbiological clearance) Contact tracing Reinforce hygiene measures

E. coli **(vero cytotoxin producing)** Most common serotype in UK is *E. coli* 0157.H7	*E. coli* 0157 – asymptomatic/ diarrhoeal illness/ haemolytic–uraemic syndrome (HUS – particularly in children)/death	Food, environmental, animal samples Stool culture Biochemical and serological testing – isolates Reference labs for VTEC	Gastrointestinal tract of cattle (and possibly other domesticated animals)	Contaminated food/ water Animal to person Person to person	**Prevention:** Hand hygiene Adequate cleaning (kitchen, toilet) Precautions during farm visits Well-cooked beef, lamb, venison products Good practice – food processing and food service industries **Control:** Hygiene advice cases/contacts/food service industry Cases not in risk groups are excluded for 48 h after first normal stool. Cases in risk groups are excluded until microbiological clearance – two consecutive stool samples 2 days apart Household contacts in risk groups are screened microbiologically
Salmonellae **(*S. enteritidis* PT4 – associated with eggs and poultry)** **(*S. typhimurium* DT104 – increased antibiotic resistance)**	Diarrhoea Rare complication is abscess formation and septicaemia	Stool culture or rectal swab Blood culture **Reference laboratory:** Serotyping Phage typing	Gastrointestinal tracts of wild and domestic animals, birds (especially poultry), 'exotic' pets (terrapins and iguana) and occasionally humans	Animal to person – contaminated food Person to person – faecal–oral	**Prevention:** Vaccination of poultry flock Food processing industry systems to identify, control and monitor potential hazards Personal/food hygiene measures (home, institutions) **Control:** Hand/food hygiene advice Exclude cases for 48 h after first normal stool

Table *contd overleaf*

Table 2G.7.1 *contd*

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Paratyphoid fever **(*Salmonella paratyphi* A (80% in UK), B (20% in UK) and C)**	Gastroenteritis Early disease may involve constipation – later – diarrhoea, vomiting Spots 5% relapse	Blood, urine, faeces, bone marrow aspirate culture	Humans	Food borne Person to person in poor hygiene conditions	**Prevention:** Sanitation Clean water Personal hygiene No effective vaccine **Control:** Notifiable Isolation in hospital is advisable Screen all household contacts Hygiene advice Exclude food handlers (2 weeks after antibiotic therapy, require six clear consecutive samples – 2 weeks apart). Other risk groups: exclude until three clear consecutive samples taken 2 weeks apart Exclude cases not in risk groups until well and have normal stools

Typhoid (*Salmonella typhi*)	Similar to paratyphoid – but more severe Rose spots (undetectable in pigmented skin) Spleen may enlarge Abdominal haemorrhage (ulceration of Peyer's patches) Intestinal perforation Renal failure	Blood, urine, faeces, bone marrow aspirate culture Phage typing – unexplained clusters	Humans	Food borne Person to person with poor hygiene	**Prevention** Sanitation Clean water Personal hygiene Vaccination for travellers to endemic countries **Control:** Notifiable Exclude food handlers (2 weeks after antibiotic therapy require six clear consecutive samples – 2 weeks apart) Exclude food handlers (2 weeks after antibiotic therapy, require six clear consecutive samples – 2 weeks apart). Other risk groups: exclude until three clear consecutive samples taken 2 weeks apart. Exclude cases not in risk groups until well and have normal stools Contact tracing of household/close contacts

Table 2G.7.2 Viral hepatitides

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Hepatitis A (HAV)	Ranges from asymptomatic to fulminant hepatitis Young children are commonly asymptomatic. Adults are more likely to have symptoms – 75% of adults develop jaundice Infectious from 2 weeks before jaundice develops	Salivary IgM and IgG – useful in outbreak investigations Detection of IgM (serum/saliva) = acute infection IgG persists for life Persistent IgG – past infection/immunisation	Human gastrointestinal tract	Person to person spread – via faecal – oral route and contaminated food Consumption of food grown or washed in contaminated water (e.g. shellfish, fruit and vegetables)	Notifiable disease Use of immunoglobulin (Ig) or vaccination for close, household and sexual contacts should be offered via GP Risk groups and all cases should be excluded for 7 days after onset of jaundice and/or symptoms Travel advice for those to areas of endemicity
Hepatitis B (HBV)	Non-specific prodromal illness Jaundice (often after fever) Can lead to long-term carriage (carriage rate in UK population ~0.5%) Cirrhosis Hepatocellular carcinoma Co-infection with hepatitis subviral satellite D	See Table 2G.7.6 (HBV marker interpretation)	Humans	Person to person by blood-borne routes. Intravenous drug users, sex, bites, scratches Perinatal most common in high prevalence countries	Notifiable Vaccination of high-risk group – all health-care workers should be immunised against hepatitis B infection and should be shown to have made a serological response to the vaccine Boosters in poor or non-responders UK schedule 0, 1 and 6 months Accelerated schedule 0, 1, 2 and 12 months If susceptible, vaccinate household and sexual contacts Screening of blood supply in UK Antenatal screening and immunisation of babies at risk

Hepatitis C (HCV)	Asymptomatic Mild infection Jaundice – unusual 80% carriers 80% chronic hepatitis 10–20% cirrhosis 1% liver cancer	Enzyme immunoassay – detect HCAb Recombinant immunoblot assay (RIBA) confirms HCAb HCAb +ve – previous exposure PCR – detect infection (HCV RNA)	Humans	Historical: blood products until screened since 1991 Now intravenous drug users: 80% Sharing razors or toothbrushes Body piercing (like tattooing or acupuncture) Vertical transmission rare Risk is highest in HIV mothers and those who have high viral load Insufficient evidence to assess the risk of transmission via breast milk	Hepatitis C National Register was established in 1998 Aim of the register is to inform the natural history of HCV infection in the UK Majority of these cases are transfusion recipients who were traced during the national HCV look-back programme DH released hepatitis C strategy for England (2002) and hepatitis C: Action plan for England (2004) to implement strategy No vaccine Interferon and ribovirin treatment for chronic infection with HCV
Hepatitis D (HDV)	Exists only in conjunction with HBV. Known also as the 'delta agent' Increases risk of cirrhosis	Serology	Humans	As for HBV	As for HBV
Hepatitis E (HEV)	Illness similar to HAV without chronic sequelae or carriage Most cases – young/ middle-aged adults High case fatality in third trimester of pregnancy	Serology Specific IgM testing		Faecal–oral Person to person – low spread	Provision of safe water supplies Pregnant and older people, those with weakened immune systems and people with chronic liver disease might need closer observation for deterioration in liver function No vaccination

Table 2G.7.3 Vaccine-preventable diseases

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Diphtheria **(diphtheria toxin produced by toxigenic *Corynebacterium diphtheriae* or by *C. ulcerans*)**	Pharyngitis, enlarged lymph nodes and 'bull-neck' appearance may cause respiratory obstruction, paralysis and cardiac failure – fatal if untreated Many cases vaccinated, so rarely recognised on clinical grounds	Nasal, throat, skin ulcer swabs – identify *C. diphtheriae* (toxigenic) Occasionally toxigenic *C. ulcerans* Confirmation of toxigenicity from reference laboratory – can be obtained within few hours by PCR	Reservoir of *C. ulcerans* is cattle	Animal to person No direct evidence of person to person – but possible	Statutory notifications Rare in England and Wales following the introduction of mass immunisation in 1942 Travel, and close contact with cattle or other farm animals, are potential risk factors for infection Contact in the previous 7 days with a case of infection caused by toxigenic *C. diphtheriae* or *C. ulcerans* should be considered at risk Cases barrier nursed, antibiotic treatment, antitoxin, booster or primary vaccination Contacts – swabbed, food handlers/unvaccinated children excluded, antibiotic, booster/primary vaccination Primary vaccine coverage (three doses) for children aged 2 has been 94% since 2001, just below the WHO target of 95% No public health action reqiured if non-toxigenic *C. diphtheriae*

Pertussis (whooping cough) ***(Bordetella pertussis)***	Cough, cold, fever progressing to paraxysmal cough and bouts of coughing ending with a whoop or vomiting Last 2–3 months <6 months at risk Adults – milder symptoms – recognised as a cause for chronic cough	Culture nasal swab – but low sensitivity and high specificity PCR EIA	Human	Droplet spread	**Prevention:** Acellular pertussis vaccine is given in the primary course with diphtheria, tetanus, polio and Hib, as DTaP/IPV/Hib, at 2, 3 and 4 months of age. A further booster dose with acellular pertussis, given as dTaP/IPV, is given with the preschool boosters between the ages of 3 and 5. See: www.hpa.org.uk/infections/topics_az/vaccination/vac_sced.htm **Control:** Cases treated with antibiotic, isolated, vaccinated (if not vaccinated). If a vulnerable contact is present in a household, all offered antibiotic prophylaxis and vaccinate those under 10 and unimmunised. If no vulnerable contacts, no prophylaxis required. Outbreaks – community-wide vaccination if coverage low/case finding/antibiotic treatment
Tetanus ***(Clostridium tetani)***	Painful muscular contractions – especially of neck and jaw Often history of tetanus-prone wound	Infrequently obtained	Animals, humans	Dirty wounds Abdominal surgery	**Prevention:** Vaccine – in UK schedule = three doses at 2, 3 and 4 months, boosters at 3–5 years and 13–18 years Case – vaccinated primary or booster – if 10 years or more since last vaccine or if acquired tetanus-prone wound **Control of outbreak:** Look for source, e.g. surgery, intravenous drug users

Table *contd overleaf*

Table 2G.7.3 *contd*

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Haemophilus influenzae type b (Hib)	Invasive disease Commonest presentation is meningitis; other – pneumonia, epiglottitis, bone and joint infection, facial cellulitis	Blood/CSF culture PCR Reference laboratory for confirmation and typing	Humans	Droplet spread/direct contact Unvaccinated – higher carriage rate – common in young children	**Prevention:** Vaccine in UK given at 2, 3 and 4 months with diphtheria/tetanus/pertusis and polio vaccines Vaccination prevents carriage **Control:** Notifiable Unvaccinated household children contacts vaccinated and adults given chemoprophylaxis, including case Vaccination programme in cluster if coverage low
Meningococcal disease *(Neisseria meningitidis)*	Non-specific early phase Babies – floppy, fever, vomiting Photophobia, neck stiffness Petechial rash, septicaemia, death Complications: deafness/convulsions/ limb amputation/ mental impairment	Blood/CSF culture PCR Reference laboratory for confirmation and typing	Humans	Direct or indirect person to person spread	**Prevention:** Vaccines against serogroups A, C, W135 and Y – short-lived vaccine Men C vaccine given in the UK at 2, 3 and 4 months (children >1 year single dose) **Control:** Chemoprophylaxis – close contacts Vaccinate if vaccine-preventable strain

***Mycobacterium tuberculosis* (MTB)** **(occasionally *M. bovis* and *M. africanum*)**	Long incubation period produces chronic disease with risk of reactivation (particularly with age) and fatal without treatment Symptoms can include: Cough Blood in sputum Weight loss Night sweats More common in immigrant ethnic groups Mortality decreased rapidly after introduction of effective chemotherapy	Chest radiograph Sputum smear Sputum culture Microscopy – sputum Sensitivity testing for multi-drug-resistant TB (MDRTB) Molecular typing for identifying clusters	Animals, humans	Direct spread from infected case Bovine TB from ingesting raw milk from infected cows	**Prevention:** BCG vaccine: UK government now recommends that the following risk groups be offered BCG vaccination: • All infants living in areas where the incidence of TB is 40 per 100 000 or greater • Infants whose parents or grandparents were born in a country with a TB incidence of 40 per 100 000 or higher • Previously unvaccinated new immigrants from high prevalence countries for TB **Control:** Notifiable UK – enhanced TB surveillance since 1999 Cases are followed up by chest clinics to ensure that adequate treatment is given and that contacts are identified, screened and given prophylaxis where necessary Measures to maximise compliance such as directly observed therapy (DOTS) Port health
Mumps (paramyxovirus)	Tenderness and parotid swelling Meningitis (commonest cause of viral meningitis pre-vaccine era) Orchitis Pancreatitis	Saliva, CSF, urine culture Serology	Humans	Direct contact with saliva or droplets of saliva of an infected person	**Prevention:** Vaccine in the UK given at 12–15 months and 3–5 years (in combination with measles and rubella) **Control:** Notifiable As vaccination rates drop, resurgence and possible outbreaks of these diseases are increasingly likely Exclusion Check vaccination status for case Consider community-wide programme if coverage low in outbreaks

Table *contd overleaf*

Table 2G.7.3 *contd*

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Measles (paramyxovirus)	Prodromal flu-like symptoms Koplik's spots inside the mouth Rash starts on days 3–4, over face, trunk and limbs Not itchy Complications include: • Pneumonitis • Acute otitis media • Pneumonia • Encephalitis	Salivary test kit for measles IgM Serology	Human	Direct contact Person to person	See mumps Human normal immunoglobulins (HNIG) for: Pregnant contacts who are measles IgG negative Babies less than 6 months old – maternal IgG negative Exclusion
Rubella **(rubella virus – a member of Togaviridae)**	Moderately infectious Pre-vaccination era Affecting primary school-aged children, susceptible pregnant women – congenital rubella syndrome Pharyngitis Conjunctivitis Fever Rash	Serum or saliva detection of IgM Viral culture from urine or serum	Human	Direct person to person	See mumps Pregnant women in contact with case – tested; if susceptible offer vaccination post partum

Table 2G.7.4 Nosocomial infections

Nosocomial (health care acquired) infections

Health care-associated infections (HCAIs) are infections acquired as a result of health care. Examples include meticillin-resistant *Staphylococcus aureus* (MRSA), *Clostrdium difficile,* glycopeptide-resistant enterococci (GRE) and *Acinetobacter* species. Higher rates of HCAIs are often found in specialist hospitals such as orthopaedic centres. However, with the move towards performing invasive procedures such as minor surgery in the community, primary care HCAIs may become more common in future

HCAI risk factors

- More **susceptible patients** being treated (elderly patients or patients with severe or chronic diseases)
- **Invasive procedures** (e.g. indwelling lines into veins, access for dialysis or artificial ventilation)
- **Immunosuppression** (e.g. chemotherapy and anti-rejection drugs used in transplant surgery)
- Increased patient movement between wards or hospitals due to pressures on hospital beds
- Wider use of **antibiotics** and emergence of antibiotic-resistant microorganisms

Impact of HCAI

Effects on patients	Effects on the health service
Severe or chronic illness	Extended lengths of stay
Pain, anxiety, depression	Costs of diagnosis and treatment of the infections and their complications
Longer stay in hospital	Costs of specific infection-control measures – cleaning, disinfection, cohort nursing, etc.
Reduced quality of life	Bed and ward closures and postponed admissions
Loss of earnings	Provision of isolation facility/rooms
Death	Antibiotic costs may be further increased if the infection is also due to a resistant microorganism

Control of HCAIs

- **Hand washing** is the most important prevention activity. In the UK the *'Clean your hands'* campaign raised awareness among staff, patients and the public
- Prudent antibiotic **prescribing**
- **Surveillance** – in the UK surveillance systems for MRSA, surgical site infection
- **Isolation**, cohort nursing
- Local and national policies

Table *contd overleaf*

Table 2G.7.4 *contd*

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
MRSA **(meticillin-resistant *Staphylococcus aureus*)**	Range from skin infection and conjunctivitis, to pneumonia and life-threatening septicaemia 30% of the general population are colonised by *S. aureus*. In hospitals the percentage is higher because of more likely contact with infected cases	Gram stain, culture and sensitivity testing on appropriate specimen 16 strains EMRSA-15 and EMRSA-16 are the dominant UK stains	Humans, rarely animals	Direct contact	**Prevention:** Good personal hygiene – hand washing Compliance with control measures Hand washing/aseptic techniques/handling waste/waste disposal/ward or equipment cleaning **Control:** Mandatory surveillance schemes in the UK: • Mandatory *S. aureus* bacteraemia surveillance • Mandatory MRSA bacteraemia-enhanced surveillance scheme Infection control policies – central and local Central government initiatives (DH) – Towards Cleaner Hospitals and Lower Rates of Infection, Saving Lives Delivery Programme to reduce HCAIs, including MRSA Antibiotic prescribing policy

Table 2G.7.5 Sexually transmitted infections (STIs)

Prevention involves:

- Health and sex education
- Early detection and prompt effective treatment
- Contact tracing and treating contacts
- Opportunistic or routine screening
- Surveillance
- In the UK KC60 returns from genitourinary medicine (GUM) clinics

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Chlamydia trachomatis	Commonest bacterial STI in the UK Highest rates in young people (especially under 24 years) Majority asymptomatic in men Untreated may lead to pelvic inflammatory disease (PID), ectopic pregnancy and ophthalmia neonatorum	PCR or culture of urethral, cervical or urine samples	Human	Sexually	**Prevention:** National Chlamydia Screening Programme commenced in 2002 in England Antibiotic treatment Contact tracing and treating partners Surveillance in England and Wales KC60 returns – GUM clinics
Gonorrhoea (*Neisseria gonorrhoeae*)	*Neisseria gonorrhoeae* is the second most common bacterial STI in the UK Men more likely to have symptoms than women Complications: • PID, ectopic pregnancies • Septic arthritis • Ophthalmia neonatorum	Microscopy or culture of urethral, cervical swabs	Human	Sexually	**Prevention:** Antibiotic treatment – but many strains resistant to commonly used antibiotics Contact tracing and treating partners Outbreak control Surveillance (KC60) GRASP (gonococcal resistance to antimicrobial surveillance programme)

Table *contd overleaf*

Table 2G.7.5 *contd*

Disease (organism)	Clinical features	Diagnosis	Reservoir	Transmission	Response/control
Syphilis (***Treponema pallidum***)	Primary ulcer: third of cases develop secondary eruption. Late lesions of skin, bone, central nervous system, heart	Microscopy in early syphilis Serological tests* – treponemal (e.g. TPHA) and non-treponemal (e.g. VDRL) test	Human	Sexually Mother to baby Blood transfusion	**Prevention:** In the UK Enhanced Surveillance – mainly due to outbreaks seen in Manchester, London, Bristol and Brighton among gay men and heterosexual men and women KC60 return Routine antenatal screening Antibiotic treatment Contact tracing Outbreak control Late syphilis – test for HIV also
Human immunodeficiency virus (HIV)	HIV continues to be one of the most important communicable diseases in the UK. It is an infection associated with serious morbidity, high costs of treatment and care, significant mortality and high number of potential years of life lost	HIV antibody test P24 antigen – for early tests and screening blood Viral load/CD4 count	Human	Person to person Sexually Blood transfusion Sharing needles Vertical transmission	**Prevention:** Surveillance Routine antenatal screening Education Contact tracing Antiretroviral treatment (no vaccine) Post-exposure prophylaxis

**TPHA,* Treponema pallidum *haemagglutination assay; VDRL Veneral Disease Reference Laboratory.*

HEPATITIS B SEROLOGY

Laboratory reports for hepatitis B contain details of a number of markers, the interpretation of which is outlined in Table 2G.7.6.

Table 2G.7.6 Interpretation of hepatitis B markers

Laboratory abbreviation	Serological marker	Description	Implication
HBsAg	Hepatitis B surface antigen	Serological marker on surface of HBV that is present in serum during acute or chronic infection	Person is infectious. Presence for >6 months implies chronic carrier status
Anti-HBs	Hepatitis B surface antibody	Antibody to surface antigen that is usually produced as part of the normal immune response	Person is immune (either due to recovery from prior infection or due to vaccination)
Total anti-HBc	Total hepatitis B core antibody	Antibodies (of all classes) to a component of HBV	Previous or ongoing infection
IgM anti-HBc	IgM hepatitis B core antibody	IgM class of antibody that persists for 6 months following exposure	Acute or recent infection
HBeAg	Hepatitis B e-antigen	Marker present soon after exposure, then absent within 3 months	High infectivity
Anti-HBeAb	Hepatitis B e-antibody	Develops after HBeAg	Low infectivity

NEW AND EMERGING INFECTIONS

Emerging infectious diseases are commonly defined as those that have newly appeared in a population or have existed but are rapidly increasing in incidence or geographical range.

DRIVING FORCES BEHIND GLOBAL EMERGING INFECTIONS:

The pattern of communicable disease occurrence is in constant flux. Current influences include:

- Global travel
- Climate change
- Global trade and importation
- Urbanisation
- Population displacement
- Animal movements
- Changes in agriculture
- Emerging zoonoses
- Deforestation
- Bird migration
- Human conflict
- Antimicrobial resistance
- Genetic mutation/recombination
- Deliberate release.

CURRENT UK AND EUROPEAN CONCERNS:

Diseases attracting particular research attention and interest in control policy include:

- Smallpox
- Anthrax
- SARS
- Avian influenza and pandemic influenza
- West Nile virus
- Changes in vector distribution
- Pet travel scheme
- Babesiosis
- Leishmaniasis
- Hantavirus
- MRSA, vancomycin-resistant enterococci (VRE).

NZ NEW ZEALAND – IMPORTANT INFECTIONS

In addition to the lists above, infectious diseases of particular importance in NZ include those shown in Box 2G.7.2.

Box 2G.7.2

Leptospirosis	This zoonotic infection follows exposure to urine from infected animals and urine-contaminated surface water. It is an important occupational zoonotic disease for farmers and abattoir workers in New Zealand. Individual cases and outbreaks are common in developing countries, particularly following flooding
Rheumatic fever	Acute rheumatic fever may occur following streptococcal throat infection and may result in serious damage to heart valves, leading to chronic rheumatic heart disease (CRHD). Acute rheumatic fever and CRHD remain an important cause of morbidity and mortality for indigenous people in New Zealand (Maori), Australia (Aborigines) and the Pacific Islands, as well as people living in many developing countries

PANDEMIC INFLUENZA VIRUS IN HUMANS (H5N1)

Outbreaks of influenza A/H5N1 (a highly pathogenic form of avian influenza) are thought to represent a serious threat to public health. Although A/H5N1 has appeared before, it has caused concurrent outbreaks in several countries and is proving difficult to eliminate.

Box 2G.7.3 summarises some terminology.

Box 2G.7.3

Seasonal influenza	RNA virus of the Orthomyxoviridae family. It rapidly spreads around the world in a seasonal pattern
Avian influenza	Form of influenza that is virulent in birds
Influenza A/H5N1	Highly pathogenic form of avian influenza
Pandemic influenza	Worldwide epidemics of human influenza such as those that occurred in 1918 and 1957

It is feared, for several reasons, that A/H5N1 might trigger the next pandemic:

- Ability to infect humans and cause severe disease
- Ability to mutate and to acquire or exchange genes from other viruses
- Ongoing spread in birds.

The likelihood of a pandemic increases as more people become infected over time, especially if they are concurrently infected with human influenza. In these circumstances a novel viral subtype may emerge that can be transmitted readily person to person. The likelihood of such a mutation occurring is difficult to predict.

EMERGENCY PLANNING

The key steps in dealing with SARS, pandemic influenza and other emerging communicable disease threats are:

- **Assess**
- **Prevent** (vaccinate)
- **Prepare** (plan for surge capacity, stocks of facemasks)
- **Respond**
- **Recover** (from the event including psychological care).

Note that these are the same principles as for other disasters such as flood, fire or terrorist incidents.

2G.8 ORGANISATION OF INFECTION CONTROL

Organisation is at several levels.

LOCAL GOVERNMENT

Local authorities have the power to take action to control notifiable diseases within their boundaries. They are required to appoint a 'proper officer' – usually a Consultant in Communicable Disease Control (CCDC) or a Designated Medical Officer in Scotland. Environmental health services have a duty to register, inspect and investigate food premises, and have legal powers of enforcement and prosecution.

HEALTH SERVICES

As part of their remit to promote health and prevent disease, health authorities/boards are responsible for the surveillance of disease, identifying problems and establishing planning measures.

HOSPITALS

Each hospital is required to have in place an infection control team (ICT) consisting of an infection control physician or microbiologist, and an infection control nurse (ICN) and in the UK one of them will act as Director of Infection Prevention and Control (DIPC). The activities of the ICT are listed in Table 2G.8.1.

Table 2G.8.1 Activities of the ICT

Planning	• Developing policies and procedures • Accommodation • Purchasing equipment, etc. • Clinical waste
Education	• Education of staff • Audit • Handwashing
Surveillance	• Antibiotic use
Outbreak	• Advise on outbreaks • Use of isolation facilities

The ICT reports to the hospital infection control committee, the members of which may include the Chief Executive (or director level deputy) and a CCDC. The committee meets at regular intervals to review infection control.

COMMUNITY

Community ICNs (Infection Control Nurses) encourage collaboration among community staff, ICTs (Infection Control Teams) and CCDC (Consultants in Communicable Disease Control) as well as with care homes, prisons, nurseries and schools.

UK NATIONAL CENTRES: HEALTH PROTECTION AGENCY – CENTRE FOR INFECTIONS

These are responsible for:

- Infectious disease surveillance
- Provide specialist and reference microbiology and microbial epidemiology
- Coordinate investigation and cause of national and uncommon outbreaks
- Helping advise government on the risks posed by various infections
- Respond to international health alerts.

Eng DEPARTMENT OF HEALTH

Central government has overall responsibility for national policy matters relating to infection control. The Chief Medical Officer and Chief Nursing Officer develop policies, guidance and tools for the NHS. They seek advice from experts within and outside the DH to advise government on infection control matters, including:

- Aseptic technique
- Cleaning
- Decontamination
- Handwashing
- Invasive devices (installing catheters, etc.)
- Isolation
- Laboratory specimen guidance
- Laundry and linen handling/management
- Handling of sharps
- Waste disposal/management.

Examples of recent initiatives are given in Box 2G.8.1.

Box 2G.8.1

Examples: recent initiatives from the DH related to infection control

- Health Bill (2006) – includes Hygiene Code of Practice for the Prevention and Control of Healthcare Associated Infections
- Saving Lives – a delivery programme to reduce health-care-associated infections, including MRSA (2005)
- A Matron's Charter – an action plan to cleaner hospitals (2004)
- Getting Ahead of the Curve – a strategy for combating infectious diseases (including other aspects of health protection) (2002)

NOTIFIABLE DISEASES

Eng Wal The diseases in Table 2G.8.2 are statutorily notifiable: all doctors working in England and Wales are required to report to the proper officer of the local authority of any **suspected** cases. The proper officer is required to provide anonymous details to the Health Protection Agency every fortnight.

Table 2G.8.2 Notifiable diseases in England and Wales

Anthrax	Meningitis	Scarlet fever
Cholera	Meningococcal septicaemia	Smallpox
Diphtheria	Mumps	Tetanus
Dysentery	Ophthalmia neonatorum	Tuberculosis
Encephalitis (acute)	Paratyphoid fever	Typhoid fever
Food poisoning	Plague	Typhus fever
Leptospirosis	Poliomyelitis (acute)	Viral haemorrhagic fever
Leprosy*	Rabies	Viral hepatitis
Malaria	Relapsing fever	Whooping cough
Measles	Rubella	Yellow fever

**Leprosy should be reported directly to the Centre for Infections at the Health Protection Agency.*

Scot All doctors working in Scotland are required to report any suspected cases to the designated medical officer of the local authority. The designated medical officer is required to provide anonymous details to Health Protection Scotland every fortnight.

Note that, in Scotland, **chickenpox** and **Legionnaires' disease** are notifiable diseases (in addition to the main UK list above).

2G.9 MICROBIOLOGICAL TECHNIQUES

Basic understanding of the biological basis, strengths and weaknesses of routine and reference microbiological techniques

There are two main methods of microbiological analysis – the traditional method involves growing a culture of the specimen in order to isolate and identify the microorganism (bacteria, fungi, viruses and parasites). The alternative is a range of modern molecular methods that involve the identification of specific DNA or RNA (e.g. RNA transcriptase) within the specimen. Tuberculosis, for example, once took 12 weeks to diagnose; now using molecular methods it takes 24 h.

MAIN CATEGORIES OF METHODS USED IN MICROBIOLOGY LABORATORIES

- Microscopy (including immunofluorescence)
- Culture
- Identification (e.g. typing of bacterial strain)
- Isolation of virus
- Drug sensitivity
- Serology (including immunoassay for antigen and antibody).

ROUTINE MICROBIOLOGICAL TECHNIQUES

Local hospital laboratories have tended to use traditional techniques, the strengths and weaknesses of which are listed in Box 2G.9.1.

Box 2G.9.1

Strengths	Weaknesses
Relatively low cost Can provide definitive diagnosis	Limited ability of laboratories to provide doctors with timely and clinically relevant information Low sensitivity, e.g. samples taken after antibiotic has been given may test negative Samples taken after onset of illness may result in difficulty isolating pathogen, e.g. viruses Limited range of tests available – may not be able to provide full identification, e.g. toxin-producing strains

REFERENCE MICROBIOLOGICAL TECHNIQUES

Molecular biological techniques form the basis of detecting and characterising an ever-increasing range of viruses, bacteria, fungi and protozoa.

Nucleic acid probes are commercially available for cytomegalovirus, human papillomavirus, hepatitis B virus, hepatitis C virus, *Chlamydia trachomatis*, *Neisseria gonorrhoeae, Streptococcus pyogenes* and mycobacteria, among others.

Nucleic acid amplification systems are available for the direct detection in clinical specimens of hepatitis C virus, HIV, *M. tuberculosis*, *C. trachomatis* and *N. gonorrhoeae*.

Strengths and weaknesses of these techniques are listed in Box 2G.9.2.

Box 2G.9.2

Strengths	Weaknesses
Increased speed Increased sensitivity and specificity Identify organisms that do not grow (or grow only slowly) in culture Identify genes that result in resistance to antibiotics 'Fingerprint' individual isolates for epidemiological tracking Recognition of newly emerging infectious diseases Control of antibiotic resistance in *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Staphylococcus aureus* and common Gram-negative bacilli	Need for specialised equipment Segregated rooms in laboratories Currently detect only microorganisms Turnaround times for existing tests are much longer than can potentially be achieved using molecular methods

IMMUNOASSAYS

These are used in the detection of microbial antigens and offer the potential for rapid diagnosis. Examples include **enzyme-linked immunoassays** and **direct immunofluorescence antibody assays**. Box 2G.9.3 lists strengths and weaknesses of immunoassays.

Box 2G.9.3

Strengths	Weaknesses
Technical simplicity	Poor sensitivity
Rapidity	Low negative predictive value
Specificity	Low positive predictive value
Cost-effectiveness	

AUTOMATED AND SEMI-AUTOMATED SYSTEMS

These fall into two main groups:

- Identification and susceptibility testing (some can provide results within a single working day); and
- Blood culture systems (most true positive results are detected within 24–36 h).

Some blood culture systems have been adapted for automated or semi-automated culture (e.g. for *M. tuberculosis* and other mycobacteria). These enable the identification and susceptibility results to be processed from large numbers of blood culture samples.

Box 2G.9.4 lists the strengths and weaknesses of automated and semi-automated systems.

Box 2G.9.4

Strengths	Weaknesses
Reduce the traditional dependence on biochemical reactions to identify organisms	Organisms may be incorrectly identified, e.g. database does not include the correct identification
Avoid the many labour-intensive steps between isolating and reporting clinically significant bacteria	Bacteria with heteroresistance to β-lactam drugs, inducible resistance mechanisms or susceptibility gene mutation may be misclassified
Provide rapid results	
Perform tests more reproducibly	May miss resistance of an organism to antibiotic, e.g. enterococci to glycopeptides; use supplemental testing with manual methods for problematic combinations of organisms and drugs

Molecular methods undoubtedly have enormous potential in diagnosing infectious diseases. New molecular methods will be widely accepted and implemented routinely within the next decade.

2G.10 INTERNATIONAL ASPECTS OF COMMUNICABLE DISEASE CONTROL

International aspects of communicable disease control, including port health

Globalisation has increased the risk of international spread of infectious diseases. In the past, the most concrete measures to stop importation of infectious diseases were thought to be quarantine and trade embargoes.

INTERNATIONAL OBLIGATIONS

INTERNATIONAL HEALTH REGULATIONS

This is a multilateral initiative by countries to develop a global tool for the surveillance of cross-border transmission of diseases. It balances the protection of public health with the avoidance of unnecessary disruption to trade and travel.

CORE OBLIGATIONS FOR WHO MEMBER STATES

Countries are obliged to notify the WHO of public health emergencies of international concern. They must also:

- Respond to requests for **verification** of information regarding urgent national risks
- **Control** urgent national public health risks that threaten to transmit disease to other member states
- Provide routine **port** inspection and control activities to prevent international disease transmission
- Apply the **measures** recommended by the WHO during public health emergencies.

CORE OBLIGATIONS OF THE WHO

The WHO has a duty to respond to the needs of member states regarding the interpretation and implementation of its regulations. It must update these regulations (and their supporting guides) so that they remain scientifically valid. In addition, the WHO must publish recommendations for use by member states during public health emergencies of international concern.

PORT HEALTH

In the UK, the regulations for ships, aircraft and international trains give local authorities and port health authorities the power to appoint medical and non-medical port health officers who can prevent the entry of communicable diseases into the country.

CIRCUMSTANCES REQUIRING THE INTERVENTION OF PORT HEALTH STAFF

- **Outbreak** of food- or water-borne disease on the vessel
- **Contamination** of aircraft by faeces or vomit
- Pests (rodents or insects) on board
- Passengers or crew who are suspected of being infected with viral haemorrhagic fever, yellow fever, plague, cholera, diphtheria or TB.

The Port Medical Inspector advises immigration officers on matters of health protection. Immigration officers may refer passengers who are emigrating to the UK, long-stay visitors and those visiting for health reasons. Some such passengers may be required to have a chest X-ray, with the findings being passed to the CCDC.

2H
Principles and Practice of Health Promotion

This chapter is about encouraging people to adopt actions that reduce the risk of developing disease. It is increasingly recognised that health promotion is at its most effective when it focuses on enabling people to increase control over their own health. For this reason it is often helpful to regard the discipline as a sociopolitical process.

2H.1 RESPONSIBILITY FOR HEALTH

Collective and individual responsibilities for health: both physical and mental

There are different views about the extent to which health is a collective or an individual responsibility. These perspectives determine how societies organise themselves to improve health.

- **Social responsibility** is a doctrine that holds that an entity (be it state, government, corporation, organisation or individual) has a responsibility to society as a whole. This responsibility can be **negative** (i.e. a responsibility to refrain from acting) or **positive** (i.e. a responsibility to act).
- In contrast, **individualism** is a moral, political and social philosophy, which emphasises the importance of the individual. Its central tenets are individual liberty, the *'virtues of self-reliance'* and personal independence.

Proponents of public initiatives and social responsibility argue that their policies are beneficial to the individual, and that excessive individualism may actually be detrimental to the individuals themselves. In contrast,

individualists hold that public initiatives may have unintended consequences beyond the issues that they are intended to address. Many commentators find the *'beneficial to the individual'* argument condescending and argue that individualism is not about individual benefit so much as individual choice.

It generally falls to politicians to decide which paradigm dominates health policy in any one country at a given time. For example, the absence of a universal health service in the USA has its roots in the political belief that individuals, rather than society, have responsibility for health care.

COLLECTIVE RESPONSIBILITIES

Approaches that emphasise collective responsibilities for health encompass population-wide measures. They include those shown in Box 2H.1.1 for the UK.

UK Box 2H.1.1

Policy	Example
Legislation	Drink–driving laws exist not just to protect the individual but to ensure that the individual does not put others at risk
Regulation	Health and safety legislation and regulations enable external bodies (e.g. the Health and Safety Executive) to inspect and ensure that businesses are working to protect their employees
Population-wide measures	Fluoridation of the water supply
Progressive health service funding systems	Universal taxation where high earners in society provide a larger contribution for NHS costs, despite the fact that those earning least may actually use the service more

INDIVIDUAL RESPONSIBILITIES

Approaches based on **individual responsibility** focus on initiatives to enable individuals to make an **informed choice**. They include those shown in Box 2H.1.2.

Box 2H1.2

Policy	Example
Information provision	Providing safe drinking limits allows individuals to choose how much they drink given knowledge of the health consequences
De-regulation	Relaxation of licensing to allow pubs and shops to sell alcohol 24 h a day relies on individual choice regarding when and how much to drink
Choice in health care	Private health-care insurance enables individuals to choose where they receive health care and whether or not to insure their health

MIXED APPROACHES

Eng England's White Paper *Choosing Health* (DH 2004) contains many proposals both for collective action and for encouraging people to assume individual responsibility. This paper set out the future direction for population health improvement in the country. It covered six main themes, including a mixture of legislation and measures to help people make positive, informed choices about risk factors associated with their future health.

The six main areas prioritised for action were:

1. Reducing the numbers of people who smoke

2. Reducing obesity and improving diet and nutrition
3. Increasing exercise
4. Encouraging and supporting sensible drinking
5. Improving sexual health
6. Improving mental health.

At a collective level, the White Paper resulted in legislation to ban smoking in enclosed public places. Other mechanisms for supporting people to make healthy choices included:

1. Marketing health
2. Improving information about health for the public
3. Tackling health inequalities
4. Partnership with industry (e.g. proposals for voluntary agreements on food marketing)
5. Promotion of healthy food for children
6. Further restrictions on tobacco advertising.

2H.2 DETERMINANTS OF HEALTH

Interaction between genetics and the environment (including social, political, economic, physical and personal factors) as determinants of health, including mental health

See also Section 2I.2.

The factors that have the most significant influence on health are termed the **determinants of health**. While health care and Social Services focus largely on dealing with the consequences of poor health, most of the key determinants of health as a *positive* attribute lie outside the direct influence of these services. They are influenced more by factors such as education, employment, housing and environmental policy.

Genetic and epidemiological studies offer understanding of the relative contributions of such factors to health and illness. This is useful when developing health promotion interventions, in particular:

- Whether to use **targeted** or **universal** programmes
- How to **allocate resources**
- Predicting **susceptibility**
- Predicting **uptake** of health promotion in the population.

Several theories exist for contemplating the range of influences on health. This is potentially confusing, but it reflects the rapid advances in this important field over the last half century, and the changing political and cultural perspectives from which the theories arise. The four frameworks on the determinants of health are summarised in Table 2H.2.1, and will then be discussed in turn.

Table 2H.2.1 Determinants of health frameworks

Framework	Author(s)	Year	Summary
Health field concept	Lalonde	1974	Health care is not the sole determinant of health Fields are biology, lifestyle, environment and health care
Policy rainbow	Dahlgren and Whitehead	1991	Determinants of health exist as interrelated layers of influence
Health field model	Evans and Stoddart	1990	Health is not only the absence of disease, but also takes into account functional status and wellbeing
Social determinants model	Diderichsen and Hallqvist	1998	Social conditions affect individuals' social situations, which in turn determine their health risks

LALONDE HEALTH FIELD CONCEPT (1974)

Marc Lalonde, Health Minister for Canada 1972–1977, proposed the health field concept in 1974 in a seminal report: *A New Perspective on the Health of Canadians*. Building on ideas of Thomas McKeown, it used evidence of mortality and morbidity in Canadians to argue that health-care services were not the most important determinant of health. He identified four fields – biology, lifestyle, environment and health care – as determinants of health (see Figure 2H.2.1) and started a new direction in Canadian health policy:

> *'The Government of Canada now intends to give to human biology, the environment and lifestyle as much attention as it has to the financing of the health care organisation so that all four avenues to improved health are pursued with equal vigour.'*
>
> *Reproduced from Lalonde (1974)*

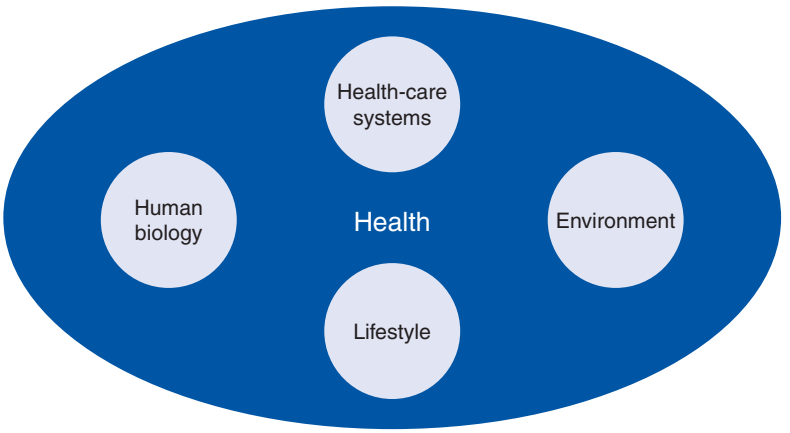


Figure 2H.2.1 Lalonde health field concept

The model has undergone refinements following criticism that it focused too much on lifestyle and too little on environment. However, it was hugely influential throughout the world and was pivotal in the growth of the discipline of health promotion.

DAHLGREN AND WHITEHEAD (1991)

Dahlgren and Whitehead's health 'rainbow' identifies a range of determinants of health (see Figure 2H.2.2). It recognises that some determinants (e.g. age and sex) are not modifiable, but that many can be altered. It builds on previous models by giving an indication of the different levels at which health is affected: see Box 2H.2.1. The model makes no attempt to explain the relationship between the different tiers, nor between factors in the same tier – but instead aims to stimulate discussion about the interrelationship of different layers and the relative importance of each to health. Furthermore, the model aims to promote debate regarding potential interventions for improving health and reducing inequalities in each of the four layers. Box 2H.2.2 illustrates how the rainbow has been used to develop policy.

Box 2H.2.1

Individuals	Although some individual factors (e.g. age) are fixed, others (e.g. lifestyle and behaviour) can be influenced. Influences include information and education, and also influences on the *'distal'* determinants in the other layers
Communities	Strengthening communities through action to improve the local environment and living conditions, or through *'bottom-up'* action led by local community groups
Access	Improving access to services such as health care, leisure, transport in terms of location, cost, appropriateness
Macroeconomics	Engendering macroeconomic or cultural change at national or global levels, possibly through legislative changes

The relative importance of factors in each layer depends on the health issue or population under discussion.

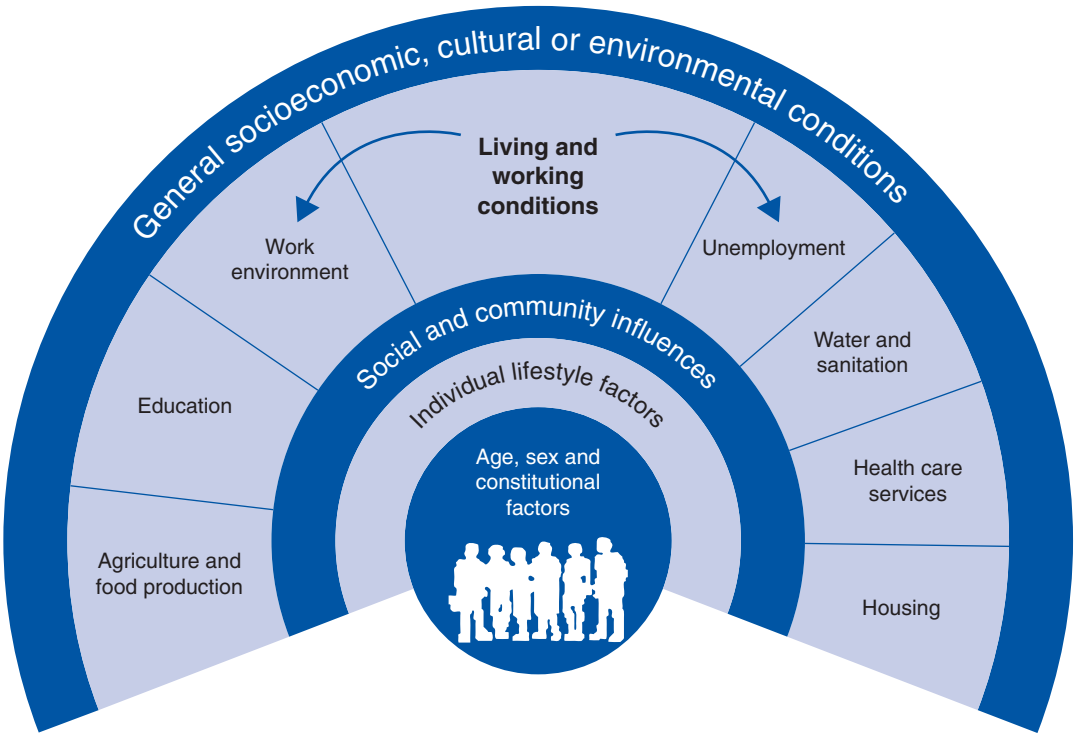


Figure 2H.2.2 Dahlgren and Whitehead's health rainbow. *Reproduced from Dahlgren and Whitehead (1991)*

Eng **Box 2H.2.2**

Example: London Health Commission reports

The work of the London Health Commission is informed by Dahlgren and Whitehead's policy rainbow. The Commission produces an annual report on the health of Londoners, focusing on variations on ten indicators including:

1. **Health outcomes**: life-expectancy, infant mortality and self-reported health status
2. **Determinants of health** related to:
   - Social and community influences – levels of crime
   - Living and working conditions – employment, education and housing
   - Environmental factors – road safety, air pollution

The Commission's analysis links individual constitutional factors with the determinants in higher levels. For example, there are clear physical and biological reasons why very young and very old people are more susceptible to disease and injury than adults of working age. The report also considers the changes in people's living conditions as they age, in particular:

- **Housing**: young households (where the oldest member is 16–24) are most likely to live in poor housing. Households with residents over 75 are also likely to live in poor housing
- **Employment**: 16–19 year olds have the highest unemployment rates
- **Crime**: young households are most likely to be burgled

Poor socioeconomic conditions are likely to affect old and young people in different ways: while young people may be physically robust enough to withstand the health risks of poor living conditions, older people are more susceptible.

Reproduced from the London Health Commission (2002).

EVANS AND STODDART (1990): THE HEALTH FIELD MODEL

In this health field model, health is explicitly conceptualised as more than the presence or absence of disease to include **functional status** and **wellbeing**. By setting out a **relationship** between determinants, the health field model helps practitioners to understand how the determinants are themselves influenced, and therefore how they might be modified (see Figure 2H.2.3).

Evans and Stoddart's framework of health fields encompasses a range of factors, including those shown in Box 2H.2.3.

Box 2H.2.3

Social environment	Education, employment, family, poverty
Physical environment	Poor housing, proximity to hazards, waste and conflict
Genetics	Genetic factors that interact with environmental conditions
Behaviour	Viewed as an '*intermediate*' determinant (i.e. not simply a voluntary act), behaviour is shaped by a range of determinants, including education, access to facilities and financial considerations
Health care	Another '*intermediate*' determinant, this encompasses access and quality of health care

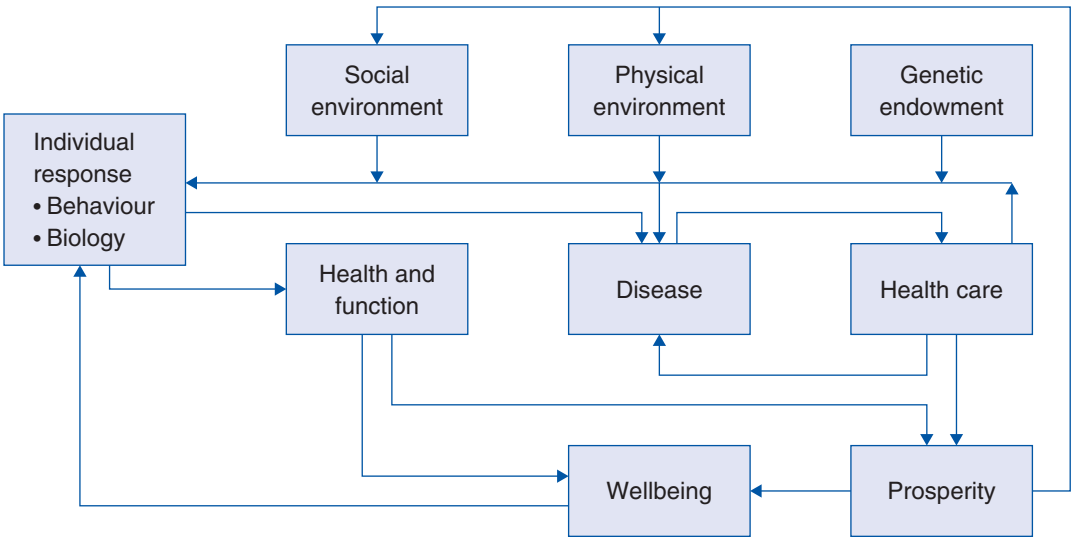


Figure 2H.2.3 The health field model. *Reproduced with permission from Evans and Stoddart (1990)*

SOCIAL DETERMINANTS FRAMEWORKS

The links between social conditions and health have been thoroughly explored. Research has strongly linked health (considered as premature mortality, vulnerability to illness and injury, self-reported health and wellbeing) to **education** and **employment**. For example, the **Whitehall study** found that those employed in lower grades of the British civil service were more likely to die prematurely than those employed in higher grades. There are several models that place social conditions as the main determinant of health and health inequalities.

Diderichsen and Hallqvist (1998) devised one of the most commonly cited '*social determinants*' frameworks. The model identifies four broad conceptual mechanisms:

1. **Social stratification** (social conditions, such as education and employment, will determine people's social situation) leading to
2. **Differential exposure** and
3. **Differential vulnerability** which together result in
4. **Differential consequences**.

These mechanisms work synergistically to generate health inequities. For each mechanism, the possible entry points for policy interventions are identified (see Figure 2H.2.4).

Other models focusing on social determinants of health include explicit recognition that:

- Health influences social position as well as vice versa
- There are interacting effects between social and biological pathways
- Access to and quality of health care are related to social determinants.

Recently the focus for action has been on the **social inequalities** of health, which Dahlgren (quoted in an interview with Koller, 2006) defines as:

> *'Systematic differences in health between socioeconomic groups [which] are socially produced, modifiable and unfair.'*

This reflects growing disparity in life-expectancy and health status between socioeconomic groups, while overall the average health and life-expectancy continue to improve.

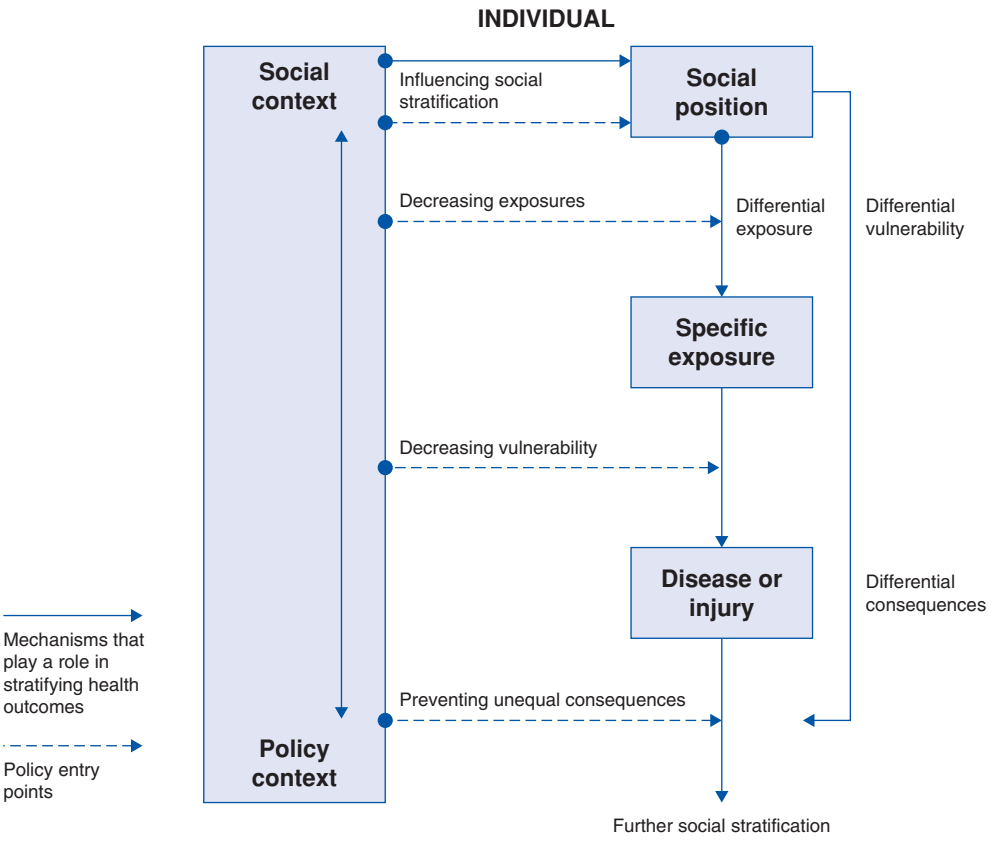


Figure 2H.2.4 Diderichsen and Hallqvist's social determinants framework. *Adapted from Diderichsen and Hallqvist (1998)*

2H.3 POLICY DILEMMAS

Ideological dilemmas and policy assumptions underlying different approaches to health promotion

While the aims of health-care treatment are usually relatively clear, the aims of health promotion may be contentious, and they determine the approaches used to improve health.

Health promotion can reach beyond the traditional boundaries of health into policy-making, personal choice and community development. In some contexts, the reach of health promotion into these spheres is more acceptable than others. The acceptable scope for health promotion will depend on a number of factors including:

- **Prevailing policy**: two contrasting health policy statements – *Health for All* (1986) and *Choosing Health* (2004) – exemplify how approaches to health promotion can be shaped by views regarding personal autonomy and the role of communities.
- **Socioeconomic circumstances** affect both the community's health needs and the resources available to promote health. Naidoo and Wills (1998) describe the dilemma of health promotion for people whose primary problem is economic poverty.
- The characteristics of the particular **health issue**, e.g. the epidemiology and perceptions of HIV and AIDS have dictated the focus of health promotion.

EFFECT OF POLICY: BALANCE OF INDIVIDUAL AUTONOMY, COMMUNITY INVOLVEMENT AND STATE INTERVENTION

The **Ottawa Charter (1986)** applied the WHO '*Health for All*' policy to the field of health promotion. The aims of the Charter implicitly assume that circumstances beyond the control of the individual are necessary to support health promotion, i.e. community involvement and state intervention are necessary. The Charter advocates that people involved in health promotion should:

- Create supportive **environments**
- Enable **community** participation
- Develop personal **skills** for health
- Reorient health-care services towards **prevention** and health promotion
- Build wide-ranging **public policy** that protects the environment and promotes health.

Reproduced from www.phac-aspc.gc.ca/ph-sp/phdd/docs/charter/index.html (The Ottawa Charter 1986).

Eng **Choosing Health (DH 2004)**: the Department of Health Public Health White Paper focuses on enabling **individuals** to make decisions about their own health. Here, the role of health promotion is to provide or facilitate the provision of information. Health promotion through legislation or community involvement is largely outside its scope (see Section 4C.3).

INDIVIDUAL ROLES

Personal roles have different requirements in terms of information, skills and involvement: see Box 2H.3.1.

Box 2H.3.1

Role	Needs and entitlements
Patients	Receive interventions only; do not need further input in the way that programmes are delivered
Consumers	Need sufficient information to make an informed choice about whether they accept or reject health-promoting behaviours
Empowered participants	Need adequate skills and resources to enable them to take part, and have a responsibility to fulfil their role in shaping and delivering health promotion

DILEMMA 1: TACKLE POVERTY OR ADDRESS NARROWER DETERMINANTS OF HEALTH?

An ideological dilemma of health promotion is whether its aim is to reduce inequalities in access to health services or to reduce inequalities in health (which would require wider actions outside the health remit).

The **Black Report** (Black et al 1980) argued that poverty was one of four possible explanations for the widening health inequalities seen in Britain in the twentieth century. See Sections 4A.8 and 4C.10.

POVERTY

Poverty (relative or absolute) is a major determinant of ill health. Low pay, inadequate benefits or unemployment lead to various types of poverty, namely lack of food and fuel, poor housing and transport, and social isolation. These in turn lead to:

- **Physical** problems (e.g. low-birth-weight babies, respiratory disease)
- **Psychological** problems (e.g. stress, depression, anxiety)
- **Behavioural** changes (e.g. smoking, low exercise, poor diet).

BARRIERS

Barriers to effective health promotion in poverty include:

- Intrinsic '**victim blaming**' culture
- Focusing on knowledge, attitudes and behaviour ignores the **constraints on choice** of healthy lifestyles that is associated with poverty
- Focusing in medicine on **one-to-one interventions** can ignore wider social influences, and underestimates the effect of poverty
- The lack of a **common approach** to poverty in health promotion
- Initiatives that may alleviate poverty (e.g. food cooperatives) are **outside the health remit**.

POLICIES

Policies to reduce poverty and improve ill health include:

- **Macro**-level changes (e.g. minimum wage)
- Collection of **data** on the wider social and economic determinants of health
- **Multisectoral action** at all levels (e.g. organisational: employment of benefits advisors in health clinics)
- **Community development** rather than simply health advice (e.g. food cooperatives to support healthy eating messages).

DILEMMA 2: TARGETED OR UNIVERSAL HEALTH PROMOTION?

Targeting offers an opportunity to **prioritise** resources and to **tailor** messages and activities to the particular characteristics of the people that need them the most. Targeting can be according to several factors. However, there are disadvantages associated with targeting health promotion, outlined in Table 2H.3.1. The problems associated with both targeting and universal health promotion approaches are illustrated with respect to HIV/AIDS in Box 2H.3.2.

Table 2H.3.1 Disadvantages of targeting health promotion

Targeted focus	Examples	Disadvantages
Behaviour	Smoking, eating, car driving	This risks **widening health inequalities** by appealing to those with the resources and circumstances available to change their behaviour. For example, advice to eat more fruit is easier to follow by people who can afford fruit and who live in an area where fruit can readily be purchased
Group	Children, older people, gay men	Assumes that groups are **homogeneous** (e.g. all gay men indulge in promiscuous, unprotected sex) Can lead to **culture blaming** (i.e. ascribing the risk to an inherent aspect of culture rather than a problem to be tackled jointly by the community and health). For example, the Asian rickets campaign focused on health education messages about increasing vitamin D in the diet, thereby placing the blame for vitamin D deficiency on Asian families. In contrast, when vitamin D deficiencies were first recognised in the 1950s as a general health problem, the response was to fortify core foods Groups that are deemed to be at risk of a certain conditions (e.g. sickle cell disease in Africans) may be **neglected** with regard to broader health problems (e.g. coronary heart disease)

Targeted focus	Examples	Disadvantages
Condition	Hypertension, diabetes	As part of the **prevention paradox**, many healthy people will be '**pathologised**' (e.g. hypertension will be diagnosed and treated in them, thereby causing unnecessary worry and side effects of treatment in people who will never become ill as a result of elevated blood pressure) In contrast, focusing only on people at high risk can **miss** many people who are at risk. For example, larger numbers of people who are normotensive die of coronary heart disease than do people who are hypertensive. Furthermore, concentrating on high-risk attributes while ignoring broader social influences may be less successful than whole-population initiatives

The alternative strategy is to attempt to reduce risk across the whole population (see Section 1C.16). Where a risk factor is widespread in society, universal approaches can have a larger effect than targeted interventions. For example, universal approaches to reducing dietary salt (e.g. lobbying manufacturers to reduce levels of salt in their products) can reduce average blood pressure and thereby reduce the burden of coronary heart disease. In this case, however, while the population's risk of disease may be substantially reduced, the average individual risk of disease is essentially unaffected. This is known as the **prevention paradox** and its recognition can cause the public to lose credibility in universal approaches (see Section 2H.4).

UK USA Box 2H.3.2

Example: Changing approaches with the changing perceptions and epidemiology of HIV/AIDS

When the HIV/AIDS epidemic began in the 1980s in the UK and the USA, health promotion messages were initially targeted at gay men. This led to:

- An illusion that heterosexual people were immune to HIV/AIDS
- Increased homophobia
- Difficulties in attracting funding and resources for what became seen as a 'marginal' illness

By the mid-1980s, the epidemic had changed and HIV was recognised as an infection that could potentially affect anyone. As a result, a whole-population approach for HIV/AIDS health promotion was adopted. However, this caused:

- A shift in resources from gay projects to professionally led mainstream interventions
- Explicit, detailed guidance on safer sexual behaviour that would previously have been unacceptable to the general public
- Change in messages from minimal behaviour change (sex, but safe sex) towards more conservative ones (monogamy)
- Ignored the fact that gay men were still disproportionately at risk
- Downgraded the strong activist/support networks that had been built up

The debate still continues with HIV about whether to focus resources on the groups primarily at risk, i.e. people from sub-Saharan Africa and men who have sex with men – or on the whole population.

Reproduced from Naidoo and Wills (1998).

2H.4 THE PREVENTION PARADOX

As described in Section 2H.3, public health practitioners often face a dilemma over whether to target people at greatest risk of a disease, or to lower the risk across the whole population.

The prevention paradox describes the effect whereby actions to reduce the risk of disease across the population successfully reduce the population's overall risk, but affects the outcome only for a minority. For example, the mandatory wearing of seat belts is a policy that affects all car users but prevents death only in the small minority involved in a road traffic crash.

> *'A preventive measure that brings large benefits to the community offers few benefits to each participating individual.'* (Rose 1981)

IMPLICATIONS FOR HEALTH PROMOTION

Because only a minority of people benefit directly from a population approach, the alternative approach of targeting health promotional activities at high-risk people may seem more attractive (see Dilemma 2, Section 2H.3). In other words, the prevention paradox can cause a loss of credibility in population-wide health promotion materials. For example, health promotion advice recommends reducing fat in the diet in order to minimise the risk of coronary heart disease. If it is evident that some individuals have a high-fat diet and do not develop heart disease, there is less incentive for others in the community to reduce their dietary fat. Hunt and Emslie (2001) describe what happens when health promotion materials do not take into account the prevention paradox:

> *'The failure to acknowledge the prevention paradox more directly in health education material thus can lead, at best, to greater mistrust among the general public of the messages contained, and at worst to their outright rejection.'*

2H.5 HEALTH EDUCATION

Health education and other methods of influencing personal lifestyles that affect health

Health education is sometimes erroneously taken to be synonymous with health promotion – the flawed assumption being that if people are sufficiently well informed about health then they will make healthy choices. Frameworks of the determinants of health illustrate that health is actually affected by more than simply an individual's personal choices (see Section 2H.2). This reality is echoed in models of health promotion, in particular Ewles and Simnett's (2003) five approaches to health promotion explicitly include other activities, such as social change, within the remit of health promotion (see Section 2H.7).

Education is a major component of health promotion: formal health education programmes usually aim to influence beliefs, attitudes and behaviours by enabling individuals to make an informed decision about those aspects of their life that affect their health.

Individuals do not receive their health education solely from planned, conventional, health promotion practitioners and materials: there is a range of other routes, such as friends, family, television and magazines, as illustrated by the example of children's knowledge of drugs and alcohol (see Box 2H.5.1).

2H.6 SETTINGS FOR HEALTH PROMOTION

Appropriate settings for health promotion (e.g. schools, the workplace)

Key locations for conducting health promotion include:

- Workplaces
- Schools
- Primary care.

As shown in Table 2H6.1, these settings offer the **structure** (e.g. physical building), **resources** (e.g. staff) and **access** to people (who spend considerable time in these settings) for health promotion to improve health and productivity. These benefits may improve the performance of the organisation as a whole.

Eng Box 2H.5.1

Example: Schoolchildren's knowledge and use of drugs and alcohol

In England, the National Curriculum requires schools to teach children about the effects of drugs and alcohol. This occurs at different stages of their education and in both science classes and in personal, social and health education (PSHE) lessons.

Regular surveys are conducted in England to explore secondary schoolchildren's smoking habits and their knowledge and use of drugs and alcohol. The 2004 survey included nearly 10 000 pupils aged 11–15.

Sources of information

The 2004 survey indicated that informal sources play as great, or a greater, role in where pupils receive their information about drugs and alcohol than does formal education. It concluded that:

> *'Pupils were most likely to say that they had received useful information on smoking, alcohol and drugs from the **television** (82%), their **parents** (75%) or **teachers** (72%). **Newspapers or magazines** (65%) and **friends** (52%) were also important sources. Pupils were least likely to have received useful information from the **government's information** and advice campaign, FRANK (15%) or from helplines (15%).'*

Effect of social deprivation

The survey also looked at reported rates of smoking and drug use according to two measures of social deprivation (namely receipt of free school meals, and few or no books in the home). Children living in higher social deprivation (according to either measure) were more likely to smoke or have taken drugs. This indicates that efforts to reduce smoking or drug taking among school-aged children need to take account of social deprivation as well as education and information.

Reproduced from the National Centre for Social Research and the National Foundation for Educational Research (2005).

Table 2H.6.1 Settings for health promotion

	Schools	Workplaces	Primary care
Target	Children and adolescents	Employees	Patients
Others	Parents	Employees' relatives and friends	Customers (e.g. in pharmacies), patients' relatives and friends
Potential hazards at the setting	**Physical:** • Exposure to unhealthy food **Psychological:** • Peer pressure • Stress from assessments and examinations • Bullying	**Biological:** • Exposure to toxic chemicals and fumes **Physical:** • Musculoskeletal symptoms lifting and handling • Sedentary lifestyle, e.g. office work **Psychological:** • Stress, bullying and harassment	Workplace hazards apply to staff

Table *contd overleaf*

Table 2H.6.1 *contd*

	Schools	Workplaces	Primary care
Potential benefits to the organisation and individual	**Performance:** • Better educational results **Wellbeing:** • Stronger links between school and home • Prevention or delay in risk-taking behaviours (e.g. drugs)	**Performance:** • Recruitment, morale, retention • Reduced absenteeism **Wellbeing:** • Exposure to healthy social norms (e.g. no-smoking workplaces) • Reduce hazards	**Performance:** • Overall cost savings • Reduced consultation rate • Reduced referral rates • Reduced emergency care
Resources available to support health promotion	**Staff:** • Teachers • School nurses **Facilities:** • Playing fields (if any) • Classrooms	**Employees:** • Occupational health • Health and safety officers • Managers • Workers **Partners:** • Unions **Facilities:** • Office space, financial	**Staff:** • GP • Practice nurses • Receptionists • Community nurses • Pharmacist • Dentist • Optometrist **Facilities:** • Practice/shop space
Examples of major policies or activities	**National Healthy Schools Programme:** • Personal, social and health education • Healthy eating • Physical activity • Emotional health and wellbeing **Ofsted:** • Inspections cover 'Healthy Schools' implementation as well as quality of teaching	**Health and Safety Executive** **COSHH** (see 2F) **Policies:** • Harassment • Smoking **Training:** • Health and safety • Manual handling **Subsidised or free:** • Counselling • Bike loans • Canteens with healthy menus • Gym membership • Vaccinations for at-risk workers (e.g. hepatitis B and influenza offered to health-care workers)	**General practice:** • Vaccinations • Lifestyle advice • Screening • Disease prevention treatment **Pharmacies:** • Display leaflets • Provide smoking cessation • Monitor blood pressure, cholesterol **Dentists and optometrists:** • Smoking advice

Advantages	Most children in school 'Captive audience' Children are in school for several years so changes can be tracked over time Opportunities for progressive programme	Protect health Access to healthy workers (who otherwise may not have access to health messages)	Patients seeking primary care may be more **receptive** to health messages Most people have **contact** with primary care at some point People are more receptive to health messages from senior health professionals
Limitations	**Work pressures** may affect implementation What about children **excluded** or away from school for other reasons? Staff (particularly non-health professionals) need to **develop skills** and confidence to deliver health messages	**Work pressures** may affect implementation **Priorities for employees** may differ from health at work (e.g. earning money) **Priorities for employers** may conflict with health priorities, particularly in **low-paid** or **illegal** work **Unemployed** people are not covered	**Short consultations** mean insufficient time to discuss health promotion If the intervention is not in the contract/not prioritised, then it may not take place Some staff not convinced of benefits or trained to deliver health promotion **Work pressures** may affect implementation

2H.7 MODELS OF HEALTH PROMOTION

Models of health promotion have been developed to:

- Define the scope and aims of the discipline (e.g. Tannahill 1985)
- Enable health promotion practitioners to understand what motivates individuals and/or communities to adopt health-seeking/harming behaviours.
- Inform the development of health promotion programmes to influence health behaviours.

The principal health promotion models are described in Table 2H.7.1.

TANNAHILL MODEL

Andrew Tannahill (1985) considered health promotion to be defined by 'three overlapping spheres of activity' (see Figure 2H.7.1):

- Health **education**
- **Protection** against harm and enhancing wellbeing
- **Prevention** of disease, disability and injury.

Table 2H.7.1 Health promotion models

Name	Author	Year	Brief description
Health belief model	Hochbaum et al	1958	Individuals will adopt health-related actions if they believe that they are faced with risk and have the potential to reduce the risk
Social learning theory	Bandura	1977	Behaviour is influenced by social norms, expectations, observations and perceived ability to control behaviour
Stages of change	Prochaska and DiClemente	1984	Individuals go through several stages to change behaviour
Modern theories	Tannahill	1985	Overlapping spheres of protection, prevention and education
	Beattie	1991	Four approaches covering a range of levels of authority and individuality
	Ewles and Simnet	2003	Multidisciplinary perspective, from biomedicine to sociopolitical change

Examples of the contents of areas of the overlapping spheres are shown in Box 2H.7.1.

Box 2H.7.1

Prevention	Organised preventive programmes of health care, e.g. immunisation
Health education	Health education to prevent disease onset, e.g. smoking cessation advice and information
Health protection	Health protection (legislation) to prevent illness or injury, e.g. fluoridation of water

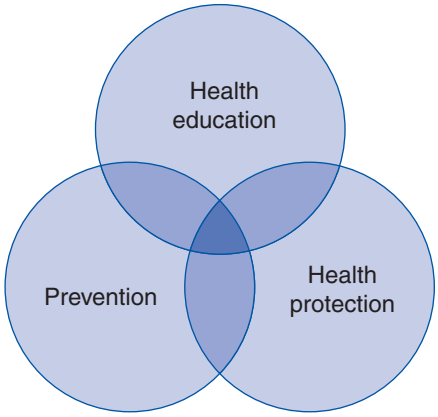


Figure 2H.7.1 Tannahill model of health promotion. *Reproduced from Tannahill (1985)*

The strengths and weaknesses of Tannahill's model are summarised in Table 2H.7.2.

Table 2H.7.2 Strengths and weaknesses of Tannahill's health promotion model

Strengths	Weaknesses
Simple to understand **Widely adopted** for defining what constitutes health promotion, and for informing health promotion practitioners how to plan and conduct their work Encompasses not simply the absence of disease but also the positive enhancement of wellbeing	Distinction between protection and prevention is **arbitrary** at times

BEATTIE MODEL

Beattie's model of health promotion (summarised in Naidoo and Wills 2000) considers not just the activities involved in health promotion but also how they are delivered, i.e. from the top down or from the bottom up. It is a useful tool for critically evaluating programmes, particularly regarding the balance of authoritative and negotiated approaches.

Beattie outlined four approaches to health promotion:

1. Health **persuasion**
2. Personal **counselling** for health
3. **Legislative** action
4. **Community development**.

Which approach should be used in a particular circumstance depends on the **mode** of the programme – ranging from authoritative (top down) to negotiated (bottom up) – and its **focus** (individual or collective), as shown in Figure 2H.7.2.

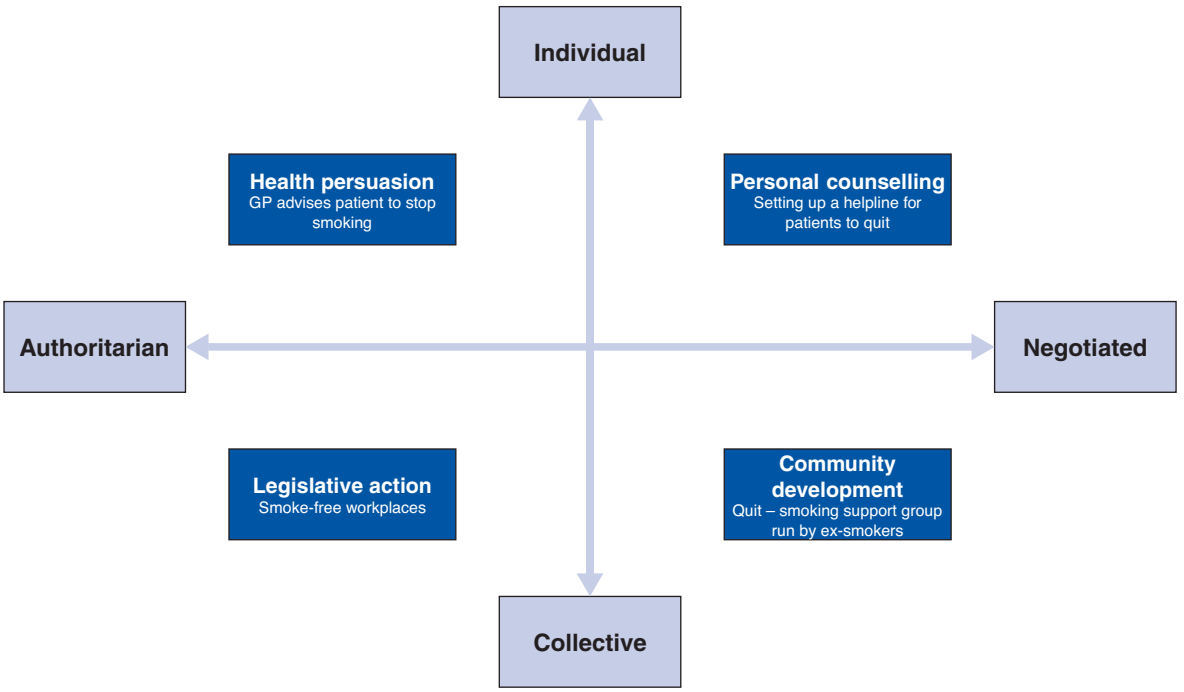


Figure 2H.7.2 Beattie's model of health promotion

EWLES AND SIMNETT

The model proposed by Ewles and Simnett (described in Ewles and Simnett 2003) considers health promotion from a multidisciplinary perspective. It explicitly incorporates biomedical approaches and activities in order to achieve policy and social change within the broader scope of the discipline. This model considers five approaches – the most appropriate combination of these depends on the starting point of the health promotion initiative (see Table 2H.7.3).

Table 2H.7.3 Ewles and Simnett's health promotion approaches

Approach	Description	Example
Medical	Focused on disease and biomedical explanations of health Narrow concept of disease (ignores social and environmental dimensions)	Immunisation Screening
Behavioural	Encourages individuals to adopt healthy behaviours	Healthy cooking classes
Educational	Provision of knowledge and information and assists development of skills for individuals to make informed decisions	Schools
Empowerment	Helps individuals to identify their own concerns and needs Health educator as facilitator	Community development work
Social change	Focus on socioeconomic environment in determining health Involves lobbying, policy planning, negotiating	Lobbying Policy planning Negotiating

HEALTH BELIEF MODEL (1958)

The health belief model (HBM) was first developed by American social psychologists Hochbaum, Rosenstock and Kegels in response to the failure of a free tuberculosis health-screening programme.

The HBM is based on the understanding that a person will take a health-related action if that person believes all three of the things shown in Box 2H.7.2.

Box 2H.7.2

Necessary beliefs

1. That a negative health condition can be avoided
2. That by taking a recommended **action, they will avoid a** negative health condition
3. Believe that they can successfully take a recommended health action

Moreover, the health belief model predicts that behaviour change requires individuals to believe all five of the things shown in Box 2H.7.3.

Box 2H.7.3

Susceptibility	They are susceptible to the condition or problem
Consequences	It could have potentially serious consequences
Course of action	A course of action is available to reduce the risks
Benefits outweigh costs	The benefits of the action outweigh the costs or barriers
Ability*	The individual perceives that they have the ability to carry out the action ('self-efficacy')

**Note that the final parameter,* self-efficacy, *was added later by Rosenstock et al (1988). This addition is thought to improve the way that the HBM meets the challenges of changing* habitual *unhealthy behaviours (e.g. being sedentary, smoking or overeating).*

The strengths and weaknesses of the HBM are listed in Box 2H.7.4.

Box 2H.7.4

Strengths	Weaknesses
Evidence supports the usefulness of the model in predicting behaviour or improving the effectiveness of interventions	Less useful for complex, long-term behaviours (e.g. alcohol dependence)
Most useful for traditional preventive behaviours (e.g. immunisations, health checks)	It does not account for other forces aside from individuals' beliefs that influence behaviour (e.g. socioeconomic circumstances or access to health care)

SOCIAL LEARNING THEORY

Also known as the **social cognitive theory** (developed by Bandura in the 1970s and 1980s – Bandura 1977), this model focuses on three influences on behaviour: see Box 2H.7.5.

Box 2H.7.5

Reciprocal determinism	The continuous, subtle and complex interactions between people's behaviour and their environment
Social norms	The effect of social and cultural conventions on behaviour
Cognitive factors	These encompass observational learning, expectations and self-efficacy

Observational learning is the concept that humans learn not just by doing (participation) but also by watching other people's behaviour and the rewards that others receive from their behaviours.

The term '**expectations**' is used to describe the capacity of a person to anticipate and value the outcomes of a behaviour. These vary between individuals, underlining the importance of exploring personal attitudes and beliefs when looking to change behaviour. For example, young women who believe that smoking helps with weight loss are more likely to be persuaded to give up if they are given information about other ways to control weight, rather than by warning them about the risks of smoking and lung disease.

Self-efficacy is a person's perceived ability to control their own behaviour. This is both person specific and environment specific. For example, someone may be very confident of their ability to avoid alcohol at home, but less so if in a social situation.

Strengths and weaknesses of social learning theory are listed in Box 2H.7.6.

Box 2H.7.6

Strengths	Weaknesses
Realistically complex solutions to health problems (i.e. not too simplistic) Unlike the *stages of change* and the *health belief* models, it explicitly recognises the impact of other factors: environmental, social and behavioural Widens the role of health promotion beyond individual persuasion about a discrete behaviour. Instead covers the entire social environment and wider personal beliefs	Can be difficult to implement because of its broad scope and complexity

STAGES OF CHANGE MODEL

This was developed by Prochaska and DiClemente (1984) over a number of years and is also known as the **trans-theoretical model**. It describes behaviour change as a process (not a one-off event) and it predicts that all individuals who change their behaviour will go through the stages shown in Box 2H.7.7.

Box 2H.7.7

Stages of change

- Pre-contemplation
- Contemplation
- Determination
- Action
- Maintenance
- (Termination)

People can enter or exit at any point – and they can stall at any stage. The model can be applied to people who initiate change themselves, as well to people in organised programmes. Programmes based on this model require initial understanding of the stage or stages at which people may enter the programme.

Strengths and weaknesses of this model are listed in Box 2H.7.8.

Box 2H.7.8

Strengths	Weaknesses
Useful for long-term, **complex** behaviour changes (e.g. giving up smoking, weight management)	Less useful for programmes aimed at **whole communities**
Useful for practitioners who wish to **tailor their counselling** (and their expectations for change) according to the stage of the model in which the individual is currently located	
Useful for **programme planning**, to organise interventions sequentially, and to match interventions to stages of the population	

2H.8 RISK BEHAVIOUR

Risk behaviour in health and the effect of interventions in influencing health-related behaviour in professionals, patients and the public

A person's aversion or predilection to risky behaviour is influenced by several factors, including:

- **Familiarity** with the **outcome** of the risky behaviour
- Degree of **personal control** over the risk factor: in contrast to environmental risks, individuals tend to downplay personal risks. This is due to beliefs of personal invulnerability and that other people are at greater risk *('It won't happen to me')*
- **Demographics** (age, gender and ethnicity) – young people are more likely to take risks (partly due to greater peer pressure) and women are more likely to be risk averse.

HEALTH RISK FACTORS

Factors known to be risks to health include those shown in Box 2H.8.1.

Box 2H.8.1

Harm	Smoking, cannabis use
Harm and benefits	Alcohol or food
Harm to others	Unprotected sexual intercourse or driving while intoxicated

RISK INTERVENTIONS

Interventions can be implemented at different levels:

- **Professionals** – to reorient services
- **Patients** – to receive preventive treatment, make lifestyle changes
- The **public** – to make healthy choices and protect or promote the health of society.

See also Section 2F.2.

2H.9 COMMUNICATION IN HEALTH EDUCATION

Theory and practice of communication with regard to health education

There are several requirements for health education to be effective; it needs:

- To be received
- To be understood
- To stimulate a change in **attitude**
- To provoke **behavioural change**.

The success of health education messages depends not simply on what is said, but on how it is said (including which media are used to communicate it).

HEALTH MESSAGES

As with all communication, there are four components of health messages: source, message, channel and receiver (see Figure 2H.9.1).

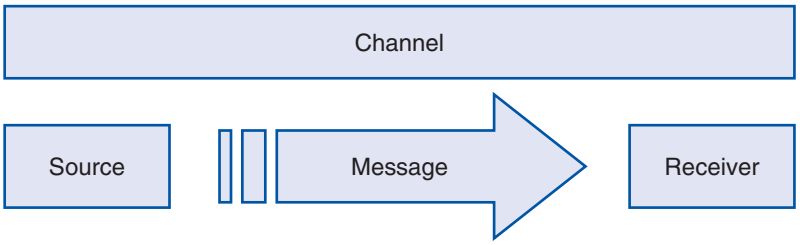


Figure 2H.9.1 Components of health messages

SOURCE

This is the person or organisation that generates the message. The credibility of a source depends on several factors, including:

- Source's **position in society**
- Training and **qualifications** of the source
- **Shared characteristics** with the recipient (e.g. age, culture)
- Perceived **conflict of interests**.

For example, a health advice message from a health minister may not be as credible as one from a doctor.

MESSAGE

A message may be **verbal** (written or spoken words) or **non-verbal** (images or sounds), and it may be horizontal or vertical (Box 2H.9.1)

Box 2H.9.1

Horizontal	General lifestyle improvement messages (e.g. eating a healthy diet)
Vertical	Specific issue messages (e.g. advice not to binge-drink)

Health promotion messages can be designed to **persuade** or to **empower**.

CHANNEL

This is the medium or media through which a message is conveyed. Channels include:

- One to one (e.g. midwife–patient consultation)
- Small groups (e.g. antenatal classes)
- Drama, storytelling or songs
- Mass media (broadcast, internet, newspapers, leaflets).

The most appropriate setting for the communication to take place will depend on the message. Examples include:

- Home
- Schools
- Community centres
- Workplaces.

Integrated marketing communication (IMC) achieves greater efficacy with:

- A communication mix of multiple channels
- Use of public relations, advertising and promotion.

RECEIVER

The target audience should be the prime consideration of any communication.

Messages may be targeted, for example, at:

- Individuals, families or communities
- Adults, adolescents or children
- Men or women.

SUCCESSFUL COMMUNICATION

In order for a communication to improve health it has to fulfil several requirements: see Table 2H.9.1.

Table 2H.9.1 Requirements for successful communication

Requirement	Description	Example
Be seen or heard	Will the target audience see/hear the message?	A poster for young people is displayed in places where young people go such as schools and colleges
Attract attention	Will the target audience notice the message?	A newspaper insert is sufficiently enticing for readers to notice it rather than throw it away unread
Be understood	Is the language intelligible to the target audience?	The wording and images on a leaflet are pre-tested on a sample of the intended audience to ensure that it is readable and unambiguous
Be accepted	Does the message reinforce current attitudes and beliefs?	Stop-smoking advertising is often effective in influencing smokers who have already decided that they wish to quit. It is less effective on those who do not want to give up
Change behaviour	Are all the factors that could prevent behaviour change being addressed?	In order to change their diet, people on low incomes may require messages about healthy nutrition to be supported by food subsidies

MEDIA

Different channels of communication may be suited to different aims and different settings. Often, however, they will be most effective when used in **combination**. Table 2H.9.2 compares some features of mass communication methods with small group methods.

Table 2H.9.2 Features of mass communication

	Mass media	Small group
Methods	Broadcast: TV, radio, cinema Print: newspapers, magazines, posters and leaflets Electronic: websites, email lists	Face to face: consultations, classes, groups At a distance: telephone, email, chat room
Scope	Many recipients	Few recipients
Flexibility	Low: one style produced for all recipients (electronic media excepted)	High (message can be tailored to the individual or group)
Feedback	Low	High – integral part of the process
Strengths	Useful for reinforcing attitudes and behaviours Simple, unambiguous messages (e.g. stop smoking) Ensures that same message reaches all recipients	Can be used to challenge current attitudes and behaviours Useful for conveying complex messages (e.g. benefits and risks of alcohol)

Table *contd overleaf*

Table 2H.9.2 *contd*

	Mass media	Small group
Weaknesses	Weak link between mass media production and the receipt of information (e.g. the information broadcast on TV may not be watched, and even if watched, it may not be listened to or understood) Limited scope to tailor information so that it is more suitable for particular individuals or groups Expensive: mass media may be unaffordable for local campaigns	Can reach only small numbers, so it is less well suited to messages that need to reach whole populations Low profile without the mass media Hard to control how information is disseminated by individual practitioners

2H.10 LEGISLATION AND HEALTH PROMOTION

Role of legislative, fiscal and other social policy measures in the promotion of health

Health promotion can make use of a range of levers at national and community level to create a supportive environment that encourages health-seeking behaviour. These include those shown in Box 2H.10.1.

Box 2H.10.1

Legislation	Description	Example
Social policy	Local, national or international health cultures and policies	Free-market or regulatory approaches to the economy
Restrictions and bans	Discourage behaviours known to have a damaging effect on health	Ban on children purchasing cigarettes
Fiscal measures	Systems of taxation to discourage certain behaviours, and subsidies to encourage others	Alchohol duty

SOCIAL POLICY

Social policy measures encompass a wide range of arrangements and structures designed to **increase harmonisation** in society – be it at an international, national or local level (Bunton 2002). While some policies will specifically be targeted at promoting health, other policies on a wide range of other issues will also influence health (see Section 2I.7).

A society's dominant **ideology** will influence its policy development, coupled with the degree to which state intervention is acceptable in that society. For example, cultures that prize the freedoms of the individual may be less inclined to promote policies of collective actions, such as banning smoking in enclosed public places.

BANS AND RESTRICTIONS

Restrictions or bans on activities or goods can take several forms, including the following.

AVAILABILITY

Eng **Cigarettes to young people**: in England the Children and Young Persons (Protection from Tobacco) Act 1991 makes it illegal to sell cigarettes to under 16s. This age limit was increased to 18 years in 2007 and is enforcable

by local authorities [www.opsi.gov.uk/ACTS/acts1991/Ukpga_19910023_en_1.htm Children and Young Persons (Protection from Tobacco) Act 1991 (c. 23)].

Eng **Alcohol at particular times of day**: in England, the Licensing Act 2003 has liberalised when alcohol can be sold so establishments can, potentially, sell alcohol 24 h a day

UK **Drugs/medicines:** the Medicines Act 1968 controls the availability of medicines – ranging from drugs that are on general sale to those that are available only on prescription.

USAGE

UK **Smoke free public places**. In Ireland this became law in 2004. Similar legislation was enacted in Scotland in 2006, Wales, Northern Ireland and England in 2007 (www.ash.org.uk).

SALES

UK **Illicit drugs**. The Misuse of Drugs Act 1971 designates controlled drugs (both 'medicinal' drugs and those with no known therapeutic benefit) into three classes – A, B or C – with corresponding restrictions on their availability and penalties for selling or being in possession. The most severe sentences are for class A drugs and the least severe for class C.

ADVERTISING AND INFORMATION

Eng **Tobacco warnings**. Compulsory labelling of cigarette packets with health warnings is stipulated in the Tobacco Advertising and Promotion (Point of Sale) Regulations 2004, which also bans tobacco advertising except at the point of sale (www.opsi.gov.uk/si/si2004/20040765.htm).

These measures often require **legislation**, although voluntary industry codes of practice also exist (e.g. major soft drinks retailers have voluntarily pledged not to target adverts at children). Restrictions and bans are only successful if they are supported by favourable public opinion or by enforcement.

FISCAL MEASURES

Fiscal measures generally take place at a national level, and can involve either **taxes** or **subsidies**. Such measures alter the price of goods to reflect the **externalities** associated with the particular action or goods. An externality is a cost or benefit from consumption that falls on someone other than the consumer (e.g. herd immunity is a positive externality of immunisation).

TAXATION

Taxation can serve two purposes:

1. To **raise revenue**
2. To **decrease demand** by increasing the price of the good to the consumer (or sometimes to the producer, e.g. the *polluter pays* principle).

In the UK, taxation to influence health is largely limited to alcohol and tobacco. Its success depends on the **price elasticity** of the goods (i.e. by how much the demand for a product is affected by its price). The price elasticity of goods will be different for different groups within society. So, for example, UK tax rises on tobacco products have had the greatest effects on:

- Young people
- Those on lower incomes.

Specific fees for particular activities can also be used to increase the price of producing or consuming goods (e.g. the **congestion charge** levied for driving into central London during busy times). The arguments for and against taxation as a health promotion measure are illustrated with respect to fatty foods in Box 2H.10.2.

SUBSIDIES

Subsidies decrease the price of consumption of goods to the consumer in order to adjust for the **positive** (i.e. beneficial) **externalities** associated with consuming them. Examples in health policy in England include the *Healthy Start Scheme* (previously known as the Welfare Food Scheme) where families with young children receive vouchers for liquid milk, infant formula milk or fresh fruit and vegetables so as to encourage the consumption of these foods among children.

Note that there may be **unintentional health effects** associated with subsidies. For example, the Common Agricultural Policy has subsidised tobacco farmers across Europe in a way that has encouraged tobacco production. This subsidy is being gradually phased out, and in the short term it is being revised so that subsidies are no longer linked to the amount of tobacco that is produced.

UK Box 2H.10.2

Example: Should fatty foods be taxed?

In January 2000, the *British Medical Journal* featured an article exploring the merits of imposing value-added tax (VAT) on a range of common foods high in saturated fat. The revenue generated would be used to compensate lower income families who would be hardest hit by the policy. The arguments for taxing high fat foods include:

- There are **precedents** for taxation to influence health, e.g. increasing tobacco duties
- Diet is partly responsible for **ischaemic heart disease**
- The costs of consumption therefore are partly borne by **health services**
- Increasing price could decrease the purchase and consumption of high-fat foods (assuming that these goods *are* **price elastic**)
- The effects would be greatest on those on low income, who would be most sensitive to price changes and who are at higher risk of heart disease

In contrast to the taxation of tobacco, however, the argument for foods was less persuasive because the relationship between fats and heart disease is complex. Some fats are necessary and beneficial, and people are affected by fat to a greater or lesser extent due to their genetic make-up. Since the article was published, the prevalence of obesity has increased in the UK and it has become a key public health problem. As a result, taxation of foods that could increase the risk of obesity is being reconsidered.

Adapted from Marshall (2000) and Marshall et al (2000).

2H.11 PROGRAMMES OF HEALTH PROMOTION

Methods of development and implementation of health promotion programmes

Health promotion programmes can be developed at local, national and international levels and therefore vary greatly in their scope and aims. However, all health promotion programmes should be based on the aspects shown in Box 2H.11.1.

Box 2H.11.1

Evidence	Available evidence of effectiveness: evidence-based practice (see Sections 1A.35 and 1A.34)
Theory	Appropriate health promotion theory and other scientific disciplines
Need	Local need (see Sections 1C.1 and 1C.2)
Resources	Resources available locally, in terms of funding and human resources
Priorities	Local and national priorities

APPROACHES TO HEALTH PROMOTION

Programmes often include a combination of the following approaches:

- Changes in **policy** to shift culture and/or behaviour (see Section 2H.10)
- Distribution of **resources** to provide incentives or remove barriers to change
- **Community development** (see Section 2H.12)
- **Information**, communication and education to inform people of the risks and benefits associated with behaviours, and to influence attitudes towards change (see Section 2H.9).

IMPLEMENTATION

Front-line clinical staff and members of the community conduct far more health promotion than do health promotion or public health staff. **Implementation** therefore requires working with diverse practitioners to ensure that they are equipped to carry out programmes. Some considerations are set out in Table 2H.11.1 and the issues in practice are described with reference to smoking cessation in Box 2H.11.2.

Table 2H.11.1 Issues affecting implementation of health promotion

Values	Do those who are involved have the **cultural** and **professional values** to support the programme's aims? For example, if members of the community are opposed to sex outside of marriage, then they will be unlikely to publicise or support community-based testing for sexually transmitted infections in young people
Motivation	What is the **motivation** for staff not connected with health improvement to change their practice to improve health? This could involve providing incentives (e.g. financial bonuses related to programme implementation)
Guidance	Is it clear from the **policies** and **guidance** what exactly is required of staff and community members? People may support a programme in principle but, if their role is unclear, then the programme may not be implemented successfully
Skills	Do those involved have the relevant **skills** and **competencies** to carry out their roles? If staff are required to perform tasks that they have not conducted before, then they may need additional training
Time	Is there **enough time** to put changes in place? While consultations in general practice could be an ideal opportunity to discuss behaviour changes with patients, it may not be feasible to do this in a 10-min consultation (or shorter)

In order to ensure that a programme has been implemented effectively, it should be **evaluated** during or after the programme's completion (see Section 2H.14).

UK Box 2H.11.2

Example: Reducing smoking among NHS staff

In an NHS trust, a programme to reduce smoking among staff involved:

- Responding to **national policy** and targets regarding smoking
- Staff **surveys** to find out the smoking prevalence and the attitudes of smokers towards giving up
- **Research** into guidelines, evidence of effectiveness and case studies of other organisations' experience
- Application of relevant health promotion **theory** (e.g. *stages of change* model)
- Consideration of the **resources** available to support the programme

The implementation comprised the following steps:

- Creation of a **policy** regarding smoking at work, with staff involvement to ensure that the policy was workable and that staff felt some ownership of it
- **Dissemination** of the smoking policy to ensure that staff were aware of whether they could smoke at work, and the support available to help them stop
- This was followed by **monitoring** and **enforcement** of the policy, e.g. recording occasions where the policy was not adhered to, and disciplinary action for staff smoking in no-smoking areas
- Ensuring a strong **partnership** with the local NHS stop-smoking service, with time being made available to enable smokers to attend support groups
- Sufficient **incentives** for staff to give up (e.g. free or subsidised nicotine replacement therapy, stop-smoking support groups and one-to-one help)
- Recruitment and training of sufficient smoking cessation **advisers** at the trust to support staff giving up

2H.12 COMMUNITY DEVELOPMENT

Building stronger communities is a key strategy for **health improvement** and for **reducing inequalities**. It is fundamental to building healthy environments and providing individuals with the social support to adopt and maintain health-seeking behaviours. Community development encompasses a range of activities to generate social networks and to, empower people to shape their local services and have an input into their community. It can stimulate innovative and creative solutions to problems that are not amenable to conventional health promotion programmes. Its challenges lie in evaluating success and in ensuring that those in greatest need of stronger communities actually benefit.

DEFINING THE COMMUNITY

A community is not a static entity. Rather, it consists of groups of people with a common characteristic at a particular time, most commonly:

- **Geographical** (e.g. housing estates, villages)
- **Social** (e.g. workers' groups, student unions, lesbian/gay communities)
- **Cultural** (e.g. religious, ethnic).

Terms such as **community development**, **community participation** and **community renewal** are often used interchangeably. In contrast to social planning and other initiatives aimed at changing communities, community development is a 'bottom-up' approach. Social planning would be described as 'top down'.

METHODS

Community development uses a combination of activities. Smithies and Adams (1990) suggest that there are five core activities of community development, described in Table 2H.12.1.

Table 2H.12.1 Activities of community development

Activity	Description	Examples
Formal participation	Formal participation in decision-making	Focus groups Consultation days
Community action	Priorities are developed by community groups	Lobbying Self-help activities
Facilitation	Health service employees promote community activities	Provision of meeting rooms and refreshments
Interface	Statutory services working closely with communities and community leaders	Consultation with local imams
Strategy	Strategic support from national initiatives	Neighbourhood renewal funds Local strategic partnerships

Projects may progress from one type of activity to another as they mature, as illustrated in the example in Box 2H.12.1.

Box 2H.12.1

Example: cooking skills in the community		
↓ More user led	Top-down target	Tackling obesity
	Formal participation	A decision by the local strategic partnership to fund training in cooking skills for local people following consultation with mothers who attend the local children's centre in a deprived area
	Facilitation	Local mothers attend cooking skills course funded by the PCT and local authority
	Community action	Trained mothers organise and provide training for other local mothers

ADVANTAGES AND CHALLENGES OF COMMUNITY DEVELOPMENT

The advantages and disadvantages of community development are outlined in Table 2H.12.2.

2H.13 PARTNERSHIPS

Working with people outside the field of public health is essential for delivering effective health promotion.

PARTNERSHIPS IN PUBLIC HEALTH

In the UK, Public Health is currently part of the NHS, itself both a provider and a commissioner of health services. However, health is influenced by factors outside the remit of health care. According to Lalonde's (1974) health field concept, health is determined by:

- Genetics
- Environment
- Lifestyle/behaviour
- Health care.

Table 2H.12.2 Advantages and disadvantages of community development

	Advantages	Challenges
Initiating projects	User led: can achieve better community support if based on community priorities	Resource intensive Time-consuming Can be difficult to secure funding (especially given the unknown outputs)
Goals	Can focus on root causes of ill health rather than simply lifestyle choice (e.g. organising a food cooperative rather than just providing dietary advice)	Long timescales: it may take years for health outcomes to appear and for communities to change
Evaluation/outputs	The process of enabling communities to participate is an end in itself Enhances self-esteem, confidence and control	Results are often intangible and unquantifiable
Communities involved	Can reach disadvantaged or excluded groups that conventional interventions would miss (see Section 2H.6)	There may be a conflict of accountability for community development workers: do they work for the community or for the statutory service?

Therefore, working with those actors (individuals and groups) who have influence over factors other than health care (e.g. local councils and businesses) can have an impact on health.

Other reasons for partnership work include:

- Avoiding duplication
- **Pooling resources** – funding, experience, contacts, information, skills
- **Political imperatives** – duty to work together (e.g. health and social care provided to patients discharged from hospital).

PARTNERS

Public health practitioners need to work with partners from a range of disciplines and organisations: see Box 2H.13.1.

Box 2H.13.1

Within organisation	Primary care directorate, modernisation team, clinical governance officers and community services
Within health	Acute hospitals, mental health hospitals, neighbouring health authorities
Within the community	Local authorities, education, police, voluntary sector, local businesses

TYPES OF PARTNERSHIP WORKING

Public health practitioners work in partnership in a number of ways: see Box 2H.13.2.

Box 2H.13.2

Type of partnership	Example
Statutory committees	Local strategic partnerships
Shared targets and monitoring	Local area agreements Community strategies
Joint projects	Production of joint reports (e.g. *Health in London* was jointly produced by the London Health Observatory, the Health Development Agency and the Greater London Authority
Specifically resourced initiatives	SureStart HealthySchools Teenage Pregnancy Partnerships
Shared posts	Eng Director of Public Health employed jointly by a PCT and a local authority
Shared budgets	Eng A cluster of PCTs may choose to implement one chlamydia screening programme across all partners' areas, with pooled budgets and pooled resources

BARRIERS AND CHALLENGES TO SUCCESSFUL PARTNERSHIP WORKING

Partnership working can be challenging and it can be easy to blame problems on the other members of the partnership. However, many of the difficulties involved in partnership working are due to a mismatch in the ways that the organisations or individuals within them work. Table 2H.13.1 outlines some of these factors and some questions that might help to identify why partnerships are not working effectively.

Table 2H.13.1 Challenges to partnership working

Aspirations	Do all partners want the same outcomes from a partnership? Are partners working to different (even conflicting) performance drivers? (e.g. NHS is driven by health outcomes whereas private industry is driven by profit) Are all partners clear about what the objectives of the partnership are?
Processes	Organisational cultures: do different organisations function in the same ways? Consider Handy's work (see Sections 5B.1 and 5B.3) Are there compatible systems for making decisions? Do some members need to take decisions back to in-house committees or boards? Can representatives authorise spending or actions at the meeting on behalf of their organisation? Team working: how do different members of the partnership work in a team? Consider Belbin's team roles (see Sections 5A.1)
Commitment	The opportunity to build up a relationship can be lost if there is no continuity of people attending meetings Who attends the meetings? Are they the right people? Do they attend regularly? Do the same people attend each meeting?

Table *contd overleaf*

Table 2H.13.1 *contd*

Influence	Are people from equal levels of seniority attending the meeting? As there may be no direct leverage over members of the partnership through line-management, how can it be ensured that group members deliver on what they promise?
Power	Big fish/little fish situations (e.g. statutory agencies versus user groups in the community) How is the agenda developed? Are all parties given plenty of time to contribute? Where are meetings held: always in the PCT? Or out in the community?

MEASURING SUCCESS

Partnerships, like other aspects of health promotion, should be evaluated to ensure that they are meeting their aims. There are a number of tools that could be used to do this but Donabedian's evaluation model (see Section 1C.9) is a simple approach for considering what a partnership has achieved. Examples of successful indicators of partnership under Donabedian's three categories are shown in Box 2H.13.3.

Box 2H.13.3

Indicator	Examples
Structure	Joint funding Joint posts
Process	Internal and partnership plans are aligned Meetings well attended
Outcome	Objectives and milestones are achieved Outputs (e.g. reports) are used and valued by the partners and by the local community

2H.14 EVALUATION

Evaluation of health promotion, public health or public policy interventions

The WHO Working Group on Evaluation in Health Promotion (2001) identifies the eight steps listed in Box 2H.14.1.

Box 2H.14.1

1. Describe
2. Identify
3. Design
4. Collect
5. Analyse
6. Recommend
7. Disseminate
8. Use

DESCRIBE THE PROGRAMME

Clarify the initiative's mandate, timeframe, aims and objectives. Create a logical model showing the activities, outputs, impacts, effects, objectives and goals.

Assemble an evaluation group

- Engage the stakeholders
- Clarify the purpose of the evaluation
- Identify key questions
- Identify evaluation resources.

IDENTIFY THE ISSUES AND QUESTIONS FOR THE EVALUATION TO ADDRESS

Consider the purpose of the evaluation:

- **Formative** (for an ongoing programme): to maximise the effectiveness of a programme. Provides ongoing feedback and improvement
- **Summative** (once a programme has finished or reached a particular point): to decide to what extent the programme met its aims
- **Persuasive**: if the aim of the evaluation was to evaluate a particular outcome.

DESIGN THE DATA COLLECTION

Choose measurement methods, considering which indicators will be used as criteria of success. Both **qualitative** and **quantitative** data may need to be collected.

COLLECT THE DATA

Consider: confidentiality, validity, reliability of sources and information produced.

ANALYSE AND INTERPRET THE DATA

Consider:

- Comparisons with similar programmes elsewhere
- Differences between expected and observed results
- Strengths and limitations of the data (e.g. statistical issues such as chance and biases)
- Different interpretations are possible for any set of data
- Involve stakeholders to consider the implications of the findings.

MAKE RECOMMENDATIONS

Identify the costs and benefits of implementing recommendations *and* the costs of ignoring them. Involving stakeholders in generating recommendations means that they will *'already be committed to acting on the findings and receptive to the results'*.

DISSEMINATE THE FINDINGS

Key audiences include:

- Funders
- Practitioners considering implementation of a programme elsewhere
- Stakeholders of the original programme.

ACTION: USE THE FINDINGS AND RECOMMENDATIONS

Change ongoing programmes in light of the recommendations. Where changes are required in existing programmes, change-management techniques should be used (see Section 5C.2).

2H.15 INTERNATIONAL INITIATIVES

International initiatives in health promotion

The principal actors in international health promotion have changed since the middle of the twentieth century, when there were relatively few agencies working across countries. The WHO (a subsidiary of the United Nations) had the greatest role and was largely uncontested in its activities. As outlined by Walt (2001) and Lincoln and Nutbeam (2005), today international health promotion is characterised by:

- A smaller, more contested role for the WHO
- A range of other UN organisations with a remit for health (such as UNICEF, the UN Population Fund and the World Bank – which is currently the largest donor for health projects)
- Increasing activity from non-governmental organisations (NGOs), particularly from the private sector
- Bilateral activity between two countries' governments (e.g. host country and the UK Department for International Development, DFID)
- Resurgent interest in 'vertical' programmes that focus on discrete diseases (e.g. the Gates Foundation's activity focuses on particular diseases).

TIMELINE

Some key health promotion landmarks are summarised in Table 2H.15.1. Box 2H.15.1 provides examples of the application of other major global health policy developments.

Table 2H.15.1 Key health promotion landmarks

Name	Setting	Year	Description
Lalonde Report	Canada	1974	Commissioned by the then Canadian Health Minister, Marc Lalonde, and generally regarded as a key turning point in health promotion. Established the **health field** concept of health (see Section 2H.2)
Health for All: Declaration of Alma-Ata	Alma-Ata (Kazakhstan)	1978	WHO stated the aim of providing universally accessible primary care, which included health education
Health for All 2000: Ottawa Charter	Canada	1986	WHO established the core **principles of health promotion** internationally
Millennium Development Goals	UN	2000	**Eight targets** aimed at improving the lives of the world's poorest peoples
Global Fund	International	2000	Global partnership among governments, civil society, the private sector and affected communities to combat **AIDS, tuberculosis and malaria**
Global Strategy	WHO	2004	Global Strategy on **diet, physical activity and health** provides member states with a range of global policy options to address the growing problems of unhealthy diet and physical inactivity
Bangkok Charter	Thailand	2005	WHO recognised the growing burden of communicable and non-communicable diseases and called for greater coherence across governments, international organisations and civil society

Box 2H.15.1

Examples:

Healthy Cities (1988)

Launched by WHO Europe to support cities in prioritising health improvement, and to provide resources and guidance for cities to improve health. Over 1000 European cities now participate, linked through national and international networks. The programme sets priorities for 5-year periods. A recent phase had three core themes: **healthy ageing**, healthy **urban planning** and **health impact** assessment. In addition, all participating cities focus on the topic of physical activity.

Reproduced from Healthy Cities and urban governance (2006).

Framework Convention on Tobacco Control (2005)

This framework was the WHO's first treaty. It is a binding international legal instrument with broad commitments and governance for national governments regarding tobacco control. It sets international standards on a range of tobacco-related issues, including tobacco price and tax, tobacco advertising and sponsorship, labelling, illicit trade and second-hand tobacco smoke.

Reproduced from the WHO Framework convention on tobacco Control (2005).

2H.16 INTERNATIONAL HEALTH PROMOTION INITIATIVES

Opportunities for learning from international experience

International agencies such as the WHO are central for disseminating learning on key issues, by means of the methods shown in Box 2H.16.1.

Box 2H.16.1

Guidance and tool kits	Where practice and evidence exists
Consensus statements	Where evidence is not conclusive but advice is helpful
Expert networks	Where opinion at an international level is required (see Section 1C.13)

Other systems for sharing experience include international scholarships, study visits, conferences and the internet.

Note that there are limitations to learning from international experience that may preclude an approach that was successful in one country from being implemented or successful in another. These include:

- Different cultures
- Different health systems and funding of health care
- Demographic structures
- Varying disease epidemiology and health needs in different regions.

The example described in Box 2H.16.2 describes some of the problems in transferring lessons learnt about HIV health promotion from western Europe to central and eastern Europe.

Box 2H.16.2

Example: Transferring lessons learnt regarding HIV/AIDS prevention from western Europe to central and eastern Europe

The HIV epidemic has been relatively minor so far in central and eastern European countries, providing an opportunity for instituting preventive measures and for learning lessons from HIV prevention in western Europe. Wright (2005), however, highlights the fact that there are diversities within Europe that may affect the extent to which western European approaches can be used. One key issue is the cultural difference between some groups of men who have sex with men and the gay identity.

In western Europe, the central role of the gay community in leading health promotion to prevent the spread of HIV/AIDS was one of the 'success stories'. In eastern European countries, it cannot be assumed that there is an established gay culture among men who have sex with men. For example, legal restrictions on homosexuality are severe in parts of central and eastern Europe

Wright argues that the focus should be on:

'Assisting each country to adapt basic principles of HIV prevention to their current political and social situation.'

Reproduced from Wright (2005).

2I
Disease Prevention and Models of Behaviour Change

Following on closely from the theories explored in Section 2H, this chapter tackles the challenges, approaches and priorities for improving health in practice. **Preventive actions** are arguably most important in children and families, whose health can be particularly vulnerable to the effects of living in **deprivation**. However, as this chapter discusses, robust evaluation and research of effective strategies is often lacking.

The principal tools used to improve health and to prevent disease are discussed. These range from **social marketing** (a relatively new technique that is becoming increasingly widespread) to **target setting** (which has become integral to the delivery of modern health services).

2I.1 PREVENTION IN THE EARLY YEARS

Evaluation of preventive actions, including the evidence base for early interventions on children and families, support and emotional development

The goal of most public health interventions is to prevent disease and to maximise health. Interventions aimed at pregnant women, babies and young children are particularly attractive to public health practitioners because of their potential to influence health and wellbeing throughout life. Working with these groups holds some particular challenges, not least because the outcomes are often realised only decades into the future.

EVALUATION OF PREVENTIVE ACTIONS

> *'There is currently limited evidence about what works in terms of preventive and public health interventions, how effectively to implement them, and even less evidence on their impact on inequalities and the cost-effectiveness of these interventions.'* (Wanless 2004 p 110)

In commenting on the evaluation of preventive action, the Wanless report emphasised the need for sustained support and funding, and observed that well-controlled studies are rarely performed in this field.

Evaluation of public health interventions may be complicated by a number of factors: see Box 2I.1.1.

Box 2I.1.1

Study design	It has been argued that randomised controlled trials can be difficult to justify ethically for preventive studies. However, without a well-designed study, biases and confounders cannot be excluded. It is therefore difficult to identify whether the intervention was directly responsible for the averted disease
Lack of control group	It is not always possible to predict or model the course of events if a preventive action was not put into practice. For example, it is impossible to tell what the disease occurrence would have been had a vaccine programme not been delivered
Lead times	Long time delays between an intervention and the manifestation of its effects can require lengthy (and therefore costly and complex) studies, e.g. effect of early years' education on achievement at 18 years of age, or effect of smoking cessation on lifetime risk of cancer. Where effects are not observed in the short term, it can be difficult to ensure continued support for preventive actions

See also Sections 1C.9, 1C.16, 2H.4 and 2H.14.

THE EVIDENCE BASE

Studies that investigate the impact of early interventions on children and families may be grouped into five areas:

1. Education
2. Health and nutrition
3. Socioeconomic benefits
4. Emotional and social support
5. Combined programmes.

EDUCATION OF THE CHILD AND PARENTS

Pre-school education can improve children's social and intellectual development and long-term outcomes, as illustrated in the examples in Boxes 2I.1.2 and 2I.1.3. Education for parents can include parenting skills but may also involve ensuring that young parents are still able to access their own education.

HEALTH AND NUTRITION

Health promotion in the early years starts before birth with interventions to reduce the risk of **low-birthweight** babies. Low birthweight is associated with higher infant mortality but also long-term effects, such as a higher risk of chronic conditions in adult life (e.g. diabetes, heart disease). It is more common for parents of lower socioeconomic background to have babies with low birthweight. A Health Development Agency systematic review of published evidence identified two major modifiable factors that influence low birthweight and explored the outcomes of intervention to reduce these risks, namely:

- Poor **maternal nutrition** at conception and during pregnancy can led to low-birthweight babies. While there is evidence that **calcium** supplements and folate can be effective, there is little evidence about the effectiveness of other supplements.
- **Smoking during pregnancy** doubles the risk of having a low-birthweight baby. Formal smoking cessation programmes with nicotine replacement therapy can enable some pregnant women to give up smoking. However, other factors (e.g. partner's smoking status and mother's socioeconomic group) affect the success of such programmes.

EU Box 2.I1.2

Example: the Effective Provision of Pre-School Education (EPPE) Project

A cohort study of 3000 children across Europe considered the effects at age 6–7 of different forms of pre-school education, and the circumstances of the home learning environment, upon children's literacy, numeracy and social development (e.g. anxiety, antisocial behaviour and positive social behaviours), together with the impact on social inequalities. The findings included:

- Children who had attended pre-school (full or part time) showed **higher educational and social attainment** than those who had no pre-school experience – even when home and social circumstances were taken into account
- The **quality** of pre-school education influenced the educational levels achieved
- The **home environment** was important, but parents' socioeconomic status did not affect children's benefit
- The report concluded that *'what parents do is more important than who they are …. Children whose parents read to them, taught them letters and numbers, songs and nursery rhymes, and took them to the library had better outcomes at 6 and 7 years'.*

Reproduced from Sylva et al (2004).

The evidence indicates that **breastfeeding** can improve the health and wellbeing of infants and mothers. Breastfeeding is less common in lower socioeconomic groups and is becoming less widespread among certain minority ethnic groups that traditionally breastfed their babies. Prenatal support and education of both mothers and health-care staff have been shown to improve breastfeeding rates.

USA Box 2I.1.3

Example: Head Start – long-term effects of comprehensive child development in the early years

The Head Start programme is a pre-school education scheme run for disadvantaged families and children in the USA to reduce social, educational and health inequalities between children from disadvantaged backgrounds and their peers. As well as pre-school education, Head Start provides a range of other services, including:

- Facilitating use of medical care for children (e.g. immunisations and dental health)
- Provision of healthy food and snacks
- Encouragement of parents' involvement in their children's education

Since Head Start was started in the 1960s, it has been possible to evaluate its long-term outcomes. A large-scale survey of social, health and economic behaviours (called the Panel Survey of Income Dynamics) has been conducted in a cohort of 8000 families since 1968. In 1995, adults aged 18–30 were asked as part of the survey whether they participated in Head Start or other pre-schools as a child. The survey found that adults who had attended Head Start were more likely to **complete high school** and **attend college** than their siblings who attended other pre-schools.

The effects of the Head Start programme are not just restricted to educational achievement, but encompass **crime** and **economic benefits**: Head Start graduates from African–American origins were less likely to have been later charged or convicted of a crime than were their siblings who attended other pre-school programmes. Also, white graduates in their 20s who had attended Head Start earned more than comparator groups.

Reproduced from Administration for Children and Families (2006) Head Start General Information, available online at: www.acf.hhs.gov/programs/hsb/, Garces ET (2002), Fight Crime: Invest in Kids (2006) Head Start reduces crime and improves achievement, available online at: www.fightcrime.org/reports/HeadStartBrief.pdf.

In the early years, interventions to promote maternal and child nutrition include:

- **Education** (e.g. nutritional content of foods, cooking skills)
- **Subsidies** (e.g. free school meals, free fruit in schools)
- **Supplements** (e.g. fluoridation of water supplies, nutritional supplements to food staples, such as bread).

SOCIOECONOMIC BENEFITS

In the UK, the **Acheson Report** (an official enquiry commissioned by the government in 1997) reviewed evidence from a range of stakeholders and concluded that families with young children were at increased risk of poverty. Many of these families found themselves in a 'benefit-dependent poverty trap', i.e. they were unable to seek work because affordable childcare was unavailable. The enquiry recommended that poverty in young families should be reduced through:

- Provision of accessible and affordable **childcare**
- Increased **benefits** (and the **uptake of benefits**) to pregnant women and to families with young children.

EMOTIONAL/SOCIAL SUPPORT

Family support programmes can be based at:

- Community level (i.e. addressing poverty, social isolation and lack of community resources)
- Individual or small group level (e.g. home visiting during pregnancy and after birth).

The aims of such support are to:

- Provide parents with respite, problem-solving skills, capacity and wellbeing (e.g. decrease incidence of postnatal depression)
- Ensure physical, emotional and cognitive development in children
- Prevent child abuse.

Few studies have evaluated the long-term outcomes of family support. However, in 2005, the European Early Promotion Project (an ongoing cohort study of approximately 1000 families across Europe) identified that training health-care workers to support early parent–infant relationships leads to fewer psychosocial problems in young children.

COMBINED PROGRAMMES

These programmes include elements of the above four types of intervention and are often run by multi-agency teams. An example in England that developed in the late 1990s was the *SureStart* programme, now developed into children's centres.

An evaluation of this programme (Belsky et al 2006) reported that differences between areas with the Sure Start Local Programme (SSLP) and comparison areas were limited, small and varied in degree of social deprivation. SSLPs had beneficial effects on non-teenage mothers (better parenting, better social functioning in children) and adverse effects on children of teenage mothers (poorer social functioning) and children of single parents or parents who did not work (lower verbal ability). SSLPs led by health services were slightly more effective than other SSLPs.

They concluded that SSLPs seemed to benefit relatively less socially deprived parents and their children, but to have an adverse effect on the most disadvantaged children. Programmes led by health services seem to be more effective than programmes led by other agencies.

2I.2 PRE-DETERMINANTS OF HEALTH

Understanding of pre-determinants of health, including the effect of social cohesion on health outcomes

Pre-determinants of health are those factors that portend the determinants of health. However, the distinction between pre-determinants and determinants is variable, e.g. income can be seen both as a determinant of health or as a pre-determinant of a determinant such as housing.

See Section 2H.2.

Pre-determinants can be grouped in terms of individual and community material goods, policies and societal factors: see Box 2I.2.1.

Box 2I.2.1

Material	Policies	Society
Sufficient and healthy food	Minimum wage	Social cohesion: the extent to which a society is mutually supportive and minimises inequalities
Pure water	Health at work	
Clean air	Maternal services	Values and attitudes, e.g. the balance between competitive and cooperative approaches
Income	Childcare	
Housing	Benefits	Ethnic diversity and the tolerance of different cultures
Green spaces	General education	
		Languages

IMPACT OF PRE-DETERMINANTS OF HEALTH

Pre-determinants mediate changes in health through various routes, as described by Kahan (2005/2006), and shown in a visual form in Figure 2I.2.1.This is also illustrated in the example of social cohesion in improving health in Europe: see Box 2I.2.2.

EU Box 2I.2.2

Example: Social cohesion as the driver for improving pre-determinants of health is itself affected by health policies

The Council of Europe directs member states to consider social cohesion as '*an essential condition for democratic security and sustainable development*'. As such, its policies are designed to reduce inequalities and enhance active participation in the community.

The Council recognises not only that social cohesion influences the incidence of disease and death, but also that it can be influenced by health policy. For example, the funding of health care through private insurance systems may cause more inequalities than social insurance or tax-based systems. Hence, private insurance may adversely affect social cohesion.

Reproduced from Council of Europe (2006).

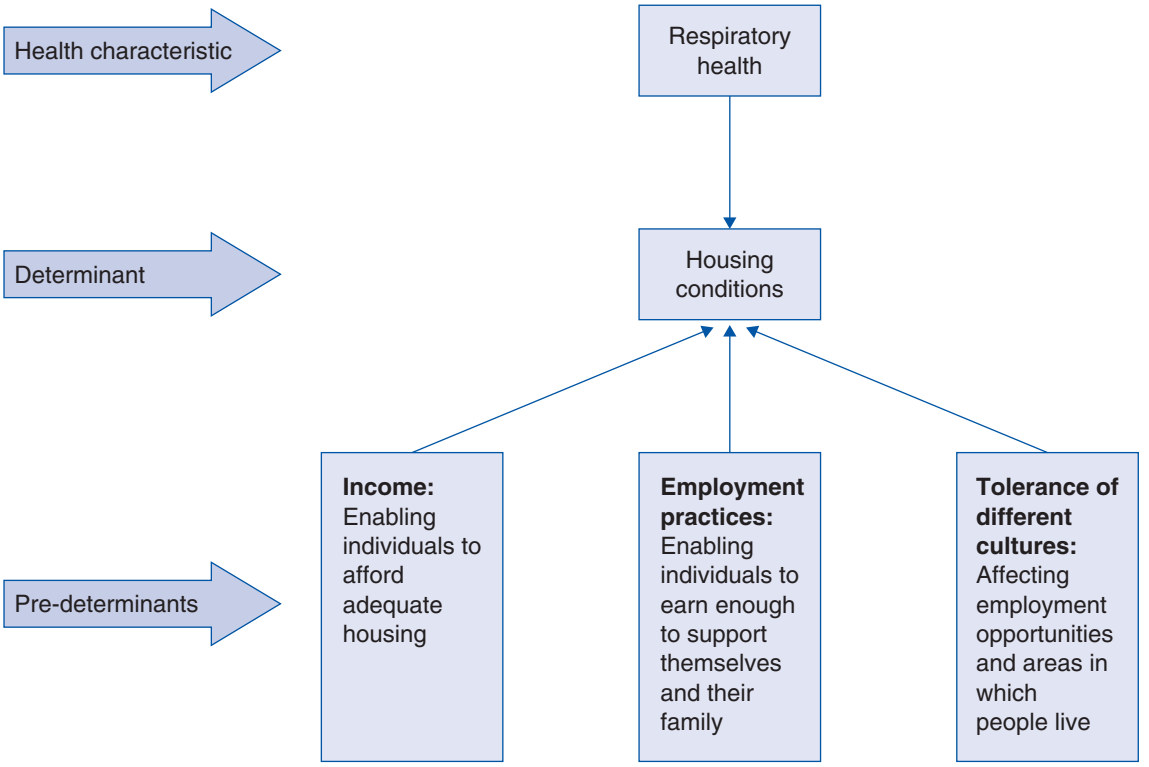


Figure 2I.2.1 Impact of pre-determinants of health

2I.3 SOCIAL MARKETING

See www.nsms.org.uk.

The term **social marketing** was first used by Kotler and Zaltman in 1971. It describes the use of techniques of commercial marketing to 'sell' a health message in order to benefit individuals and society. Social marketing approaches involve the six steps described in Table 2I.3.1.

Table 2I.3.1 The social marketing process

Identify the target group	Health approaches have traditionally segmented groups by risk/disease categories, e.g. diabetic patients, HIV+ve or sociodemographic characteristics (e.g. age group, income bracket). Marketing approaches use other tactics, e.g. psychographic, lifestyle characteristics and consumption patterns
Research target group	Surveys and focus groups are used to assess: • Attitudes and beliefs • Habits and lifestyles • Needs
Set objectives	Clear objectives are needed for the campaign, e.g. raise awareness or change behaviour
Develop the message	The message needs to be thoroughly pre-tested with target group to ensure that it is credible

Sell the message	Sell the message through a mixture of considerations (4 Ps, see Box 2I.3.1)
Evaluate	The campaign should be evaluated so as to monitor success and to guide refinements during the campaign

Selling a health promotion message, like the sale of any commodity, involves consideration of the product, its price, placement and promotion: see Box 2I.3.1.

Box 2I.3.1

Product	It must be clear what exactly is being 'sold'. For example, in a campaign to boost the proportion of children receiving MMR, the product could be the **procedure** (delivery of vaccine), the **service** (visit to nurse) or the **outcome** (immunity from measles). Each of these will have different appeals for different groups
Price	This is the relationship between the costs and benefits of the programme to the behaviour change. For example, in the case of vaccination, some parents do not see measles as a serious disease and therefore may not value the benefits of immunity. The price should be considered in economic terms, i.e. not simply the financial cost but also the **opportunity cost**
Place	The **channel** through which the message is communicated will affect who has exposure to it, and therefore levels of awareness of the message among the audience. The type of message will also affect which channels should be used
Promotion	This can be achieved through various media and advertising Marketing campaigns can include emails, mail-outs, text messaging, events, merchandising (e.g. red ribbon for AIDS) and partnerships with commercial companies and third sector organisations

Social marketing has a number of strengths and weaknesses, as shown in Box 2I.3.2.

Box 2I.3.2

Strengths	Weaknesses
Based on a clear understanding of the **target group**, not on the health promoter's perceptions of the group **Clear objectives** are integral to the approach Makes use of techniques that have been successfully applied **commercially**	Assumes that the individual is fully able to choose to change behaviour, i.e. that health is an **individual choice** and that socioeconomic barriers are not a factor in health choices As in commercial marketing, there is a danger of portraying only **partial information** in an effort to change behaviour, e.g. 'Just Say No' is catchier than a rounded picture of positives and harms of drug taking Danger of reinforcing the same **stereotypes** and attributes used by commercial marketing to sell products such as equating health with physical characteristics (e.g. youth, attractiveness, health, being thin) or with moral attributes (e.g. being in control)

2I.4 INVOLVING THE PUBLIC

Involvement of the general public in health programmes and their effect on health care

Reasons for including the public in health programmes are shown in Box 2I.4.1.

Box 2I.4.1

Improved treatment outcomes	Patients who are not informed or involved in treatment decisions are less likely to be concordant with treatment
Empowerment	The process of participation can empower individuals and communities to understand their own situations and to assume increased control over the factors affecting their lives. This process can, in turn, enhance people's sense of wellbeing and quality of life
Democracy	Community participation in decision-making, planning and action is a human right
Integrated approaches	Communities that are not restricted in their thinking by organisational boundaries can help to develop integrated, holistic and cross-cutting approaches to address complex issues
Better decisions	Involving people can result in more responsive, effective and appropriate services
Ownership and sustainability	Community participation is essential if interventions and programmes aimed at promoting health, wellbeing, quality of life and environmental protection are to be widely owned and sustainable

LEVELS OF INVOLVEMENT

The general public can be involved as patients or community members: see Box 2I.4.2.

Box 2I.4.2

Patients	Individual patients discussing their own treatment decisions with a practitioner
Community members	Geographical areas (e.g. housing estate, village) Age group (e.g. children) Condition-related groups (e.g. diabetes, mental health, self-help)

Brager and Specht (1973) described a **health ladder** that specifies levels of community involvement: see Table 2I.4.1.

A criticism of the health ladder approach is that it implies that organisations should be striving to reach the top of the ladder with the community. However, full involvement of the community may not always be feasible or desirable: often the appropriate level of involvement will be lower down the ladder.

PUBLIC INVOLVEMENT

Ways to involve the public in health-care programmes include working with individuals and groups. See also Section 2I.6.

WORKING WITH INDIVIDUALS

Individuals may be co-opted to work alongside professionals in health and social care partnerships.

Eng Local strategic partnerships (LSPs) are multi-agency bodies that function over an area of local government or social care. They provide a forum where local chief executives can meet to discuss interagency working, and to hold partnership structures in an area to account. Many partnership structures (e.g. prescribing committees, mental health fora, coronary heart disease networks, cancer networks) contain individuals who will represent themselves and their experiences to the professionals running the groups. Such people may not be typical of people with the

Table 2I.4.1 Brager and Specht's ladder of involvement

Control	Participants' action	Examples
High	Has control	Organisation asks a community to identify the problems and make all key decisions on goals and means; it is willing to help the community at each step to accomplish goals
	Has delegated authority	Organisation identifies and presents a problem to the community; it defines limits and asks the community to make a series of decisions that can be embodied in a plan that it will accept
	Plans jointly	Organisation presents tentative plans that are open to change by those who are affected; it expects to change its plans at least slightly and perhaps more subsequently
	Advises	Organisation presents a plan and invites questions; it is willing to change plans only if absolutely necessary
	Is consulted	Organisation tries to promote a plan; it seeks to develop support to facilitate acceptance
	Receives information	Organisation makes a plan and announces it; the community is informed and compliance is expected
Low	None	Community told nothing

same conditions and members will often need support from the chair of such groups to participate fully, but the perspective that service users offer is invaluable. The impact of children's involvement in the construction of Evelina Children's Hospital in London is described in Box 2I.4.3. The Expert Patient Programme in England also illustrates ways to involve patients in their own care: see Box 2I.4.4.

Eng Box 2I.4.3

Example: Evelina Children's Hospital

A children's board composed of children treated at the existing hospital and living nearby helped to design the new Evelina Hospital building. There is clear evidence of how children's views influenced the building design: children said that they did not want long straight corridors so instead the building has a curvy 'snake' floor plan.

Reproduced from the Evelina Children's Hospital Appeal, available online at www.evelinaappeal.org/hospital/index.html.

WORKING WITH GROUPS

Groups run by voluntary services, and community groups such as self-help groups, can be used to facilitate participation in:

- Advocacy activities
- Social networking through organising face-to-face consultation events or attending events held in the community.

Eng **Box 2I.4.4**

Example: Expert Patient Programme

People with chronic illness receive training from lay volunteers who themselves have chronic illness. The training helps them to understand their illness, develop coping skills and minimise the impact of symptoms.

2I.5 DEPRIVATION AND ITS EFFECT ON HEALTH

Concepts of deprivation and its effect on health of children and adults

Deprivation can be material or social in nature and can encompass all aspects of life. It is associated with:

- Lower life-expectancy
- Higher risk of tobacco, alcohol and drug dependence
- Higher chance of developing a long-term illness.

See Section 1C.8 for measures of deprivation.

CONCEPTS OF DEPRIVATION

Deprivation manifests itself in a number of ways (Box 2I.5.1), and the longer that people are exposed to deprivation, the greater its effects.

Box 2I.5.1

Area	Issue
Housing	Temporary accommodation, damp, overcrowded, poorly maintained housing
Environment	High levels of crime; poor access to facilities and transport
Income	Low income
Employment	Low status posts, hazardous work, job insecurity
Education	Lack of education during childhood and adolescence
Social exclusion	Poor social support, isolation or abusive relationships

Poverty can be absolute or relative: see Box 2I.5.2.

Box 2I.5.2

Type of poverty	Definition
Absolute poverty	Lacks the basic material necessities for life
Relative poverty	Lives on under 60% of the median national income

FACTORS THAT REINFORCE DEPRIVATION

A number of phenomena are known to result from and exacerbate the effects of deprivation: see Table 2I.5.1.

Table 2I.5.1 Factors that reinforce the effects of deprivation

Social exclusion	Relative poverty leads to exclusion from society, e.g. from decent housing
Discrimination	Racism, homophobia, ageism and discrimination against those who have received psychiatric treatment or been in prison or care all serve to exclude people from accessing services, and act as barriers to the opportunities open to other people (such as jobs, housing and social networks)
Employment	Unemployment and job insecurity lead to stress and health effects The degree of control and level of demand in a job both influence health There is a gradient across high-, middle- and low-ranking staff even when employed by the same employer (as shown by, for example, the Whitehall study)
Stress	The effects of *'continuing anxiety, insecurity, low self-esteem, social isolation and lack of control'* have both mental and physiological effects, such as increased risk of heart disease
Antenatal effects	Stress and lifestyle (e.g. smoking or ill health) during pregnancy have an effect on the fetus, which may manifest as fetal abnormalities and low birthweight Low birthweight is associated with increased risk in adult life of diseases such as diabetes and coronary heart disease (see also Sections 2A.1, 2B.1.5 and 2B.1.11)

2.I6 COMMUNITY DEVELOPMENT

Benefits and means of community development, including the roles and cultures of partner organisations such as local authorities

See Section 2H.12.

2I.7 HEALTH IMPACT ASSESSMENT

Health impact assessment of social and other policies

See Section 1C.17 for details of HIA. Box 2I.7.1 provides an example of the use of HIA on policies in London.

2I.8 STRATEGIC PARTNERSHIPS

Role of strategic partnerships and the added value of organisations working together

See Section 2H.13.

2I.9 SETTING TARGETS

Role of target setting, e.g. public service agreements, local authority agreements

See also Section 4D.4.

Reasons for setting targets in health care include:

- Adoption of management practices and culture in health services
- Improvement of performance and accountability
- Ensuring consistency across services.

UK Box 2I.7.1

Example: Health impact assessment of policies in London

The London Health Commission conducted rapid health impact assessments of the Mayor of London's draft strategies that aimed to improve the lives of Londoners. These strategies included policies on social issues (such as culture, children and young people). The health impact assessments took a largely pragmatic approach to meet time and resource constraints. They involved workshops with key stakeholders, written submissions, and the synthesis and review of available evidence.

An evaluation noted that the health impact assessments succeeded in:

- Influencing strategy: those drafting the strategy considered health because they knew that it would be subject to an HIA and they later made revisions as a result of HIA recommendations
- Involving a wider group of stakeholders than would otherwise have been involved in the policies
- Providing the evidence base for decision-makers to make choices
- Raising the profile of HIAs

Challenges included:

- Short timescales
- No established quality standards for HIAs
- Gaps in research evidence
- Involving the right stakeholders
- Relative priority of different types of research

Reproduced from HIA Gateway, available online at: www.nice.org.uk, London Health Commission, available online at: www.londonshealth.gov.uk/hia.htm#Top.

Targets can relate to each component of Donabedian's framework (i.e. structure, process, output or outcome – see Section 1C.6): see Box 2I.9.1.

Box 2I.9.1

Indicator	Example
Structure	Director of Public Health in post at 85% of PCTs
Process	80% of GP practices maintain a register of their diabetic patients
Output	98% of patients seen within 4 h of arrival in A&E
Outcome	Cancer deaths reduced by 20% by 2010 relative to 1990 baseline

Targets can be set at the micro, meso or macro level: see Box 2I.9.2.

CHARACTERISTICS OF 'GOOD' TARGETS:

'SMART' targets comply with the points in the mnemonic of Box 2I.9.3.

Box 2I.9.2

Level	Example
Organisational	Personal development plan (PDP) agreed with line manager
Local	Local area agreement (LAA) negotiated at council level and agreed with central government to ensure that local targets reflect the local situation
National	Public service agreement: in return for investment, the government sets minimum standards to be delivered across a range of public services, including health, education and crime prevention (HM Treasury 2007)
International	WHO's '*Health For All*' targets

Box 2I.9.3

S	**Specific**	Relate to what they want to achieve
M	**Measurable**	Defined indicators (quantitative if possible) to show if the target is met
A	**Achievable**	Achievable but challenging, in order to improve performance
R	**Relevant**	Relevant to current performance. For example, lung cancer deaths are mainly due to smoking practices 30 years ago rather than current health service interventions
T	**Timescales**	These must be defined in advance for each target

Rewards (or punishments) attached to meeting (or failing to meet) targets can lead to particular areas of health care receiving particular managerial interest. For example, smoking cessation and tobacco control are now discussed at board level in English PCTs as a result of the quitters' target. However, the high stakes involved (e.g. risk of job loss or attainment of foundation trust status) can lead to gaming and distortion of priorities.

ADVANTAGES AND DISADVANTAGES OF TARGETS

Strengths and weaknesses of targets are described in Table 2I.9.1.

Table 2I.9.4 Targets – strengths and weaknesses

Strengths	Weaknesses
• Provide a focus to performance improvement • Priorities are set explicitly • Create a level playing field for the organisations under comparison (e.g. hospitals, local authorities) • Number and domains of targets can be set to reflect organisational priorities and goals • Opportunity for sharing good practice and learning	• Often, information is unavailable to measure meaningful health outcomes • Distortions of practice or gaming. For example, some hospitals attempted to reclassify their A&E trolleys as beds in order to meet the waiting time target. Elsewhere, ambulance services were de-prioritising people in rural areas because they knew that they had no chance of meeting transport time target • Distort priorities – those areas of health care that are not amenable to targets are paid insufficient attention • Disengagement of clinicians if targets are externally set, or set from the top down

Section 3
HEALTH INFORMATION

Health information is essential for health service planning and evaluation. Without it, services would be unresponsive to changes in circumstances – and it would be impossible to make predictions or to increase efficiency.

There is a vast array of information available on populations, sickness and health. However, the quality and appropriateness of this information vary widely too.

Section 3 provides practitioners with an appreciation of what information is available, how data are collected, and the advantages and disadvantages of their use in public health.

3A Populations

Public health practitioners require an understanding of the population, e.g. calculating rates and risk ratios requires a viable denominator as well as a count of disease occurrences. The population size, **demography** and social characteristics in most developed countries are now enumerated by regular censuses. This chapter discusses the methods by which **census** information is obtained and, where this information is not available or sufficient, the methods by which **populations are estimated**. It also summarises **trends** in population structures across time, comparisons between different regions and approaches to address the health consequences of **population changes**.

3A.1 CONDUCT OF CENSUSES

A census is a snapshot **enumeration survey** of all the people in a country; as such, it provides the most complete set of population data. Its findings are of use within the health, housing, education and transport sectors for:

- Making **comparisons** between regions
- **Targeting** resources
- Planning future **resource allocations**.

Most countries hold a census every 5–10 years, often enshrined in the constitution. They can be conducted either by **interviewing** or by **self-enumeration** (with a post-enumeration survey to assess under-enumeration).

Questionnaires can be delivered and collected by the following means:

- Traditional 'drop-off' and 'pick-up'
- Post-out and post-back
- Internet.

UK UK CENSUSES

The first UK census was held in 1801. They have been held decennially since then, except during wartime. The 2001 census cost ≈ £250m. The Census Act makes completion of the census form compulsory: non-responders face a £1000 fine.

UK 2001 CENSUS

See Table 3A.1.1.

Table 3A.1.1 A summary of the conduct of the 2001 census

Preparation	Planning began with a **consultation exercise** There was then a parliamentary **debate** on the choice of questions, followed by a **publicity** exercise Census forms were designed for **self-completion** by the so-called 'head of the household'
Delivery	Census forms were **delivered** to each household by a *'field force'* The forms requested information about 'persons present' and 'usual residents temporarily absent' 88% of forms were **posted back** to temporary local offices and 6% **handed back** at the doorstep, with 4% **collected** by the field force
Content	Accommodation Relationships between people living in the household Demographics (age, sex, marital status) Migration Ethnicity (see Section 3A.4) Self-reported health Qualifications, employment, journey to work In Wales there was also a question about the Welsh language
Analysis	Completed forms were first **scanned** Scanned forms were then **coded** using automated and manual systems (a process that took approximately 1 year to complete) Paper forms were destroyed (with the scanned images being stored to be **made public** 100 years later) Data were **quality assured:** 2001 was branded the *'one number census'* since missing responses were estimated in order to adjust for under-enumeration Results were then tabulated into **databases** according to output areas and super-output areas (see Section 1A.18 for details), ready for analysis A **Census Output Prospectus** was published

UK CENSUS SMALL AREAS

Historically, census information was collected by **enumeration district**, and local data were generally analysed at the level of the **electoral ward**. However, enumeration districts and electoral wards varied greatly in size (the latter ranging from 100 residents to 30 000+) and were subject to frequent boundary changes – causing problems for longitudinal study.

For the 2001 census, output areas (OAs) were designed (using a geographical information system) that are:

- Co-terminous with postcodes
- Uniform in population size
- Compact in shape
- Socially homogeneous.

A new hierarchy of super-output areas (SOAs) groups OAs into units that are similar in population size and are highly stable. There are three levels of SOA, the lower SOA representing four to six OAs: see Box 3A.1.1.

Box 3A.1.1

Unit	Approximate population size
Output area	300
Lower SOA	1500
Middle SOA	7000
Upper SOA	25 000

COLLECTION IN OTHER COUNTRIES

HK In Hong Kong, censuses are conducted every 10 years, with by-censuses in the middle of the intercensal periods. The 2006 **by-census** involved 1 in 10 households. Data were collected by thorough interviews. The 2001 **population census** used two types of questionnaires:

- A short form (in six of seven) households on basic characteristics. Self-enumeration forms were mailed to the householders and collected by enumerators.
- A long form administered to one of seven households collecting data on a broad range of socioeconomic characteristics through face-to-face interviews.

The thematic household survey (THS), a further data collection tool, is used in Hong Kong. This provides information on health status, patterns of health service utilisation and health-care service expenditure profiles, and was performed in 1999, 2001, 2002 and 2005.

NZ Aus In New Zealand and Australia, censuses are administered using a similar approach to the UK:

- The Australian Bureau of Statistics conducts a nationwide census every 5 years
- Statistics NZ conducts a census every 5 years. The 2006 census allowed completion via the internet for the first time in this country.

HARD-TO-COUNT GROUPS

The following groups are typically difficult to enumerate:

- **Young, inner-city men**
- **Multiple-occupancy** buildings and student houses

- Those who do not speak the official language(s) of the country
- **Babies**
- Very **elderly** people
- **Military** personnel.

ALTERNATIVES TO TRADITIONAL CENSUSES

Between census years, planners must compromise between using either out-of-date results from the most recent census or more up-to-date results from less robust sources. The subsequent census often leads to dramatic revisions in statistics. For these reasons, some continental European countries are switching to:

- **Rolling censuses**
- **Population registers** with sample surveys (advantages and disadvantages of which are listed in Box 3A.1.2)
- **Population projections**.

Box 3A.1.2

Advantages of population registers	Disadvantages of population registers
More up-to-date results	Not a snapshot, so complicated to compare regions
Linked statistical database allows multivariate analysis	Loss of 'brand' ➔ lower response rates
Improved planning, provision and monitoring of surveys	➔ results have less impact
More consistent statistics	Depends on quality and availability of administrative data
Supports evidence-based policy	Actual risk of confidentiality breaches
Improved efficiency and quality from permanent systems	Perceived risk of confidentiality breaches

3A.2 COLLECTION OF ROUTINE AND *AD HOC* DATA

The raw data used by researchers and public health practitioners either can come from routine sources (see Section 1A.1) or may be specifically commissioned.

ROUTINE DATA

Databases of routinely collected data are a plentiful source of health information – often covering entire populations and spanning many years. Their advantages and disadvantages are listed in Box 3A.2.1.

Box 3A.2.1

Advantages	Disadvantages
Cheap	Limited to what is actually collected
Can be complete (e.g. register of births)	Difficult to assess quality control
Large numbers of subjects	Access sometimes restricted
Prospective, therefore avoid recall bias	Potential delays in publication
Particularly useful when different data sources are linked	Data linkage complex

AD HOC DATA

These data can be obtained either by commissioning a specific data collection exercise (e.g. a patient survey), or by requesting ad hoc extracts from a routine data source (e.g. from a cancer registry). Advantages and disadvantages of such data are listed in Box 3A.2.2.

Box 3A.2.2

Advantages	Disadvantages
Can specify exactly what data are to be collected	Sampling frame may be unknown
Can target data collection to the subgroup of interest	Potentially costly
Can collect qualitative data	May be difficult to link to routine data sources
Rapid data collection sometimes possible	Typically the number of subjects is small
Quality can be readily assessed	Data linkage is complex

3A.3 DEMOGRAPHY

This is the study of the characteristics and dynamics of human populations. Population change may arise because of any of the following four factors:

- Births
- Deaths
- Migration
- Ageing.

Table 3A.3.1 provides a summary of important demographic concepts.

Table 3A.3.1 Important demographic concepts

Concept	Description
Total population size	Measured at each census, this represents the historical trend in population growth or shrinkage
Age structure	Ages are divided into a series of bands (e.g. 0–4 years, 5–9 years), and the population is described according to the number of people in each band
Fertility	This is the number of offspring per female by age band Overall fertility rate = number of live offspring per 1000 per year in women aged 15–49 (Note: **fecundity** = number of offspring biologically possible per female)
Mortality	This is the count of the number of deaths per year by age bands. It is the calculated probability of a person in that age band dying each year: $= \frac{\text{(Number that died per year)} \times \text{(Length of age band)}}{\text{(Total number of individuals in each age band)}}$
Survival	Probability of survival = (1 – Probability of dying)

Once these parameters have been calculated, then population projections can be made (see Section 3A.7).

3A.4 MAJOR DEMOGRAPHIC DIFFERENCES

Important regional and international differences in populations, in respect of age, sex, occupation, social class, ethnicity and other characteristics

Some of the differences in demographics between different parts of the UK are shown in Table 3A.4.1.

UK **Table 3A.4.1** Differences in demographics between different parts of the UK

Constituent country	Number of people at 2001 census	Percentage of UK population (58 789 194)	Percentage population growth 1993–2003	Population density (number/km²)	Percentage of population non-white
England	49 138 831	84	+3.6	383	9
(London)	(7 172 036)	(12)	(+5.4)	(4700)	(29)
Northern Ireland	1 685 267	3	+4.1	125	1
Scotland	5 062 011	8	–0.7	65	2
Wales	2 903 085	5	+1.9	142	2

Reproduced from Office for National Statistics (2001 Census).

AGE

Age distributions can be represented graphically as **population pyramids**: see Figures 3A.4.1–3A.4.5.

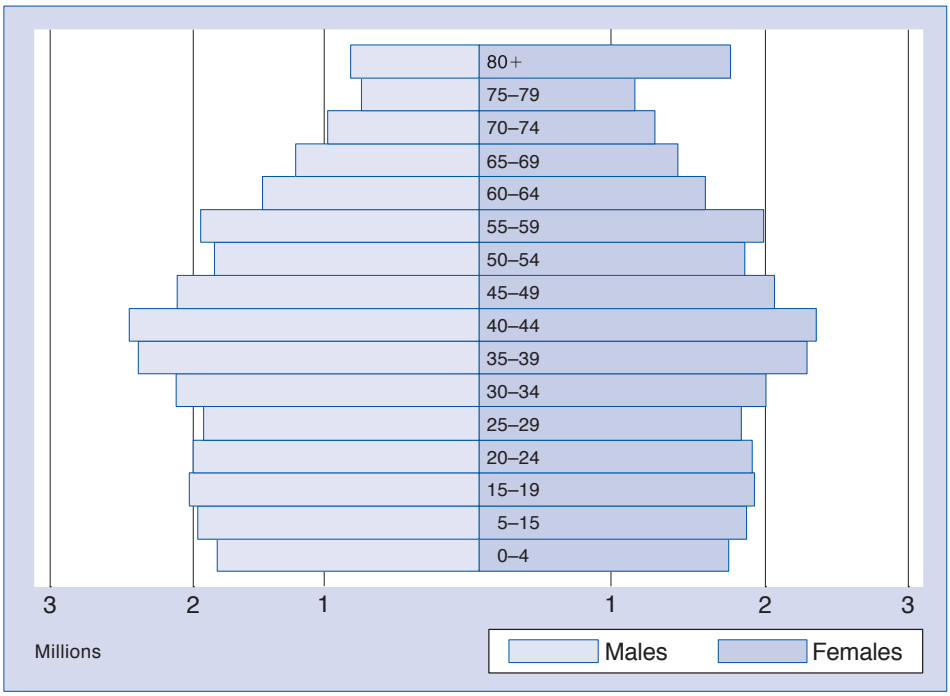


Figure 3A.4.1 Population pyramid – England 2005

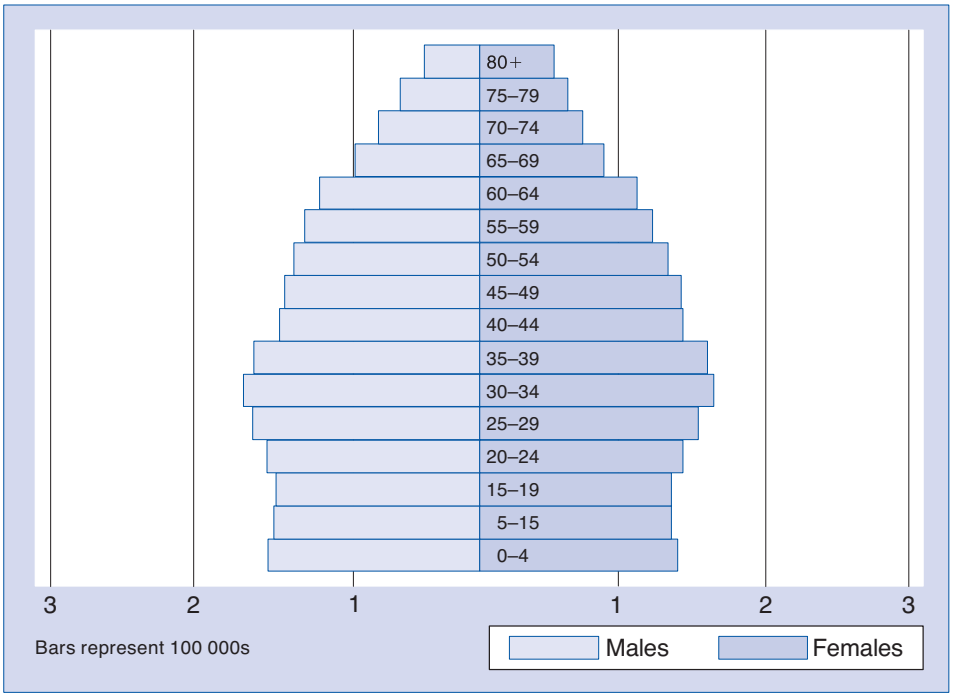


Figure 3A.4.2
Population pyramid – Ireland 2005

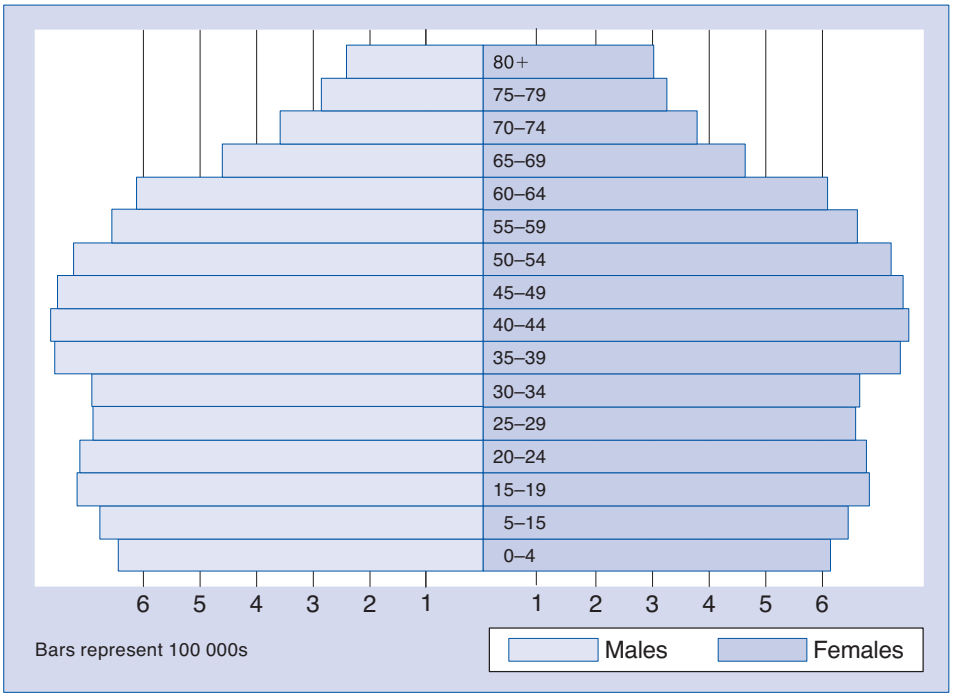


Figure 3A.4.3
Population pyramid – Australia 2005

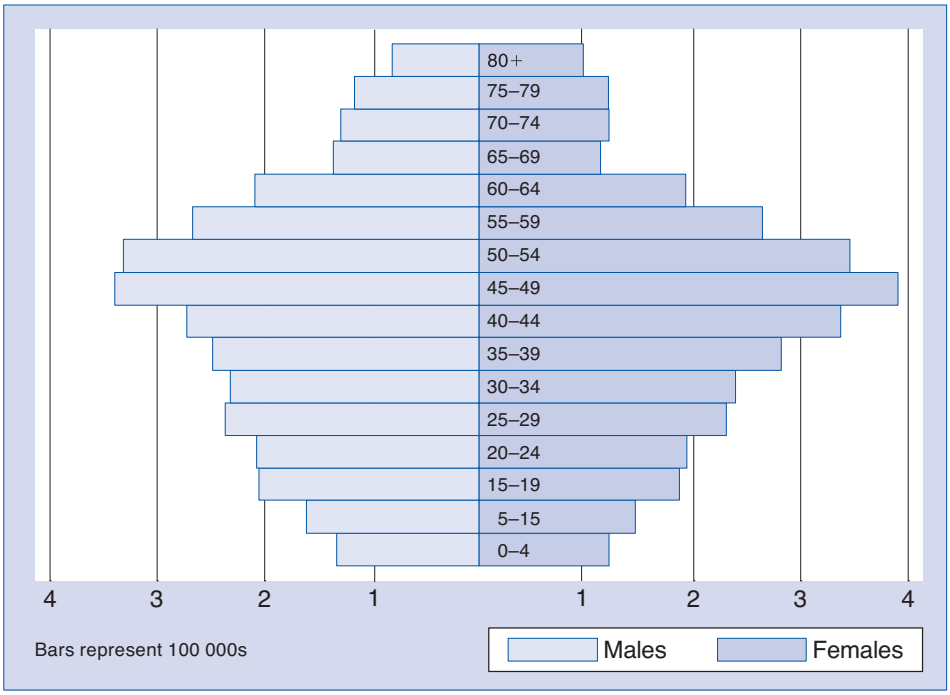


Figure 3A.4.4
Population pyramid – Hong Kong 2005

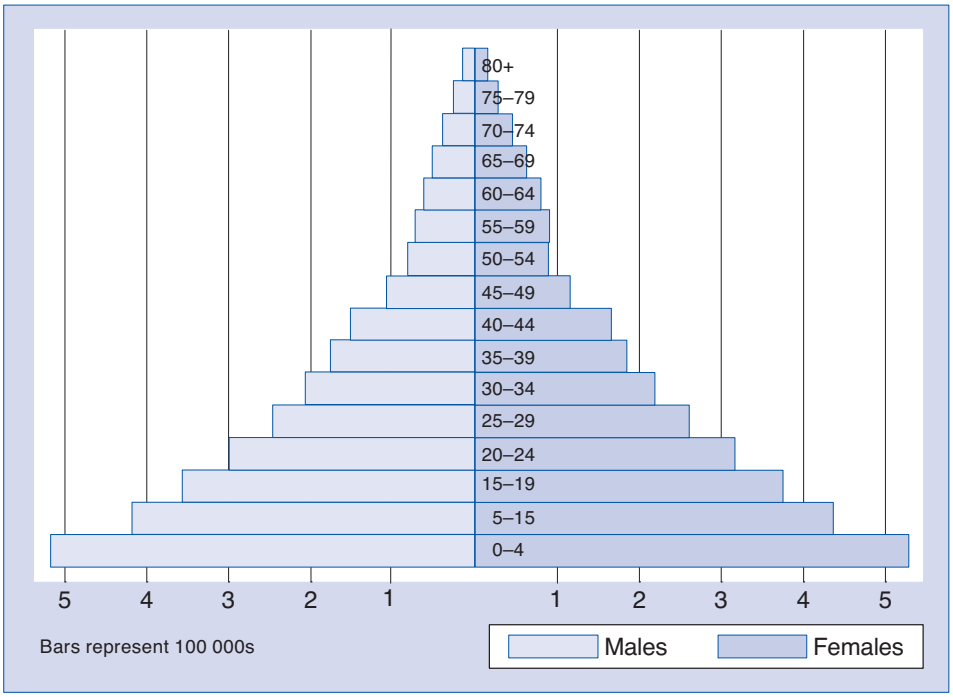


Figure 3A.4.5
Population pyramid – Sierra Leone 2005

GENDER

UK In the UK, slightly more boys are born each year than girls (330 600 boys were born in England and Wales in 2005, compared with 315 235 girls), but overall there are fewer males than females in the UK population (29 271 000 males in the UK in mid-2004, compared with 30 563 000 females). This is because, from age 22 upwards, there are more women in each age group due to higher female immigration and lower female deaths from accidents and suicide. The gap narrows for those in their 40s (more immigration of men in this age group) and then widens (due to longer female life-expectancy and the effect of World War II).

There are differences in the male-to-female ratio across ethnic groups: see Box 3A.4.1.

Box 3A.4.1

More men	More women
Pakistani	White
Bangladeshi	Black
Chinese	Indian

ETHNICITY

UK People whose ethnicity is non-white tend to live more commonly in England than in other parts of the UK, especially in London and the West Midlands. Black, Bangladeshi and Irish populations are particularly concentrated in London. See Box 3A.4.2.

Box 3A.4.2

	England	London	Wales	Scotland	Northern Ireland
White (%)	91.0	71.2	97.9	98.0	99.3
Mixed (%)	1.4	3.2	0.6	0.3	0.2
Asian (%)	4.6	12.1	0.9	1.1	0.2
Black (%)	2.3	10.9	0.2	0.2	0.0
Chinese and other (%)	0.8	2.7	0.4	0.4	0.2

Reproduced from ONS, Northern Ireland Statistics and Scottish Executive Statistics, with permission from the Controller of HMSO.

ETHNICITY AND AGE

UK Non-white groups have a younger age structure than the white population, as a result of past immigration and fertility patterns. The Irish population is the oldest, and the mixed group has the youngest structure. Because of differences in mortality rates, women aged 65+ outnumber men. However, the migration patterns of certain groups mean that there are fewer elderly women in the Pakistani and Bangladeshi populations.

MEASURING ETHNICITY

UK The 2001 Census question on ethnicity asked, *'What is your ethnic group?'* and allowed respondents to choose from one of the 16 options shown in Table 3A.4.2.

As different ethnic populations mix and inter-marry, the fastest growing ethnic group is now people of mixed ethnicity. The 2001 UK census was the first census to ask about the backgrounds of this ethnic group (Table 3.4.2) and, although very useful, this has made comparisons between censuses difficult.

Table 3A.4.2 Classification of ethnicity according to the 2001 census

A	White	1. British 2. Irish 3. Any other white (write in)
B	Mixed	4. White and black Caribbean 5. White and black African 6. White and Asian 7. Any other mixed (write in)
C	Asian or Asian British	8. Indian 9. Pakistani 10. Bangladeshi 11. Any other Asian (write in)
D	Black or black British	12. Caribbean 13. African 14. Any other black (write in)
E	Chinese or other	15. Chinese 16. Any other (write in)

Other related questions were, *'What is your country of birth?'* and, *'What is your religion?'* The latter was a voluntary question.

3A.5 METHODS OF POPULATION ESTIMATION AND PROJECTION

UK The UK's Office of National Statistics produces mid-year population estimates on 30 June each year. These are estimates of the resident population and are calculated using the **cohort component method**.

COHORT COMPONENT METHOD

See Table 3A.5.1.

Table 3A.5.1 Cohort component method for mid-year population estimates

Steps	Data sources
1. Take previous mid-year estimate	Previous mid-year estimate (or the census findings in census years*)
2. Increase the population's age by 1 year	
3. Add births	Registrar General's office
4. Subtract deaths	Registrar General's office
5. Adjust for external migration	**International Passenger Survey** **Irish National Household Survey** (for migration between the UK and the Republic of Ireland)
6. Adjust for internal migration	GP registration data (linked by NHS number)
7. Quality control (check for consistency and against previous estimates)	

**Population estimates are more reliable in census years: the cohort component method is still used, but the census population needs to be adjusted only by the number of weeks between the census and 30 June.*

SPECIAL GROUPS

The groups listed in Box 3A.5.1 are not included in the general population calculation, but are estimated separately.

Box 3A.5.1

Special group	Information source
Boarding school pupils	DfES/Welsh Assembly Government
Prisoners (only if >6 months in prison)	Home Office
Home Armed Forces	Defence Analytical Services Agency
Foreign Armed Forces	US Forces

ESTIMATES

UK Population estimates are produced for the populations listed in Box 3A.5.2, according to age, sex and marital status.

Box 3A.5.2

- UK as whole
- Constituent countries
- Government office regions
- Local authorities
- Health authority areas
- Primary care areas

PROJECTIONS

UK In the UK, population projections are produced by the **Government Actuary's Department** (GAD) and are used for planning across government sectors. The core function of the department is to produce, every 2 years, 25-year projections of the population of the UK and its constituent nations, according to age, sex and marital status. Less accurate 70-year projections are also made.

REPLACEMENT FERTILITY

In the absence of migration, the growth or decline of a population depends on sustained patterns in replacement fertility. Changes in replacement fertility are slow to take effect because of population momentum, i.e. large cohorts of the population in childbearing years will continue to have high numbers of births even if fertility falls. However, in the long run, replacement fertility depends upon three factors:

- Fertility
- Birth sex ratio (males to females)
- Female mortality before the end of reproductive age.

These factors are usually predicted on the bases shown in Box 3A.5.3.

Box 3A.5.3

Factor	Basis of forecast
Birth sex ratio	Stable at 1.05 males per female
Mortality	Extrapolate historical trends
Fertility	Judgement
Migration	Judgement

The uncertainty is greatest regarding fertility and migration. Traditionally, all factors have been forecast along smooth paths, but there is now a move towards **stochastic predictions**, where the probabilities of random fluctuations are incorporated. Likewise, the uncertainty regarding population forecasts has traditionally been

expressed as a range (lower to upper), but **probabilistic population forecasting** (with 95% confidence intervals) is now superseding this.

3A.6 LIFE-TABLES

Life-tables and their demographic applications

Life-tables, also known as **mortality-tables**, list the probabilities (according to age and sex) that a person will die before their next birthday. In the UK these tables are generated by the Government Actuary's Department for the country as a whole and for its constituent nations.

Life-tables are of two principal types: **period** and **cohort**.

PERIOD LIFE-TABLES

These are calculated using the age-specific mortality rates for a given historical period (either a single year, or a run of years), with no allowance made for any later actual or projected changes in mortality.

A period life-table displays the life-expectancy of people of a given age in a given year if they experienced **that year's age-specific mortality rates** for the rest of their lives: see Table 3A.6.1.

Table 3A.6.1 Example of a period life-table of people aged 70 in 2005, calculated using mortality rates

Age (x)	Central rate of mortality*	Probability that someone will die before their next birthday ($x + 1$)	Number of people who survive to age x $l_x+1 = l_x - (1 - q_x)$	Number of people who die aged x $d_x = l_x - l_{x+1}$	Average life-expectancy at age x
Age x	m_x	q_x	l_x	d_x	e_x
70	0.027274	0.026907	100 000	2690.7	12.86
71	0.030869	0.030400	97 309.3	2958.2	12.20
72	0.034271	0.033694	94351.1	3179.1	11.56

For **period** life-tables, use the mortalities:
- For age 70 in 2005
- For age 71 in 2005
- For age 72 in 2005, etc.

**Number of deaths in people aged x over the past 3 years divided by the average population at that age in the same period.*

Adapted from GAD website: www.gad.gov.uk.

Official life-tables that relate to past years are generally period life-tables.

COHORT LIFE-TABLES

In contrast, cohort life-tables are calculated using age-specific mortality rates which do allow for known or projected changes in mortality in later years.

A cohort life-table displays the average life-expectancy of a group of people of a given age in a given year if they experienced projected future age-specific mortality rates from the series of future years in which they will actually reach each succeeding age if they survive: see Table 3A.6.2.

Table 3A.6.2 Example of a cohort life-table of people aged 70 in 2005, calculated using projected future mortality rates

Age (x)	Central rate of mortality*	Probability that someone will die before their next birthday ($x + 1$)	Number of people who survive to age x $l_{x+1} = l_x - (1 - q_x) = l_x - p_x$	Number of people who die aged x $d_x = l_x - l_{x+1}$	Average life-expectancy at age x
Age x	m_x	q_x	l_x	d_x	e_x
70	0.027274	0.026907	100 000	2690.7	12.86
71					
72					

Cohort life-tables would use the predicted mortalities of:

- Age 70 in 2005
- Age 71 in 2006
- Age 72 in 2007, etc.

**Number of deaths in people aged x over the past 3 years divided by the average population at that age in the same period*

So, if mortality rates are projected to fall in the future, then the cohort life-expectancy at a given age will be longer than the period life-expectancy for that age.

APPLICATIONS OF LIFE-TABLES

Life-tables are used for planning in all sectors of government, but particularly with regard to pensions and social insurance. Several values can be derived from life-tables, including:

- Proportion of people born in different years who are still alive
- Remaining life-expectancy of people at a particular age
- Probability of surviving to a particular age.

As well as separate life-tables for men and women, it is also possible to distinguish other factors that affect mortality, including ethnicity, social class and smoking status.

3A.7 POPULATION PROJECTIONS

See Sections 3A.5 and 3A.6.

UK Note that, in making official predictions, the Office of National Statistics assumes that there will be improvements in mortality and increases in net immigration.

3A.8 EFFECT ON POPULATION STRUCTURE ON FERTILITY, MORTALITY AND MIGRATION

If the effects of migration are ignored then, for the population size to remain stable, the average female needs to have one female child who survives long enough for one female grandchild to be born. This level of fertility – called the **replacement fertility** – needs to be higher than two children per female in order to compensate for:

- **Mortality** (not all daughters will survive to the end of reproductive age, especially in developing countries)
- **Unbalanced sex ratio** at birth (in the UK the ratio of male births to female births is approximately 105:100).

Replacement fertility values vary between countries (see Box 3A.8.1), and the commonly quoted figure of 2.1 children applies only to the developed world. International variation is mostly due to mortality differences – particularly with respect to HIV/AIDS.

Box 3A.8.1

World region or country	Replacement fertility
Reunion	2.05
Europe	2.10
Africa	2.70
Sierra Leone	3.43

UK In the UK, replacement fertility has fallen because of decreasing mortality of the young. Again ignoring immigration, if period fertility drops below replacement fertility, then the size of the population will eventually fall. However, population decline may not be observed immediately because of the buffering effects of other factors, including:

- Age structure of the population
- Changes in age-specific mortality rates
- Childbearing postponement.

The last phenomenon has occurred in several developed countries in recent decades. It has the effect of stretching out the population into the future so that there are fewer people alive at any moment in time.

POPULATION TIME BOMB

Population decline in the UK would lead to problems often referred to as the *population time bomb*: see Figure 3A.8.1. Falling birth rates and low mortality in older age groups both result in an ageing population. As the number of people who are economically active relative to those requiring health and social care starts to fall, the resulting effects may have profound implications for society.

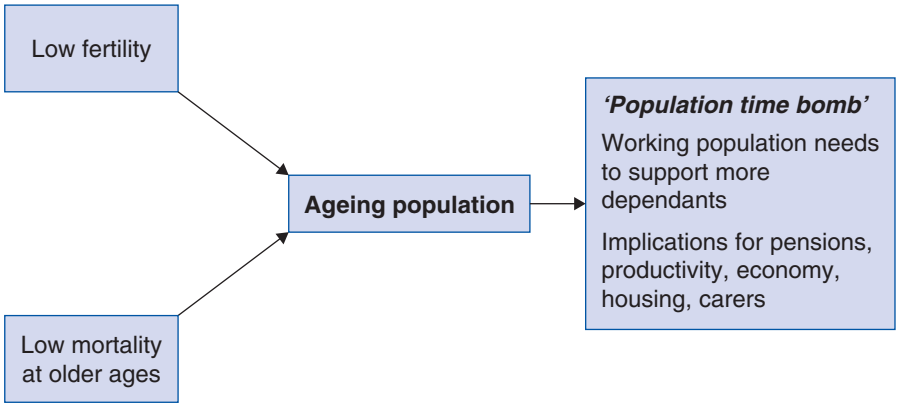


Figure 3A.8.1 Population time bomb

In economic terms, a country achieves its optimum population when productivity per capita is highest.

- Under-populated countries are those that can increase their productivity by increasing their population.
- Over-populated countries are those that can increase productivity by reducing their population.

IMMIGRATION

See 4C11.

Net migration is the difference between the number of people emigrating and immigrating. Both immigration and emigration occur most commonly among young adults, with slightly more males than females migrating each year overall.

UK In the UK, both immigration and emigration have increased in recent years. Immigration into the UK has been predominantly of citizens from the 10 accession countries that joined the European Union in 2004. Emigration has been mostly in people from the 25–44 age group, and has been mainly to other EU countries, and to Australia and New Zealand.

3A.9 HISTORICAL CHANGES IN POPULATION SIZE AND STRUCTURE, AND FACTORS UNDERLYING THEM

Over the course of centuries, large populations have changed with regard to their age structure and geographical distribution. The populations of many countries can be seen to have fitted into three phases: pre-industrial, industrial and post-industrial: see Box 3A.9.1

Box 3A.9.1

Historical phase	Population growth	Fertility and mortality	Age
Pre-industrial	Slow	High	Young
Industrial	Fast	Intermediate	Intermediate
Post-industrial	Slow	Low	Old

A more detailed '**demographic transition model**' describes five phases rather than three:

1. Pre-industrial
2. Developing
3. Urbanisation
4. Developed
5. De-industrialised (i.e. switch from manufacturing to service-based economy).

See Figure 3A.9.1.

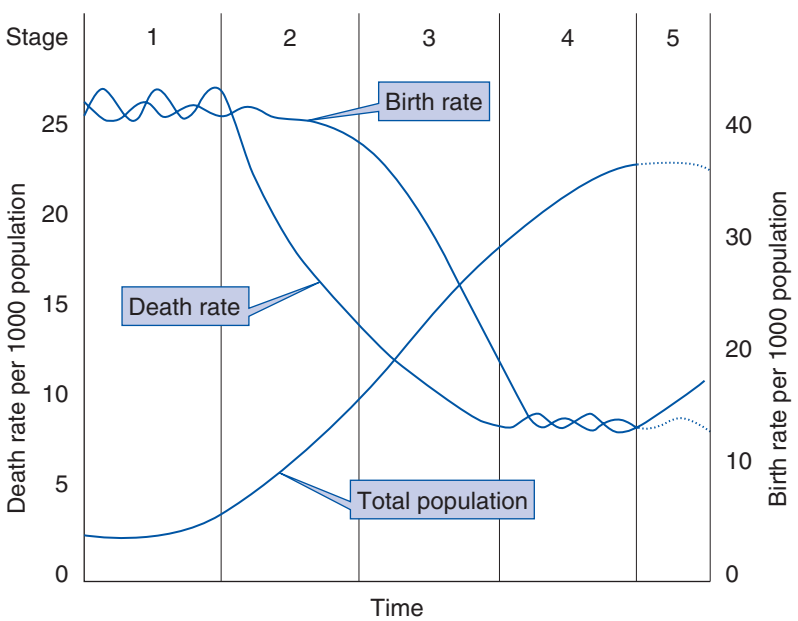


Figure 3A.9.1 Illustration of birth rates, death rates and population sizes for the five phases. *Reproduced from http://en.wikipedia.org/wiki/Demographic_transition.*

UK UNITED KINGDOM

The UK population of around 60 million (2001 census) has increased by 18% since 1951 due to births exceeding deaths. There were more births than deaths in the UK in every year since 1901 (with the exception of 1976). Net immigration has also been a factor since the mid-1990s. The population was projected to peak at 67 million in 2005 and then gradually fall.

Underlying factors affecting population size include changes in **fertility**, **migration** and increasing **urbanisation**: see Box 3A.9.2.

Box 3A.9.2

Fertility	There are fewer under-16s and more over-65s **Birth rates** rose after both World Wars, 'baby boom' in 1960s, steadily falling until reaching a trough in 1970s, some increases in 1980s and 1990s but lowest levels early 2000s The average age at which women give birth to their first child is increasing. More women remain childless (1 in 5 now compared with 1 in 10 for women born in the mid-1940s). See also Section 3B.4
Migration	**Net immigration** into the UK is an increasingly important factor in population size
Urbanisation	England is one of the most crowded countries in the world. Over **90% of inhabitants live in urban areas** – covering just 8% of the land area

WORLD

The United Nations Population Division expects the absolute number of the world's infants to begin falling in 2015, and the number of children under 15 to begin falling by 2025. Other forecasters have calculated that the world population will peak at **9 billion** in **2070**, with the average age of the population steadily rising.

3A.10 EFFECTS OF DEMOGRAPHIC CHANGE ON HEALTH CARE

Significance of demographic changes for the health of the population and its need for health and related services

See Table 3A.10.1.

3A.11 POLICIES TO ADDRESS POPULATION GROWTH NATIONALLY AND INTERNATIONALLY

Policies in a particular country are shaped by the nature of demographic challenges faced by that country. In many developed countries, relatively low birth rates require policy-makers to address issues associated with an ageing population. Worldwide, there is a need for policies to restrict or limit population growth but also to recognise that there will be an increasingly ageing population.

UK UNITED KINGDOM

The ageing population of the UK means that there will be a relative shortage of working-age adults compared with the demand from the total population. This is particularly the case regarding people with key professional qualifications (e.g. clinical staff).

Policy-makers can increase the number of working-age adults through encouraging:

- More people to have larger families by means of **family friendly policies** (e.g. subsidised paternity and maternity leave, tax allowances for parents)
- People of working age to move to the UK through **managed migration** strategies.

Table 3A.10.1 Changes in population size and structure and effects on the need for health care

Demographic changes		Effect on health and need for health and related services
Population structure	Age	An ageing population will place greater demand for geriatric, intermediate and social/personal care
	Ethnicity	Greater ethnic diversity will lead to different risk factors for disease, different patterns of disease and demand for different models of provision (e.g. bilingual health-care workers, culturally specific services such as women-only group sessions) Access to health care should be **monitored** to ensure that there is no discrimination due to language, cultural or knowledge barriers
Population mobility	Short term (travel)	Increased global spread of infectious disease and pandemics (e.g. SARS)
	Longer term (migration)	International spread of emerging diseases (e.g. TB)
Urbanisation		Increased spread of infectious diseases
Housing	People living longer	Demand for new homes has increased over the past 30 years.
	More people living alone	Housing shortages in south-east England are acute, leading to a shortage of workers. The situation is due to get worse if current trends continue
	Lower building rates	Risk of health problems associated with overcrowding/poor housing, e.g. respiratory illnesses

Demand for affordable housing is increasing, and can be met by:

- Increasing the supply and affordability of housing in areas of high demand (e.g. Thames Gateway)
- Keyworker policies (providing subsidised housing for staff, e.g. science teachers)
- Neighbourhood renewal to tackle problems in declining areas
- Sustainable development.

INTERNATIONAL

According to circumstances, strategies may be aimed at:

- Restricting population growth (e.g. China's one-child policy)
- Limiting rate of growth (e.g. developing countries encouraging birth control through providing subsidised contraception or sterilisation)
- Targeting resources into the research of diseases associated with ageing (e.g. Alzheimer's disease)
- International agreements to tackle environmental problems and the demand for natural resources.

3B
Sickness and Health

Information regarding sickness and health can be derived from a variety of sources, not just from health service data. Effective use of this information requires a familiarity with the major resources, an understanding of the limitations to data validity, and an appreciation of methods for relating one set of data to another by means of data linkage.

3B.1 ROUTINE MORTALITY AND MORBIDITY DATA

Sources of routine mortality and morbidity data, including primary care data, and how they are collected and published at international, national, regional and district levels.

Routine information is that which is regularly collected and made partially or fully available. It provides information on mortality or morbidity in a standardised format (see Section 3A.2). The sources described below relate mainly to national systems in England and Wales, although similar systems exist in Scotland and certain other countries. Additional sources may be available locally and regionally.

MORTALITY DATA

Sources of mortality data are listed in Table 3B.1.1.

UK Table 3B.1.1 Sources of mortality data

ONS – death certificates and death registration	
Information	Cause of death (inserted by medical practitioner on the certificate) Additional information collected when the death is registered includes: • Date and place of death • Name and surname of the deceased, date and place of birth, occupation, etc. • Details of any spouse or civil partner • Usual address
Collection, coding and analysis	Deaths must be registered within 5 days by a relative or friend of the deceased to the local registrar (8 days in Scotland) Statutory duty for death to be registered before funeral can occur Register information is stored at district level registry office
Uses	Legal requirement Analysis of trends, comparisons between areas Health needs analysis for serious conditions
Strengths	Complete and timely Relatively accurate
Weakness	Clinical code less accurate for older patients with several co-morbidities Coding accuracy varies Problems comparing different years if different ICD classification used (e.g. shift from ICD-9 to ICD-10 in January 2001 – see Section 3B.3) Risk of bias in social class measures due to occupational advancement (see Section 3B.2)

Additional information is collected from deaths that result from a road traffic accident: see Table 3B.1.2.

UK Table 3B1.2 Police road traffic accidents – 'STATS19'

Information	Injury/death due to road traffic accidents
Collection, coding an analysis	All reported road accidents* involving personal injury are recorded at the time of the accident on the STATS19 form Police process the STATS19 data and send the information to the Department for Transport, which adds the information to the National Road Casualty database Both the police and the Department for Transport validate the information received
Uses	Indications of mortality from incidents occurring on roads Can link with A&E data on deaths from road accidents
Strengths	More detailed information about each incident and the types of vehicles involved than recorded by A&E May include accidents not seen by health services

Weakness	Not all accidents are reported to the police Morbidity is rated differently by the police and the health service

**Note that, in Australia, public health practitioners encourage use of the term 'road traffic collisions' rather than 'accidents', since the latter carries a connotation of non-preventability.*

MORBIDITY DATA

Information on morbidity in the UK comes from:

- Condition-specific registers and datasets (see Tables 3B.1.3 and 3B.1.4)
- Individual-level secondary care databases (inpatient, see Table 3B.1.5, and outpatient, see Table 3B.1.6)
- Aggregate-level community records (see Table 3B.1.7)
- Primary care: limited information (not described here) is collected on optometry, pharmacy and dental services, mainly for payment purposes. More information is available on general practice and prescribing (see Tables 3B.1.8 and 3B.1.9).

International data sources are described in Table 3B.1.10.

UK SOURCES OF PUBLIC HEALTH INFORMATION

Table 3B.1.3 Condition-specific information: regional cancer registers on every new diagnosis of cancer

Information	Personal identifiers (needed in order to eliminate duplicates) Socioeconomic characteristics Disease status (cancer type, stage) Treatment Outcomes
Collection, coding and analysis	Sources for collection of information include cancer centres, treatment centres, hospices, private hospitals, cancer screening programmes, other cancer registers, general practices, nursing homes and death certificates Cancers are coded using a system common to all the registries in the UK The Office for National Statistics collates, analyses and publishes the registers' data
Uses	Monitor trends for incidence and survival of cancer Compare epidemiology and performance in different areas
Strengths	Complete Useful for continuing conditions with reliable diagnosis
Weakness	Time lag between collection of data and available information Time and labour are intensive to set up and maintain Risk of inaccuracies if diagnostic criteria change

Table 3B.1.4 Condition-specific information: minimum datasets specified in areas such as mental health and cancer

Information	Standardised collections of information about specific conditions, including: • **Personal:** ethnicity • **Clinical:** treatment received, inpatient/outpatient, re-admissions • **Administrative**
Collection, coding and analysis	Information collected by clinicians at the point of service and collated at regional and national levels May be fed to other databases, e.g. cancer registers
Uses	Monitor trends Performance management, e.g. regarding meeting the milestones set out in National Service Frameworks Allow health and social care professionals to measure and compare the care that they provide
Strengths	Potential to standardise care and reduce inequities in service provision
Weakness	Many still in development

Table 3B.1.5 Inpatient and day-case treatment: hospital episode statistics (HES)

Information	Consultant episodes **Personal:** name, NHS number, date of birth, ethnicity (address added but stripped out if NHS number present) **Administrative:** start and end dates of the stay, hospital, ward, specialty code, waiting time **Clinical:** ICD-10 code, ONS Classification of Surgical Operations and Procedures
Collection, coding and analysis	Hospital episode coded by hospital coders not clinicians Collected as part of monthly mandatory information submission by hospitals Sent to Secondary Uses Service (SUS) HES provided as nationally available extract Trusted organisations (e.g. public health observatories) have full database access
Uses	Payment from commissioner to provider Analysis of hospital usage, waiting times Assessment of quality and outcomes of care (also external performance management by Department of Health, and inspection by Healthcare Commission) Estimation of health need for conditions routinely managed in secondary care
Strengths	Generally complete – hospitals must submit information if they are to be paid Timely – information collected routinely; available quarterly and yearly from HES Mortality data can be linked to HES databases to produce statistics linking episode to outcome and to individual patients

Weaknesses	Accuracy dependent on quality of NHS coders (very variable) Variable completeness Relates to episodes, not patients (therefore may overestimate need if one patient has several episodes) Only useful for conditions that are generally admitted to hospital

Table 3B.1.6 Individual outpatient activity: secondary uses service (SUS)

Information	**Personal:** NHS number, date of birth, postcode, not ethnicity **Administrative:** referral and appointment dates, attendance/DNA (do not attend), outcome **Clinical:** name of clinic recorded but no information on diagnosis
Collection, coding and analysis	Collected as part of monthly mandatory submission by hospitals
Uses	Not always used at PCT level
Strengths	Measure of outpatient service demand Useful if clinic name gives indication of need (e.g. HIV clinic, diabetes clinic)
Weakness	Missing useful information, e.g. ethnicity, clinical data

Table 3B.1.7 Community level aggregated data: Körner data*

Information	Sexually transmitted infections: **KC60** GUM clinic data on attendances, demographic characteristics Immunisations: **KC50**, e.g. MMR, flu Adult screening programmes: **KC53, 63**
Information	Anonymous information collected at clinic; community level information required for national returns to DH
Collection, coding and analysis?	Performance management, e.g. uptake of routine immunisations, screening programmes Indication of use of services, level of infections, e.g. GUM
Strengths	Confidential
Weaknesses	For KC60, patients are not obliged to give truthful demographic details; therefore difficult to draw conclusions from data analysis Not straightforward link to geographical area: people can self-refer and may choose to go far from home in order to protect confidentiality

**Data-set devised by Dame Edith Körner's working party to record NHS service activity.*

Table 3B.1.8 Primary care data at aggregate level: QMAS (Quality Management and Analysis System)*

Information	Disease registers – see Section 3B.3 146 evidence-based indicators forming the quality and outcomes framework (QOF) over four domains: 1. Clinical, based on 11 conditions including hypertension, asthma, diabetes, coronary heart disease, mental health 2. Organisational, including records, practice management issues 3. Patient experience (assessed through surveys and consultation length) 4. Additional services
Collection, coding and analysis	Monthly extract from a general practice sent to QMAS either automatically or manually Payments are linked to evidence-based indicators through the Quality and Outcomes Framework (QOF) QMAS accessible online to PCTs
Uses	Paying general practices according to the services delivered and the degree to which certain milestones are met Registers can indicate disease prevalence QMAS helpful to plan primary care/referral services
Strengths	Many conditions treated only in primary care Incentive for improving the quality and comprehensiveness of information from general practice (linked to payment) QOF score gives an indicator of the quality of clinical care in a practice Relatively complete – most people are registered with a general practice
Weaknesses	Accuracy dependent on the coding and currency of the registers in general practice QMAS primarily for payment, not performance management Cannot use QOF scores to compare practices on performance: different list sizes and different population characteristics may affect QOF reached QOF is voluntary, though most practices have signed up to it

**The national system supporting the new GMS GP contract.*

UK **Table 3B.1.9** Primary care data prescribing information (see Section 3B.5): PACT, ePACT

Information	By quantity – number of items prescribed or number of tablets/grams of active ingredient By cost, e.g. net ingredient cost (NIC)
Collection, coding and analysis	Information from prescriptions made (in large part in general practice) and sent to Prescription Pricing Authority Data on each practice sent back to PCTs for payment
Uses	Monitoring of prescribing practice Payment General practice prescribing incentive scheme (being replaced by QOF in many places)

Strengths	Complete
Weaknesses	Not linked to patients or conditions, so difficult to use for performance management

INTERNATIONAL COMPARISON

NZ The Ministry of Health collects, analyses and disseminates information on population health either directly or through contracted service providers. Most of this work is conducted by two Ministry units, the **New Zealand Health Information Service** (NZHIS) and **Public Health Intelligence** (PHI). The main systems are listed in Table 3B.1.10.

NZ **Table 3B.1.10** New Zealand sources of morbidity information

Health outcomes	Mortality data Hospital discharges (public and private) Cancer registrations and deaths Mental health service use Maternal and newborn data Surveillance (see surveillance section)
Health status and behaviour (regular surveys programme)	NZ Health Survey NZ Alcohol and Drug Use Survey NZ Sexual and Reproductive Health Survey NZ Tobacco Use Survey NZ Adult and NZ Child Nutrition Surveys
Health-care services and interventions	Health workforce data Surgical throughput and waiting lists data Population screening data Pharmaceutical prescribing data
Other information sources	Census data Social wellbeing (indicators collected by Ministry of Social Development) Environmental quality (information collected by multiple government agencies, e.g. air quality, water quality)

3B.2 BIASES AND ARTEFACTS IN POPULATION DATA

Potential sources of bias and artefact should always be borne in mind when working with population data.

BIASES

A bias is a systematic error (see 1A.14). Biases can affect population data in a number of ways, including the following.

UK LIST SIZE

GP practices sometimes overestimate their list of registered patients by failing to remove, or delaying the removal of, people who have died or moved away from the practice. This leads to a systematic error with regard to the estimated population size and structure. As a consequence, it can lead to an underestimate of service provision, e.g. in terms of vaccine uptake.

UK CENSUS ERRORS

Census estimates of population size can be biased by differential response rates: in 2001 there was under-enumeration of men in their 20s (the group least likely to respond to the survey). Other groups with low response rates included: inner-city areas (particularly London) and areas with a high proportion of people with difficulties filling in the census (e.g. language problems).

UK STATUS INFLATION

There is a tendency for people to inflate the socioeconomic status of the deceased when registering the death or completing surveys. This 'occupational advancement' biases the population structure with regard to social class.

ARTEFACTS

Artefacts in population data are caused by spurious differences between an observed population characteristic and the true underlying characteristic. Artefacts may hinder accurate comparisons between areas and trends over time. Adjusting for or reducing artefacts requires an understanding of how data are collected and the potential sources of inaccuracy, bias and confounding.

INTERNATIONAL CLASSIFICATION OF DISEASE

In January 2001, ICD-9 was replaced by ICD-10 (see Section 3B.3), thereby causing an artefact in mortality data when comparing deaths before and after this date.

UK CENSUS CHANGES

Census ethnicity information codes changed between 1991 and 2001, when for the first time people were allowed to identify their ethnicity as 'mixed'. This complicates comparisons between the two censuses.

3B.3 THE *INTERNATIONAL CLASSIFICATION OF DISEASES*

The International Classification of Diseases *and other methods of classification of disease and medical care*

The *International Classification of Diseases* (ICD) belongs to the WHO Family of International Classifications (see Box 3B.3.1), the purpose of which is to provide a common language for health-related topics.

Box 3B.3.1

WHO Family of International Classifications

- *International Classification of Diseases* (ICD)
- *International Classification of Functioning, Disability and Health* (ICF)
- *International Classification of Health Interventions* (ICHI)

Between them, these classifications enable the consistent collection, analysis and presentation of data, to enable comparisons over time and between populations. In particular, they allow:

- Analysis of population health
- Monitoring of disease frequency
- Classification of death certificates and hospital records
- Mortality and morbidity statistics to be collated using a common framework.

ICD is primarily an international standard classification for mortality statistics and contains a standard format for death certification. Its history dates back to the 1850s (*International List of Causes of Death*), and since 1948 it has been administered by the WHO.

Diseases mentioned on the death certificate are translated into codes, ranging between A00 ('Cholera') and Z99.9 ('Dependence on unspecified enabling machine and device'): see Table 3B.3.1. By applying coding rules contained in the ICD (which prioritise and consolidate codes), a single cause of death will be selected.

The ICD is revised every 10–20 years. The WHO advises that it is problematical to translate between the codes of one revision to another. Deaths during the bridging period should be dual-coded (according to the old and new revisions) to enable comparisons to be made of mortality measures as collected under the two systems.

The 10th Revision (ICD-10) came into use in 1994 and contains twice the number of codes in ICD-9. A clinically modified version of ICD-9 (ICD-9-CM) contains more precise detail for describing mortality as opposed to morbidity. Although ICD-10-CM has now been written, at the time of writing ICD-9-CM remains the standard for reporting morbidity.

ICD-10

ICD-10 was the first revision to adopt alphanumeric codes (and hence potential problems of confusion between zero and letter O, and between I and 1). These codes range between three and six characters in length (with a decimal after the third character if the code is four characters long or greater). Box 3B.3.2 shows examples.

Box 3B.3.2

M60	Myositis
J85.2	Abscess of lung without pneumonia
M11.9	Crystal arthropathy, unspecified

All codes for injuries (range S00–T98) must have a corresponding external cause code (range V01–Y98).

Neoplasms are coded by morphology and site codes. Code U is reserved for future use.

Table 3B.3.1 ICD-10 disease codes

Code	Contents
A00–B99	Certain infectious and parasitic diseases
C00–D48	Neoplasms
D50–D89	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
E00–E90	Endocrine, nutritional and metabolic diseases
F00–F99	Mental and behavioural disorders
G00–G99	Diseases of the nervous system
H00–H59	Diseases of the eye and adnexa
H60–H95	Diseases of the ear and mastoid process
I00–I99	Diseases of the circulatory system
J00–J99	Diseases of the respiratory system
K00–K93	Diseases of the digestive system
L00–L99	Diseases of the skin and subcutaneous tissue
M00–M99	Diseases of the musculoskeletal system and connective tissue
N00–N99	Diseases of the genitourinary system
O00–O99	Pregnancy, childbirth and the puerperium
P00–P96	Certain conditions originating in the perinatal period
Q00–Q99	Congenital malformations, deformations and chromosomal abnormalities
S00–T98	Injury, poisoning and certain other consequences of external causes
Z00–Z99	Factors influencing health status and contact with health services
U00–U99	Codes for special purposes

DIAGNOSIS-RELATED GROUPS AND HEALTH-CARE RESOURCE GROUP

Diagnosis-related group (DRG) is an American system to classify hospital cases into one of approximately 500 groups that are expected to require similar hospital resource use. DRGs are based on ICD diagnoses, together with procedures, age, sex and the presence of complications or co-morbidities.

UK In the UK, the analogous health-care resource group (HRG) is the unit of currency for commissioning health services.

3B.4 MEASUREMENTS OF HEALTH STATUS

Rates and ratios used to measure health status, including geographical, occupational, social class and other sociodemographic variations

See Section 1A.1.

UK HEALTH SURVEYS IN THE UK

Health surveys supplement data provided from patients' interaction with the health-care system. Different models are in use in different parts of the UK.

Eng HEALTH SURVEY FOR ENGLAND

The Health Survey for England seeks to obtain a representative sample of people living in private households each year and it includes both adults and children. The survey consists of a health interview and an examination conducted by a nurse who takes a variety of measurements (e.g. height, weight and blood pressure). The Health Survey for England contains a core that is repeated each year, and has one or more supplements: modules on subjects of special interest.

The 'core' includes questions on:

- General health and psychosocial indicators
- Smoking
- Alcohol
- Demographic and socioeconomic indicators
- Use of health services and prescribed medicines.

Scot SCOTTISH HEALTH SURVEY

The Scottish Health Survey was undertaken in 1995, 1998 and 2003. As in England, the Scottish surveys also target both adults and children, and utilise a health interview and an examination by a nurse. In addition to demographic data, the survey in 2003 included questions on:

- General health and illness
- Cardiovascular disease
- Respiratory disease (including asthma)
- Dental health
- Health-related behaviours
- Use of health services
- Height and weight.

The nurses measured blood pressure and lung function, and they collected saliva, blood and urine samples. Some subjects were asked to have an ECG recording.

Wal WELSH HEALTH SURVEY

The Welsh Health Survey (WHS) was first carried out in 1995 and repeated in 1998 and 2003–5. This survey does not include a nurse examination, and it was targeted at adults only during the first two iterations. It is a postal survey designed to provide a picture of health of the people of Wales, the way the NHS is used and areas where services could be improved. The 2003–05 survey included a 15-minute face-to-face questionnaire. Results from the 2003–05 survey are not comparable with those from the previous surveys because of differences in the questionnaires and the ways in which the survey was designed and conducted.

In addition to demographic data, the content of the 2003–05 survey included:

- General health status (SF-36)
- A range of reported illnesses and other conditions (such as eyesight or hearing difficulty)
- Reported lifestyle behaviours, including smoking, drinking, fruit and vegetable consumption, physical activity and body mass index
- Reported use of a range of health services.

NI NORTHERN IRELAND HEALTH AND SOCIAL WELLBEING SURVEY

The Northern Ireland Health and Social Wellbeing Survey was conducted in 1997, 2001 and 2006. It is designed to yield a representative sample of all adults aged 16 and over living in Northern Ireland.

The questionnaire consists of a household interview followed by an individual interview with each person in the household aged 16 and above. The individual interview consists of core modules and modules that will recur on a regular cycle. Core items include accommodation, tenure, employment status, educational qualifications, family information, smoking and drinking, and health and ill health. Non-core items include physical activity and sexual health.

In 1997 the survey included physical measures for one respondent selected at random from each household. Qualified nurses recorded details of prescribed medication, and measured height, waist, hip, weight and blood pressure. A blood sample was also taken to measure the level of cholesterol (non-fasting).

In addition to demographic details, the questionnaire covers:

- Lifestyle habits, including sexual health
- General health, long-standing illness
- Common chronic diseases and conditions
- Stressful life event and possible mental health problems (SF-36 was also used in 1997)
- Informal care and the lifestyle of carers
- Characteristics of the people whom carers are looking after.

3B.5 PRESCRIBING DATA AND PHARMACOVIGILANCE

Prescribing data (information about the volumes, costs and types of drugs prescribed and dispensed) can be useful for both clinical and managerial reasons, as well as for identifying adverse drug effects.

Reasons for monitoring prescribing include:

- **Cost** containment – prescription costs are rising; money spent on prescriptions means that less is available for other areas of health care
- Monitoring adherence to **guidelines** (e.g. NICE)
- Detecting aberrant or **inappropriate** clinical performance (e.g. controlled drug prescriptions, antibiotics)
- Addressing local **priorities** (e.g. reduction of heart disease through prescribing drugs such as statins)
- Performance-related **incentive** payments (e.g. prescribing elements of QOF – see Section 3B.1).

CHALLENGES

While prescribing data are accessible and useful for measuring costs, it is more difficult to measure the quality of prescribing. Information on why a drug was prescribed or for whom it was prescribed is not collated, and information currently can be obtained only by the laborious process of auditing individual patient records.

SOURCES OF DATA

PACT

Prescribing **A**nalysis and **C**os**T** (PACT) data provide GPs and other prescribers with reliable and regular information on their NHS prescribing habits and costs.

EPACT

This is an electronic system for pharmaceutical and prescribing advisors. It allows real-time on-line analysis of the previous 5 years' prescribing data held on the NHS Prescribing Database. The data available include:

- Budgets and expenditure forecasts
- Costs and volumes of prescribing
- Prescribing totals by prescriber at all *British National Formulary* (BNF) levels
- Prescribing from the *Nurse and Extended Nurse Formularies*

- Working environment for nurses and supplementary prescribers (i.e. community or practice)
- Patient list sizes
- Low Income Scheme Index scores for practices (released in May 2004)
- Average daily quantities and defined daily doses.

UNITS OF MEASURING PRESCRIBING

ITEM

This is the number of prescription items listed on a prescription form. While easy to measure, caution is needed regarding repeat prescriptions. For example, a GP who writes monthly prescriptions will appear to prescribe more than a GP who prescribes quarterly – even though the total amount of drug is identical.

QUANTITY

- Number of tablets or millilitres, milligrams, etc.
- Strength of active ingredients: stronger active ingredients may require fewer milligrams for the same number of tablets
- Dosing schedules

Note that caution is needed regarding potency (e.g. fewer milligrams of one type of statin are needed to achieve a particular drop in blood cholesterol compared with another).

NET INGREDIENT COST

This is the basic price of a drug. It can be used to measure the volume of similarly priced groups of drugs at equivalent doses. However, where there is a large price difference, it is not an accurate measure of use.

ACTUAL COST

This is calculated by taking the basic price of the prescription items, deducting the *National Average Discount*, and then adding an allowance for the container. Actual cost is used in Prescribing Monitoring Documents (PMDs).

DEFINED DAILY DOSES

These are based on maintenance doses and are not suitable for one-off doses. Note that the defined daily dose (DDD) is not the recommended dose, nor is it necessarily a dose that a patient could practically receive. For example, simvastatin has a DDD of 15 mg but is available only in 10 or 20 mg tablets.

AVERAGE DAILY QUANTITIES

This is an England-specific system developed by the Prescribing Support Unit (PSU) and is equivalent to the DDD.

PRESCRIBING UNITS

Prescribing costs can vary across different organisations because of different prescribing practices, but also because of the features of the local population. As a result, it is not valid simply to use an average cost per patient as a measure of prescribing spend. In England, **prescribing units (PUs)** have been developed to take account of the fact that older patients have a greater need of medication. This allows comparisons of prescribing costs across areas with different age structures in their populations. Since 1983, prescribing units have been further refined into:

- **ASTRO-PUs** (age, sex and temporary resident originated prescribing units): take account of a wider range of demographic factors other than age in comparisons of prescribing or resource allocation decisions
- **STAR-PUs** (specific therapeutic group age–sex weightings related prescribing units): units specific to different therapeutic areas which also take into account demographic factors.

PHARMACOVIGILANCE

Pharmacovigilance aims to prevent or reduce harm resulting from medicines through processes described in Table 3B.5.1.

Table 3B.5.1 Processes involved in pharmacovigilance

Monitoring	Monitoring the use of medicines in everyday practice to identify previously unrecognised adverse effects or changes in the patterns of adverse effects
Risk assessment	Assessing the risks and benefits of medicines in order to determine what action, if any, is needed to improve their safety
Information provision	Providing information to health-care professionals and patients to optimise safe and effective use of medicines
Measuring	Assessing the impact of any action taken

INFORMATION SOURCES USED FOR PHARMACOVIGILANCE

A range of different national and international systems is used:

- Spontaneous adverse drug reaction (ADR) reporting schemes, such as the Yellowcard and black triangle systems (see below)
- Clinical and epidemiological studies in the worldwide medical literature
- Information from pharmaceutical companies
- Information from worldwide regulatory authorities
- Morbidity and mortality databases.

Information from any of these sources may identify unexpected side effects or indicate that certain side effects occur more commonly than was previously believed. They may also reveal that certain patient groups are more susceptible to particular problems than others. Such findings can lead to changes in the marketing authorisation of the medicine, through:

- Restrictions in use
- Changes in the dose of the medicine
- Introduction of specific warnings of side effects in product information.

UK THE YELLOW CARD SCHEME

The Medicines and Healthcare Products Regulatory Authority (MHRA) and the Committee on Safety of Medicines (CSM) run the UK's spontaneous adverse drug reaction reporting scheme, called the Yellow Card Scheme. Yellow Cards are distributed to health-care professionals, including as an appendix to the BNF. This receives reports of suspected adverse drug reactions from health professionals and patients.

UK BLACK TRIANGLE SCHEME

New medicines and vaccines are labelled with a black triangle symbol (▼) in the BNF and on all product information and advertisements. Health professionals are urged to report any suspected adverse reactions that might involve one of these products. For established medicines that are not marked with a black triangle, health professionals should report only serious or unusual suspected adverse reactions.

RISK MINIMISATION

Occasionally, where the risks of a medicine are found to outweigh the benefits, the drug (or even an entire drug class) may be removed from the market. More commonly, the risk of a side effect may be avoided or reduced by:

- Including **warnings** in the product information or on the package labelling
- **Restricting** the indications for use of a medicine
- Changing the **legal status** of a medicine (e.g. by switching from pharmacy to prescription only).

COMMUNICATION WITH HEALTH-CARE PROFESSIONALS AND PATIENTS

UK The MHRA communicates with health-care professionals and patients to warn about adverse effects and to provide feedback of information. It does this through:

- Patient Information Leaflets (PILs) and Summaries of Product Characteristics (SPCs) for medicines that are updated when new safety issues are identified
- Urgent warnings about drug hazards via letters to all doctors and pharmacists
- MHRA and CSM regular drug safety bulletin, *'Current Problems in Pharmacovigilance'*, sent to doctors and pharmacists
- Fact sheets of major safety issues for both health-care professionals and patients
- Safety alerts on the MHRA website.

3B.6 DATA LINKAGE

Data linkage is the process of matching information in one data source with that in another. Linking two databases can provide useful extra information and enable different analyses to be carried out. For example, linking hospital episode statistics (HES) databases to mortality data can provide information about the outcomes of hospital activity. One of the earliest record linkage studies was the Oxford Record Linkage Study (ORLS), which started in 1963. The ORLS consists of computerised abstracts of records of all types of hospital inpatient care, and records of births and deaths in the Oxford region. When data collection ceased in 1999 all patient identifiers were removed from the database. The data-set includes 10 million records relating to over 5 million people.

Eng Box 3B.6.1

Example: Linkage in the NHS – *Connecting for Health*

Connecting for Health is a national programme for improving the information technology of the NHS to ensure that the appropriate information is available to any clinician regardless of geography. It should mean that patients will no longer be asked the same information every time they see a different health professional. Furthermore, it should reduce the chances of information (e.g. X-rays, patient notes) becoming lost. *Connecting for Health* involves the development of several information systems containing patient information, which will be linked to provide a complete record.

- **NHS Care Records Service**: patients will have a single overarching electronic care record, which will hold all of their health information in one place
- **Choose and Book**: GP systems will connect to secondary care providers to enable electronic booking of hospital appointments
- **Electronic transmission of prescriptions system**: this will mean that patients can have a repeat prescription sent electronically to a nominated pharmacy without having to visit the GP practice
- A new national broadband IT network for the NHS
- **Picture Archiving and Communications System** (PACS): for the electronic storage and communication of X-rays, scans, etc.
- **IT supporting GP payment**s: including QMAS (Quality Management and Analysis System)
- Email and directory service

Eng Data linkage requires a unique identifier across both data sets. In England, the ideal unique identifier in health systems is the NHS number (see Box 3B.6.1). An alternative is the National Insurance number (now issued at birth).

However, where a unique identifier is not available, other information such as patient name, date of birth and postcode can be used (see Section 3A.2 and Boxes 3B.6.2 and 3B.6.3).

Eng Box 3B.6.2

Example: 'Combined model' predictive risk algorithm

This model, developed for the NHS, links together several routinely collected data sources (inpatient, outpatient, A&E, GP and Social Services) to make predictions of future emergency hospital admissions. See Section 3C.4.

Aus Box 3B.6.3

Example: Data linkage in Western Australia

Since the 1980s Western Australia (WA) has had a population health data linkage system that is unique in Australia, and one of only five similar systems worldwide. Since its inception more links have been added as its value, particularly to health researchers, becomes increasingly apparent.

The data linkage initiative is a collaborative venture of the WA Department of Health, the Centre for Health Services Research at the University of WA, the Institute for Child Health Research and the Health Science Division at Curtin University. The Data Linkage Unit, part of Information Collection and Management in the Statewide Health Support Service (WA Department of Health), is responsible for creating and maintaining the links within and between the core population health data collections, including births, and mental health, hospital inpatient, emergency, cancer and death records.

Very stringent ethical requirements and strict protocols are applied to the use of linkable data and, to the extent possible, de-identified data files are used. Links to ambulance services and home nursing clients have been added and special arrangements with the Australian government have allowed links to aged care and Medicare records to be created.

www.health.wa.gov.au/icam.

3C
Applications of Health Services Information

Health service information is used in a number of ways to improve health. These include: informing, planning and commissioning; providing key data for epidemiological studies; and identifying unusual or substandard practice through audit. This chapter illustrates how information from Section 1 and other parts of Section 3 can usefully be linked. It contains examples of recent practice in order to demonstrate how the collection and collation of this information may be achieved in practice. The strengths and limitations of health service data are considered, together with the technologies used for their manipulation.

3C.1 USE OF INFORMATION FOR HEALTH SERVICE PLANNING AND EVALUATION

Health information is essential for health service planning and evaluation. Without it, services would be unresponsive to changes in circumstances, and it would be impossible to make predictions or to increase efficiency.

Eng In England, the responsibility for local health service provision lies with PCTs, which plan health services over a 3-year period: see Table 3C.1.1.

Wal Similar arrangements apply in Wales, where responsibility for securing health services lies with local health boards (LHBs). The 22 LHBs are co-terminous with local authorities. This facilitates joint working across the health and social care agenda. However, it also means that arrangements for commissioning secondary care are complex, with a number of commissioning partnerships in place for each hospital.

Scot In Scotland, the health boards, and acute and primary care organisations all work within a single system. Planning and evaluation are led by the health boards with all organisations working together to develop improved systems and solutions.

Table 3C.1.1 Steps in health service planning and evaluation

Step	Description
Priorities	Identifying the national and local priorities and the key targets for delivery over the next 3 years
Capacity	Agreeing the capacity needed to deliver them
Responsibilities	Determining the specific responsibilities of each health and social care organisation
Plans	Creating robust plans which show systematically how improvements will be made and which are based on the involvement of staff and the public
Monitoring	Establishing sound local arrangements for monitoring progress and NHS performance management which link into national arrangements
Accountability	Improving communications and accountability to the public locally so as to demonstrate progress and the value added year on year

INFORMATION FOR HEALTH-CARE PLANNING

Data can be used to assess changes in health and health care over time, thereby enabling demand for services to be forecast: see Box 3C.1.1.

Box 3C.1.1

Health-care **needs**	The identification of health needs and priorities involves epidemiological, qualitative and comparative methods to describe the health problems of a population. These are assessed in terms of **inequalities** in health and in degrees of access to health services. Population trends (e.g. age distribution, relative deprivation) will affect need and these can also be monitored from routine data
Health-care **priorities**	The next step is to determine priorities for the most effective use of resources. This is achieved by reviewing routinely collected national and local data, the literature (both published and the grey literature) and best practice (as defined in national guidance and guidelines). Workforce trends should also be taken into account
Health-care **review**	This begins with a description of the existing service, in terms of utilisation and distribution. An analysis is then undertaken of the differences between needs and existing services (deficient and superfluous services)

ROUTINE DATA SOURCES

See also Section 1A.1.

Eng Many population measures can be obtained from routinely collected data. Several of these are published annually by the Department of Health, with the statistics made available at national, regional, health authority and local authority levels: see Box 3C.1.2.

PERFORMANCE MANAGEMENT

Performance management indicators are used to identify whether services are delivering against a dimension agreed to be of importance for that service. They are used to highlight inadequate performance or problems with service delivery. Where problems are identified, more detailed work may be required to elucidate the underlying causes so that they can be addressed.

Eng **Box 3C.1.2**

Measure	Example
Mortality	Office for National Statistics
Serious morbidity	Hospital episode statistics data Cancer registries
Minor morbidity	GP consultation rates
General health	General Household Survey (self-reported long-standing illness)
Deprivation	Jarman score
Demographics	Census (age, sex, ethnicity)

HEALTH SERVICE EVALUATION

Health service activity may be assessed in terms of **process** or **outcome** measures. Evaluations may be **formative** or **summative**: see Box 3C.1.3.

Box.3C.1.3

Formative	Assesses whether a problem is occurring while the activities are being developed. For example, in a pilot scheme, continuous feedback is obtained from service users and service providers to revise the original plans
Summative	This focuses on the impact and the effectiveness of an established programme. For example, data may be collected over an extended period to assess the impact of a service on a community

Health service evaluations typically consider a number of dimensions. The **impact** and **cost** of services are almost invariably evaluated. However, service quality can also be assessed through a consideration of **acceptability** to service users, **access** to care and the impact of a service on **health inequalities**.

The nature of a health service evaluation often depends on the level of confidence held in the effectiveness of the service in question. If there is high confidence then the evaluation may well be limited to performance management using routine data. If there is low confidence, then the service may be subjected to more rigorous evaluation, e.g. by means of a randomised controlled trial. There are levels of confidence between these extremes. See Box 3C.1.4.

Wal **Box 3C.1.4**

Example: Evaluation of CHD Secondary Prevention Programme

The literature suggests that not smoking and maintaining appropriate levels of body mass index, blood pressure and blood cholesterol can all help to prevent secondary cardiac events. Evaluation of people re-admitted to hospitals for cardiac events in the borough of Rhondda Cynon Tâf showed that, although people had been supported to bring these parameters back to normal at time of discharge home, they rapidly relapsed once discharged home. A new community-based pilot scheme has recently been developed to support people after discharge from hospital. Initially this service will be evaluated to see whether it can help people to maintain healthy blood pressure and cholesterol levels in the longer term (at 12 months post-discharge). If the pilot is shown to be effective and cost-effective, and secures ongoing funding, then performance management will be used to monitor for inequalities in access to the service and drop-out rates by social class.

BENEFITS OF EVALUATION

The prime aim of an evaluation is to determine whether the objectives of a programme are being met. The evaluation of the project should therefore reflect those objectives, but it will be affected by what can readily be measured and by the budget for the evaluation.

Depending upon the findings, an evaluation may justify continuation of a programme or else question whether it should continue as currently provided. If an evaluation is of sufficient breadth and depth, then it may identify changes that could be made to improve the programme, or it may highlight problems requiring more detailed investigation.

A balance must be struck between the benefits and risks associated with the programme, and the costs of the programme and its evaluation. However, it is important to evaluate work – otherwise limited resources will be unlikely to achieve the maximum potential impact. A high-quality evaluation of a programme in one location can often provide sufficient evidence to support rolling out that programme elsewhere – with further monitoring limited to process outcomes.

Evaluation can assess performance with regard to:

- Effectiveness
- Cost-effectiveness
- Efficiency
- Quality outcomes
- Accountability
- Impact on the community's health
- Inequality and other adverse impacts
- Access.

3C.2 SPECIFICATION AND USES OF INFORMATION SYSTEMS

An information system is a process in which raw data are transformed into meaningful information: see Figure 3C.2.1.

Figure 3C.2.1 Simple diagram of an information system

SPECIFICATION

The specification of an information system is the set of requirements agreed between the user and the producer of the system. In modern information systems, information creation and processing require a high level of engagement of all users during the developmental stages.

USES

See Table 3C.2.1.

3C.3 COMMON MEASURES OF HEALTH SERVICE PROVISION AND USAGE

Health service provision refers to the supply of health care (buildings, staff, services), while access is affected by opening hours, demand, capacity, travel times, language and cultural barriers. Health service usage is a complex

Table 3C.2.1 Uses of information systems in health care

Clinical information	To ensure that all health-care professionals have access to the relevant clinical information necessary to support patients under their care
Clinical guidance	To provide health-care professionals with on-line access to up-to-date guidance and evidence on effective treatment, both local and national, and to the information required to evaluate the effectiveness of their work, and to support professional education, development, research and the planning of services
Standardisation	To provide the basis for uniformity of clinical systems, so that comparable recording of information on activity and quality will be available to assist in the maintenance of the highest possible standards of practice
Aggregation	To analyse aggregate data for monitoring of quality (including outcomes), planning new services and supporting research activity
Security	To ensure that sensitive or critical electronic information and systems are not lost, destroyed, misappropriated or corrupted
Exchange	To exchange information securely and automatically and to protect data
Linkage	To manage, link and process the different types of data electronically
Analysis	To analyse, display, report and map accumulated data and share data and technologies for analysis and visualisation with other professionals
Regulation	To conform to national standards recommended by the government or professional bodies, including those for a common clinical terminology and headings for communication from the record

product of provision and access issues and is affected by health service need. The concepts of use and need are more fully explored in relation to equity in Sections 1C.10 and 4C.1.

Measures of health service use can be valuable in assessing the quality and appropriateness of health care provided. Indicators of health service utilisation are also described in Section 1C.3, but those particularly useful in primary and secondary care are shown in Table 3C.3.1.

Table 3C.3.1 Primary and secondary care indicators of service utilisation

Primary care	Secondary care
Consultations per patient Secondary/tertiary care referral rates	Hospital episodes: • Emergency or scheduled/elective • Outpatient appointments or inpatient stays • Re-admission rates
Time spent with practitioner	Length of stay
General practice list sizes	Bed occupancy
Prescriptions: • Number of items • Cost • Type of medication	No comparable data
Preventive health services: • Screening uptake rates • Scheduled immunisation uptake rates	No comparable data

Examples of service usage information in England are shown in Box 3C.3.1.

Eng Box 3C.3.1

Practice profiles: service usage information to improve the quality of primary care

Several PCTs in England provide general practices in their area with summarised information on their performance in a number of domains. In Croydon, the PCT compiles practice profiles, a comprehensive summary for each practice which shows their performance compared with all other practices in the PCT.

Practices receive information on four domains, to enable them to link activity with health service need, behaviour and outcomes:

- Characteristics of patients that may affect their use of primary care services
- Indicators that may be affected by GP clinical behaviour
- Patient survey results
- Disease-specific areas

Each indicator is coloured to give an instant indication of how well the practice is performing. For services with a proven health benefit, e.g. uptake of screening, high activity is marked green, low activity red, with gradations in between. For services with little evidence of clinical benefit, e.g. X-rays for knee problems, high activity is represented by red, with low activity as green.

Excerpt from standard Croydon PCT practice profile

Traffic light indicators 2006 for Practice A

Characteristics of your patients that may affect their use of primary care services

	Percentile Rank	Indicator		Actual Value	Rank in 2005	Rank in 2004
Families	40	AA01	Patients aged under 5 years	5.75	32	55
	20	AA02	Mothers aged under 20 years (new mothers)	2.33	15	6
	26	AA03	Lone parent households with dependents	6.11	28	27
Older People	73	AB01	Patients aged over 65 years	14.3	73	72
	78	AB02	Patients aged over 75 years	7.38	75	78
	54	AB03	Lone pensioner households	11.0	53	53
	95	AB04	Nursing and residential home residents	5.33	98	96
Ethnicity and Immigrants	63	AC01	Asian or Asian British	12.5	65	28
	25	AC02	Black or Black British	5.60	23	22
	15	AC03	Mixed	2.33	14	22
	72	AC04	Chinese or other ethnic group	2.64	71	62
	28	AC05	Born outside the UK	14.6	28	28
	53	AC06	Immigrants from selected countries	6.55	25	36

Indicators that may be affected by GP clinical behaviour

	Percentile Rank	Indicator		Actual Value	Rank in 2005	Rank in 2004
Prescribing Quality	34	BA01	Antibiotics volume	0.18	37	38
	31	BA02	Generic prescribing*	82.9	32	27
	17	BA03	Oral NSAIDS volume	1.86	29	38
	54	BA04	Ulcer healing drugs volume	1.85	42	60
Prescribing Cost	87	BB01	Antibiotics cost	4.71	73	44
	60	BB02	Inhaled corticosteroids cost	0.26	54	30
	60	BB03	Oral NSAIDS cost	0.23	64	76
	32	BB04	Ulcer healing drugs cost	0.48	9	3
	20	BC01	General surgery OPD	77.8	17	26
	25	BC02	Urology OPD	78.9	35	4
	57	BC03	Trauma & orthopaedics OPD	106	78	67
	56	BC04	Ear, nose & throat OPD	100	82	12
	57	BC05	Ophthalmology OPD	96.9	53	72

3C.4 USES OF MATHEMATICAL MODELLING IN HEALTH SERVICE PLANNING

Modelling is the actvity of bridging the real world (where observations of phenomena or behaviours occur) with the conceptual world (where understanding of real world observations occurs). In a model, the observations are analysed using statistical or other analysis, then used to predict events or to provide solutions. An iterative or feedback process allows the model to be refined. See Figure 3C.4.1.

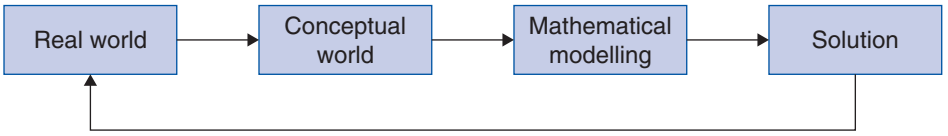


Figure 3C.4.1 Simple concept of mathematical modelling. *Courtesy of Professor E Carson and Dr A Roudsari, Centre for Health Informatics School of Informatics, City University, London*

LIMITATIONS OF MODELS

The usefulness of a model may be constrained by:

- Availability of data
- Quality of data
- Incorrect assumptions in statistical analysis
- Application of an inappropriate or flawed techniques.

ADVANTAGES OF MATHEMATICAL MODELLING

Used properly, mathematical modelling can assist in health-care planning by means of:

- Aiding decision-making
- Dealing with complexity (both complex organisations and complex activities)
- Creating alternative scenarios
- Modelling at any level of detail
- Creating plans for a few months or several years into the future.

PREDICTIVE RISK MODELLING

Predictive risk modelling is a technique that has been used in the financial and banking sectors for many years, and has recently been applied to health care. Predictive modelling involves building an algorithm by examining the historical relationships between known data and a variable of interest (an **outcome**) contained within the data. A model is built either by means of **multiple-regression** or by using **neural network** technology. Once the algorithm has been built, it is used to process contemporary data in order to make predictions about the future. There are six prerequisites to implementing a health intervention based on predictive risk modelling: see Table 3C.4.1.

An example is given in Box 3C.4.1.

Table 3C.4.1 Prerequisites to implementing a health intervention based on predictive risk modelling

Prerequisites	
Sufficient **historical data** must be available to build the algorithm.	Algorithm
The data sources must be **routinely available** and frequently updated	
It must be possible to **link** these data sources	
There must be a meaningful **outcome contained** within one of the data sources	
An **effective intervention** must be available to prevent the selected endpoint	Intervention
The intervention must be **economical**, considering the specificity of the algorithm for the chosen outcome at a particular level of risk	

UK **Box 3C.4.1**

Example: King's Fund/New York University/Health Dialog algorithms

The NHS has commissioned the King's Fund and partners to develop a series of algorithms that predict which people in a population are at highest risk of future multiple emergency hospitalisations. Advance predictions are needed because of the rapid turnover of the group of patients who are frequently admitted to hospital: within 18 months they will have regressed to the mean admission rate for the population.

Several years' worth of historical data were used to build the algorithm. The data sources included inpatient, outpatient, A&E and GP practice data. These were linked using NHS number and the outcome variable chosen was multiple emergency admissions (information that was contained within the inpatient data records). The algorithms allow PCTs to rank their population in terms of likelihood of needing multiple emergency admissions, and an 'upstream' intervention (e.g. case management by a community matron, or admission to a *virtual ward*) is offered to the patients at highest risk.

3C.5 INDICES OF NEEDS FOR AND OUTCOME OF SERVICES

Service indices used by health care are listed in Table 3C.5.1.

Table 3C.5.1 Service indices used by health care

Output index	How much of each service is being produced (e.g. number of patients treated)
Welfare index	Value to final users (e.g. degree of pain reduction)
Performance management index	How the services are being produced (e.g. extent to which doctors use appropriate treatment)
Composite index	This includes elements of the above three indices. With multiple services, weights are added to each service, often combining indicators that are measured in non-comparable units

OUTPUT INDICES

The outputs of a health service may be weighted in a number of ways.

IDEAL VALUE-WEIGHTED OUTPUT INDEX

This index has two fundamental features:

- A value attached to each output reflects its relative contribution to health outcomes; and
- The values of other important characteristics of health care (such as the process of care delivery) are also incorporated.

A lack of health outcome data makes calculation of this index unfeasible.

COST-WEIGHTED ACTIVITY/OUTPUT INDEX

This weights separate activities or outputs using the costs of providing them. The implication of this is that decision-makers are equating costs and benefits of the service provided. There are many arguments for and against this assumption: some might feel that doctors do or should do what is best for the patient whatever the cost. Decisions about capacity may use cost–benefit analysis in a more explicit way, which will affect the decisions made by doctors.

COMPOSITE INDICES

Composite indices of health-care performance are aggregations of several underlying individual performance measures. They are used worldwide to rank health-care organisations or systems: see Box 3C.5.1. They are designed to be easy to interpret and present the 'big picture'. However, there is a need to pay attention to methodological issues in their construction, otherwise misleading conclusions may be drawn. For example, some hospitals may be promoted up the league table of performance as a result of subtle changes in the methods of creating the composite.

Box 3C.5.1

Example: The World Health Report 2000 – *Health Systems: Improving Performance*

In this report the WHO ranked the **overall health system performance** world's health-care systems according to an index composed of five dimensions:

1. Overall population health
2. Health inequalities
3. Health system responsiveness
4. Distribution of responsiveness
5. Distribution of financial burden for the health system

France was ranked top and Sierra Leone bottom. The UK was ranked 18th, Ireland 19th, Australia came 32nd, the USA 37th, New Zealand 41st and South Africa 175th.

Reproduced from www.who.int/whr/en.

PROPERTIES OF PERFORMANCE MANAGEMENT INDICES

Features of an ideal performance index are listed in Table 3C.5.2 and an example is given in Box 3C.5.2.

Table 3C.5.2 Features of an ideal performance index

Fit for purpose	Which index to use will depend on the question being addressed
Consumer's wellbeing	Welfare index is required (of interest to organisations providing the service as well as central government)
Local information	Departments, as providers of services, require performance management indices

UK Box 3C.5.2

Example: Star rating system

Originally introduced in 2000 by the Department of Health, the following *key targets* were used in the star ratings for PCTs in 2004–05:

- Access to a GP
- Financial management
- 4-week smoking quitters
- A larger set of 'balanced scorecard' indicators

The *performance indicators* used in star ratings for PCTs in 2004–05 included:

- Risk management
- Quality of ethnicity data
- Workforce indicators

In 2005–06, the star ratings were replaced by the Healthcare Commission's *'Annual Health Check'* which includes a small set of 'key targets' and 'performance indicators'.

3C.6 ISSUES WITH ROUTINE HEALTH INFORMATION

Strengths, uses, interpretation and limitations of routine health information

Routine data are derived from automated, ongoing, data collection systems. They include large local and national databases associated with health and social services. Data are collected irrespective of the procedure or outcome. Examples include HES, deaths, births, statutory notifications of infectious diseases and cancer registry data. See Figure 3C.6.1.

STRENGTHS AND WEAKNESSES OF ROUTINE DATA

See Box 3C.6.1.

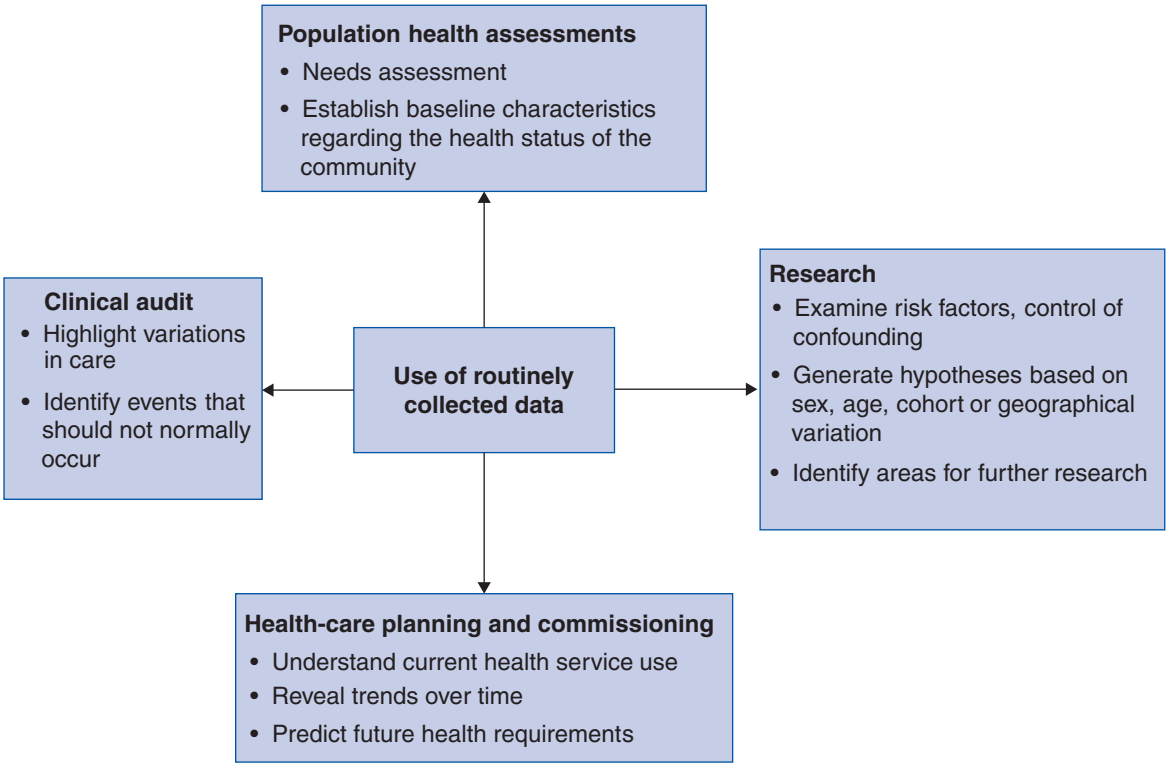


Figure 3C.6.1 Uses of routine data

Box 3C.6.1

Strengths	Weaknesses
Readily available Low cost Up to date Large population coverage Collection usually spans a significant time period Breadth and diversity to explore unexpected avenues Useful for initial assessment: provides baseline data on expected levels of health/disease Simple statistical analysis often sufficient due to size and completeness of database	**Incompleteness:** • Statutory returns do not guarantee completeness, e.g. even meningococcal septicaemia is only 70% notified • Poor levels of ethnicity coding **Bias:** • If those who participate in providing data are significantly different from those who do not, the collated data on health disease may be biased • Political interference in what data are collected (or not collected), and how they are presented Poor collation, presentation and analysis limit value in informing practitioners and providing policy-makers with usable information More information on process rather than health status and outcomes Lack of uniformity in data structures, coding systems and definitions

IMPROVING ROUTINE DATA

Improving the reliability, validity and completeness of routine data is important to avoid waste and to maximise the use of resources. There should therefore be a good reason to begin or to stop collecting each item of data. Ways of improving data quality are listed in Table 3.6.1.

Table 3C.6.1 Ways of improving the quality of data

Computerised data collation and analysis	Improves the accuracy and timeliness of the preparation and dissemination of information
Feedback	Improving feedback of collated data to providers of primary care is essential if their interest is to be maintained and their attention to providing quality data is to be sustained
Presentation	Data should be presented in a variety of ways which are meaningful to policy-makers, media, professionals and the lay public
Training	Training the coders and those responsible for data entry in the use of standard definitions, terminology, etc.

3C.7 INFORMATION TECHNOLOGY AND HEALTH-CARE PROVISION

Use of information technology in the processing and analysis of health service information and in support of the provision of health care

Health systems in many countries are implementing large-scale information and communication technology (ICT) projects to transform the delivery of services. Some important features of electronic communication include:

- Faster and more varied methods for communicating (e.g. email, videoconferencing and mobile telecommunications)
- Near-instant access to vast amounts of information (e.g. through the internet)
- Emergence of new forms of inequality (the so-called *'Digital Divide'*).

Consequently, for public health, ICT presents both developmental opportunities and new challenges to equity. Information underpins:

- Assessment of health **needs**
- Development of health **strategies**
- Monitoring of **progress**.

Communication provides essential links for consultation, discussion and dissemination of knowledge between all those individuals and organisations with a role to play in improving the health of the public. New technologies can offer public health practitioners rapid access to:

- Key data from international down to local levels
- Networks of professionals in health and related disciplines (e.g. managed public health networks)
- The public's views on health service development
- Electronic libraries of evidence, peer-reviewed research and practice guidance (e.g. Medline, Cochrane, NICE).

Sensitive data can be securely transferred between organisations by encrypted email.

HEALTH INFORMATICS

The field of health informatics is concerned with the application of information technology to the acquisition, processing, interpreting, storage, transmission, and retrieval of health and health care-related data. Moreover it uses this knowledge to facilitate health-care delivery, education, management and research.

Health informatics tools include computers, clinical guidelines, diagnostic and monitoring equipment, and information and communication systems.

In practice, health informatics may be broadly divided into: public health informatics, clinical informatics, nursing informatics, dental informatics, bioinformatics and pharmacoinformatics.

APPLICATIONS

The uses of informatics within health care are expanding rapidly. Current applications include:

- Patient monitoring
- Clinical care
- Geographical information systems in public health surveillance
- Integrating data sources for improved decision-making
- Electronic health records (especially in general practice)
- Remote consultations (especially in dermatology).

UK NATIONAL IT PROGRAMME

The Connecting for Health project is reportedly the world's largest ever IT investment. The programme has four goals:

1. Electronic **appointment** booking
2. Electronic care **records** service
3. Electronic and fast transmission of **prescriptions**
4. A fast, reliable, interconnected IT infrastructure.

Section 4

MEDICAL SOCIOLOGY, SOCIAL POLICY AND HEALTH ECONOMICS

Public health is concerned with generating high-quality evidence and credible health advice. However, both lay and professional behaviours often diverge from scientific evidence or professional guidance.

Two disciplines help explain what people actually do and why. Sociology describes the rules and processes of groups – communities, cultures and organisations. Economics provides insight into how decisions are made in a world with scarce resources and infinite needs.

Section 4 explores different concepts of health, and the factors that underlie the ways that people seek and use health care, in order to influence behaviours. The distinctions between equity and equality are discussed, together with their impacts on health policy. Finally, the section on health economics provides the background and core principles of the discipline to equip practitioners with the insight to lobby more effectively for equitable and cost-effective approaches to health-care provision.

4A
Health and Illness

Reaction to illness has certain common features across cultures and some important differences. This chapter explores the **social role** of illness, together with the norms and behaviours that are tacitly expected of those who are ill. In **disability and handicap** and **stigma** the chapter touches on the preconceptions surrounding health and ill health. Finally, it considers how personal characteristics and position in society affect how, and whether, people seek help for their symptoms.

4A.1 HUMAN BEHAVIOUR

Theoretical perspectives and methods of enquiry of the sciences concerned with human behaviour

The social sciences aim to understand the attitudes, motivations and behaviours of human social behaviour and why these change over time. Society is a group of interacting people who share a geographical region, a sense of common identity and a common culture. As such it is more than an aggregate of individuals. The social sciences encompass the fields of study in Box 4A.1.1.

Box 4A.1.1

Psychology	Study of individuals' mental processes and behaviour
Sociology	Study of social processes and interactions in societies, groups and institutions. Sociology recognises that people in societies may behave in ways that differ from the behaviour of individuals
Anthropology	Study of human cultures
History	Recording and interpretation of past events

These disciplines are of importance to public health insofar as they can help explain:

- Individual behaviour patterns
- Behaviour of groups within a population
- Behaviour of health-care organisations.

Data from social research may be **quantitative** (numerical data) or **qualitative** (textual data), although in practice most research considers data of both types.

QUANTITATIVE METHODS

See also Sections 1A.25–1A.29.

This asks questions such as *'How many?'* and *'What proportion?'* Examples include:

- Questionnaires
- Surveys (face-to-face or telephone)
- Routine data sources (e.g. mortality data).

As with all quantitative research, three potential causes of error that should always be considered are:

- Chance
- Bias
- Confounding.

QUALITATIVE RESEARCH

See also Section 1D.

This asks questions about *'How?'* and *'Why?'* Methods include ethnography, interviews, focus groups and case studies: see Box 4A.1.2.

Box 4A.1.2

Ethnography	This is an anthropological research method in which the investigator studies a group's behaviour in great detail, often by living among them for a protracted period of time
Interviews	This can either be a semi-structured interview (loose set of questions) or an in-depth interview (respondent guides the conversation)
Focus groups	The researcher brings together a small group of up to 15 people. The researcher uses in-depth interview techniques and is interested in the group's opinions and deliberations
Case studies	Multiple data-collection methods are used to generate a rounded picture of a 'bounded system', i.e. an organisation fixed in place and time. An example might be a particular GP surgery in 2007

TERMINOLOGY

The theoretical terminology shown in Box 4A.1.3 may be encountered in the social sciences literature.

4A.2 ILLNESS AS A SOCIAL ROLE

The concept of illness as a social role introduces the notion that people who feel ill, and those who care for and treat them, behave in ways that are related to society's implicit ideas of what it means to be sick. The American sociologist, **Talcott Parsons**, described this as 'the sick role'.

Box 4A.1.3

Epistemology	This is the study of knowledge: its origin, nature, methods and limits. Epistemology is concerned with what it means to be *'true'* or *'false'*, what constitutes valid *'information'*, and whether information is absolute or relative
Ontology	Ontology is the study of being. It considers whether facts are constructed in people's minds or whether they exist in an external world
Positivism	Social scientists who advocate positivism value the scientific method. They believe that the social world can be studied in the same way as the material world: hypotheses can be tested according to observable facts. Positivists often employ a quantitative approach
Constructivism	This philosophy is based on the premise that our understanding of the world is constructed by reflecting on our experiences. Each of us generates our own *'rules'* and *'mental models'* which we use to make sense of our experiences
Reflexivity	This acknowledges that, through the process of observing, researchers always affect the environment that they are studying

THE SICK ROLE

Parsons (1951) wrote that people who are ill have rights and responsibilities that work together in the interest of society: see Box 4A.2.1.

Box 4A.2.1

Rights of the sick	Responsibilities of the sick
Exemption from blame for having their illness	Duty to seek medical assistance
Exemption from normal responsibilities such as work	Duty to want to get better

These rights and responsibilities are all both **temporary** and **universal**.

Strengths and weaknesses of this approach are listed in Box 4A.2.2. An example of its use is presented in Box 4A.2.3.

Box 4A.2.2

Strengths	Weaknesses
Applies well to **acute** infections, e.g. cold, flu	Applies less well to **chronic** conditions (or some acute illnesses) where: • Individuals can be 'blamed' for their 'illness', e.g. obesity, sexually transmitted infections • There is no need for individuals to be exempted from normal duties, e.g. controlled diabetes, asthma • Medical assistance not always perceived as 'helpful', e.g. Huntington's disease, inoperable cancers • A duty to 'want to get better' is part of the 'sick role', e.g. one of the symptoms of schizophrenia is a lack of insight into the condition

Box 4A.2.3

Example: The sick role in a New Guinea village

Gilbert Lewis's ethnographic descriptions of sickness in a New Guinea society highlight variations in the way that different cultures treat those with sickness. In his account, both western medical approaches and local rituals are used to attempt to cure a sick man. The source of illness is sought in his own previous behaviour (fights, disputes with others); the spiritual cures attempted include crucifixes round the bed and a Malyi ceremony. Many of these behaviours contrast with the Western ideas of the sick role but also emphasise that a 'sick role' is not confined to Western cultures.

Reproduced from Lewis (2000).

THE DOCTOR–PATIENT ROLE

Doctors often face a conflict between acting in their patients' best interests and serving the wider interests of society. For example, if a doctor saw a patient who worked as a lorry driver and the patient reported having had a blackout, then the doctor would be obliged to inform the *Driver and Vehicle Licensing Authority (DVLA)*, thereby jeopardising the lorry driver's livelihood.

Scambler (1997) describes the traditional doctor–patient role as '*paternalistically doctor centred*' but in recent years there has been a shift in some countries towards more patient-centred care. See Box 4A.2.4 and Figure 4A.2.1.

It is increasingly recognised that patients and professionals each have their own area of knowledge and expertise, and that both parties benefit from working together.

UK The NHS is promoting such cooperation by means of the **Expert Patients Programme** (www.expertpatients.nhs.uk), which uses lay facilitators to empower patients to make the most of their contact time with professionals and to engage in shared decision-making.

See Box 4A.2.5.

Box 4A.2.4

Patient centred	Doctor centred
Consultation style: • Open questions Focused on: • Patients' experience of illness • Reaching concordance	Consultation style: • Closed questions • Disease centred Focused on: • Reaching a diagnosis • Patient compliance

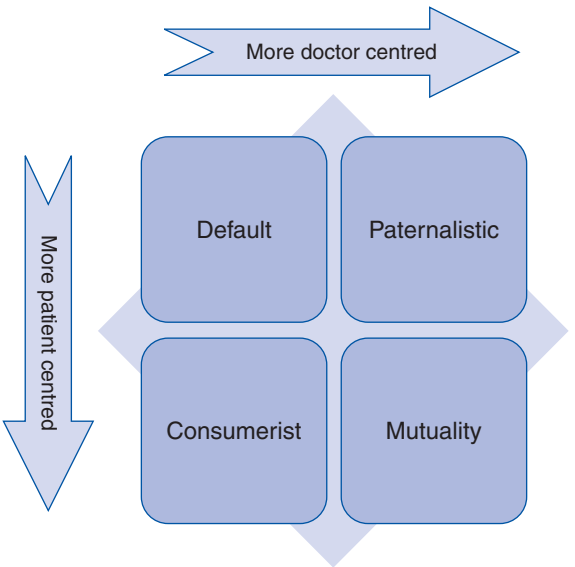


Figure 4A.2.1 The doctor–patient role

Eng **Box 4A.2.5**

Example: Combating drug taking and deviant behaviour through Drug Testing and Treatment Orders

Use of illicit drugs is a prime example of deviant behaviour in Western societies, and it is often linked to a range of other deviant behaviours (e.g. stealing, prostitution). However, it is also recognised in many Western cultures as an addiction – and, as such, an illness. In England, the programme of Drug Testing and Treatment Orders (DTTOs) for drug misusing offenders explicitly links this deviant behaviour with the sick role. Those arrested who have a history of taking illicit drugs are *required* to attend intensive treatment and rehabilitation for their drug use in an attempt to break the cycle of drugs and crime. The DTTOs are a community sentence and are provided as an alternative to prison (the traditional way of dealing with deviant behaviour).

For more information on DTTOs and their effectiveness, see: Comptroller and Auditor General (2004) The Drug Treatment and Testing Order: early lessons. London: National Audit Office. Available online at: www.nao.org.uk/publications/nao_reports/03-04/0304366.pdf.

4A.3 CONCEPTS OF PRIMARY AND SECONDARY DEVIANCE

Unacceptable behaviour within a particular culture is known as **deviance**. Behaviour seen as perfectly acceptable in one culture may be regarded as unacceptable in another (Scambler 1997). On being recognised, deviant behaviour is subject to sanctions, punishment, *'correction'* or *'treatment'*.

In medicine, deviance has labelling implications with regard to organic and psychiatric disease. Parsons (1951) considered illness as a form of deviance where the doctor is an agent of social control (i.e. the doctor *'restricts'* access to the sick role by determining who is sick and who is healthy).

PRIMARY AND SECONDARY DEVIANCE

Lemert (1967) differentiated between **primary** and **secondary** deviance: see Box 4A.3.1.

Box 4A.3.1

Type of deviance	Description	Example
Primary	Relates to the deviant **behaviour**	Rape
Secondary	Relates to the deviant **status**	Rapist

It is important to be aware of secondary deviation because society's reaction to labelling can sometimes hamper treatment and thereby reinforce the deviant behaviour.

4A.4 STIGMA AND HOW TO TACKLE IT

A stigma is a mark of disgrace or infamy. The American sociologist Erving Goffman (1963) defined stigma as:

> *'An attribute that is deeply discrediting within a particular social interaction.'*

This underlines two features of stigma: that it is an **undesirable characteristic** in a particular **context** (i.e. in a particular time and society). For example, in the 1950s in England, knowledge that a person was born outside of marriage would have been considered shameful, but in twenty-first-century England, it is considered normal.

In the medical literature, stigma is used to describe diseases or conditions that lead to exclusion from society: see Box 4A.4.1.

Box 4A.4.1

Examples of stigmatised diseases

- Leprosy
- Psychiatric illnesses
- Epilepsy
- HIV/AIDS
- Sexually transmitted infections (STIs)

CAUSES OF STIGMA

Stigma is rooted in ingrained cultural norms. Stigma thrives on inequalities, fear and misinformation: see Box 4A.4.2.

Box 4A.4.2

Inequalities	Women (e.g. HIV-positive women in Africa) Marginalised groups (e.g. homosexual men, transgender people, prostitutes)
Fear	Fear among the public of having to deal with a person having a fit is often a cause of the stigma attached to epilepsy
Misinformation	There is a popular misconception that the term 'schizophrenia' means dual personality, whereas in fact schizophrenia is characterised by impaired social functioning, distorted thought and hallucinations

The media and religious groups may perpetuate these stigmas, and in so doing will bolster those who are not currently in a stigmatised group.

CONSEQUENCES OF STIGMA

Stigma affects individuals and society, and is manifested as either **expressed** stigma or as **enacted** stigma: see Box 4A.4.3.

Box 4A.4.3

Felt stigma	Enacted stigma
Shame and guilt	Loss of job
Self-stigmatisation	Compulsory testing
Depression	Violence
Unwillingness to speak up	Quarantine
Withdrawal from society	Denial of health services

TACKLING STIGMA

Tackling stigma benefits stigmatised individuals and society as a whole. For example, reducing the stigma of STIs removes barriers to diagnosis for people with genitourinary symptoms. Early treatment of such symptoms benefits the individual (e.g. reduces the risk of infertility) and reduces the spread of the infection among the population.

Tackling stigma requires changes in the attitudes and behaviours of both the stigmatised and society at large. Ways of tackling stigma include those shown in Box 4A.4.4.

Box 4A.4.4

Measure	Description	Example
Education	Public education by means of challenging negative stereotypes and raising awareness of illness	World AIDS day initiatives
Language	Challenging the language that is used to describe illness	Promoting the term 'people with schizophrenia' instead of 'schizophrenic' because the illness does not define a person's entire identity
Public acknowledgement	Public acknowledgement of illness by celebrities	In 1985, Rock Hudson publicly declared that he was homosexual and that he was dying from AIDS
Treatment	Advances in the management of illness	Newer antipsychotic therapies that do not produce parkinsonian-like symptoms reduce the visible marks of illnesses such as schizophrenia
Legislation	Certain manifestations of stigma can be outlawed	The UK's Disability Discrimination Act provides a legal framework for promoting the rights of disabled people

A 'virtuous circle' may be established where positive challenges to stigma can change attitudes and thereby reduce felt stigma. This in turn can reduce enacted stigma: see Figure 4A.4.1.

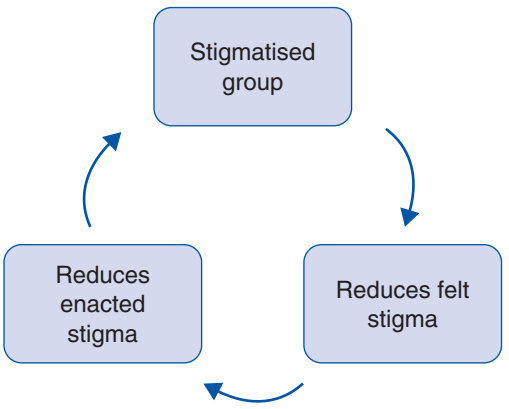


Figure 4.4.1 A virtuous circle to challenge stigma

4A.5 DISABILITY AND HANDICAP

See also Section 1C.

In 1980 the WHO defined impairment, disability and handicap as distinct but interrelated concepts: see Box 4A.5.1. These have since been updated, see Table 1C.7.1.

Box 4A.5.1

Concept	Impairment	Disability	Handicap
Definition	A loss or abnormality of a body function (anatomical, physiological or psychological)	An inability or restricted ability to perform an activity (in the normal human range)	A disadvantage due to impairment or disability that limits the role of an individual
Description	Malfunctioning body parts or systems	Activities a person cannot do	Social sequelae of impairment or disability
Example	Bilateral above-knee amputations	Unable to walk: wheelchair bound	Unable to mingle while socialising in bars

MEASURING DISABILITY

In hospital settings, the Barthel 'Index of Activities of Daily Living' (ADL) score is generally used to assess disability. Patients' independence is assessed in 10 domains: see Box 4A.5.2.

Box 4A.5.2

Domains of the Barthel index	
Bowels	Transfer
Bladder	Toilet use
Feeding	Walking
Grooming	Stairs
Dressing	Bathing

The index is scored out of 20 and describes the level of support that will typically be required (e.g. a patient scoring 10 will require a maximal package of home care; patients with scores >10 will need residential or nursing home care).

MEASURING HANDICAP

Various tools exist for assessing handicap, including:

- **Rankin scale** (used in stroke research)
- **Hearing Handicap Inventory**
- **London Handicap Scale** (domains include mobility, physical independence, occupation, social functioning and economic self-sufficiency).

THE SOCIAL MODEL OF DISABILITY

In contrast to the above indices, the **social model** of disability considers how society disables those with physical impairments by means of a wide range of barriers. These barriers are both environmental (e.g. no wheel ramps at entrances to buildings) and cultural (e.g. patronising attitudes towards people with impairments).

Eng In England, the Disability Rights Commission (now replaced by the Equality and Human Rights Commission) describes how the:

> *'Arrangement of transport, leisure facilities, public services and work excludes disabled people and how people's attitudes also demean and isolate.'*

4A.6 SOCIAL AND STRUCTURAL IATROGENESIS

In his book *Medical Nemesis,* Ivan Illich (1975) introduced the concept of **iatrogenesis**, i.e. disease caused by medicine. He described three ways in which medicine can cause illness: **clinical**, **social** and **structural**.

CLINICAL IATROGENESIS

Medical treatment sometimes worsens the original illness or creates a new illness. Examples include:

- Adverse drug reactions
- Diabetogenic drugs (e.g. steroids, certain combinations of antihypertensive drugs)
- Hospital-acquired ('nosocomial') infections.

SOCIAL IATROGENESIS

This describes the way in which medicine invades normal life: see Box 4A.6.1.

Box 4A.6.1

Aspect of life	Medicalisation
Normal childbirth	Caesarean sections on demand
Ageing	Cosmetic surgery
Unruly children	Attention-deficit hyperactivity disorder

This is reflected by the growing proportion of the GDP that is spent by many countries on health care.

STRUCTURAL IATROGENESIS

This is the impact that the medical profession has upon a population. As a result of increasing reliance upon medicine, the public has lost its traditional ways of coping with illness, death, pain and misfortune.

> *'The so-called health professionals have an even deeper, structurally health-denying effect insofar as they destroy the potential of people to deal with their human weakness, vulnerability and uniqueness in a personal and autonomous way.'* (Illich 1975, p26).

4A.7 ROLE OF MEDICINE IN SOCIETY

The role of medicine has been described as being:

> *'To cure sometimes, to heal often and to comfort always.'* (Variously attributed to Hippocrates, to a fifteenth-century French proverb, and to Sir William Osler)

In particular, the role of biomedicine (i.e. branch of medical science that applies biological and physiological principles to clinical practice) has become an increasingly dominant – though not always unchallenged – part of western society.

EXPANDING BOUNDARIES OF MEDICINE

Health-care professionals and professional organisations are now routinely involved in areas of life that were previously outside the scope of medicine. Key examples of the role of medicine in areas previously not the preserve of health care are:

- Childbirth
- Euthanasia
- Abortion.

CHALLENGES TO THE ROLE OF MEDICINE

The place of medicine in Western society is not undisputed. Some of the areas where it is frequently challenged are shown in Box 4A.7.1.

Box 4A.7.1

Clinical iatrogenesis	Well-publicised accounts of side effects and poor outcomes as a result of medical care have led some commentators to question the dominant role of medicine in society See Section 4A.6
Anti-psychiatry	More than other medical disciplines, psychiatry has been challenged. Many commentators, but especially RD Laing in the 1960s, have questioned the whole notion of mental illness. In western societies, however, psychiatry remains the dominant, if contested, model of care for mental ill health
Health literacy	Growth in consumer access to health information (e.g. via the internet) is removing the absolute dependence of the public on the opinions of doctors. Medical opinion becomes only one of many sources of information
Complementary and alternative health care	Other approaches to treating illness are available, and many such treatments are growing in popularity

4A.8 SOCIAL PATTERNS OF ILLNESS

Explanations for various social patterns and experiences of illness (including differences of gender, ethnicity, employment status, age and social stratification)

Subjective (i.e. experienced) health differs markedly from objective health. According to the Health Survey for England, approximately 1% of the population regards itself as being in *'very bad health',* with the proportion being correlated to age. A study by Scambler 1997 found that a typical person experiences symptoms of illness approximately one day in three. However, only on about 6% of occasions will the person consult a health professional. Whether the person seeks medical advice is determined by a number of triggers, including:

- Presence of a concomitant crisis
- Sanctioning by others
- Interference with normal activity
- Temporising deadlines.

Some explanations for the varying patterns of illness seen between different social groups include:

- **Biology** and genetics
- **Behaviour** (differences in health seeking or risk taking between groups)

- **Social** circumstances (e.g. disproportionate effects of poverty on health)
- **Artefact** (the categories used to distinguish social groups are in themselves problematic).

GENDER

Clearly, certain diseases affect only men (e.g. prostate cancer) or women (e.g. endometriosis). Others are commoner in men (e.g. renal stones) or in women (e.g. gallstones). Differing behavioural patterns also affect morbidity and mortality, e.g. smoking, dangerous driving.

Although life-expectancy is lower for men, women report more ill health. This is thought to be partly due to different reporting behaviours between the sexes.

ETHNICITY

Certain diseases are commoner in some ethnic groups, e.g. sickle cell disease in black people. Reasons include genetic differences, consanguineous marriage, as well as:

- Poverty: in the UK, minority ethnic groups are mostly less affluent
- Migration: loss of social capital (see Section 4A.10)
- Behavioural: differing patterns of smoking, diet, etc.
- Racism (direct and indirect) affects health through increased stress and social isolation
- Access to health services.

EMPLOYMENT

Fulfilling and secure employment provides not only material resources for individuals and their families, but also psychological and social support. A summary by the Health Development Agency (now part of NICE) underlined the effects of unemployment on:

- **Physical health:** unemployment is associated with mortality (greater suicide rates and cardiovascular mortality)
- **Mental health:** those who are unemployed or in insecure employment are more prone to common mental disorders, e.g. depression.

The relationship between unemployment and health is complex. In some circumstances, health problems may be a factor in losing a job, while, in others, the loss of work may precipitate health problems.

AGE

There are clearly different patterns of mortality at different ages, but experiences of health services and illness also vary with age. For example, surveys indicate that older people are generally more positive about the standard of care that they have received than are younger people.

SOCIAL CLASS

Social class gradients in health occur at every age (e.g. low birthweight) and for all major causes of death.

UK In the UK, the Black Report (1980) found that, despite general improvements in health and prosperity in the UK, there were still pronounced, and possibly increasing, disparities in health and illness across the five social classes. The report was the work of the Department of Health Research Working Group on Inequalities in Health and was led by **Sir Douglas Black**. It suggested four possible explanations for differences in health: see Box 4A.8.1. The response to the report is discussed in Section 4C.10.

Box 4A.8.1

Artefact	The association between social class and health is an artefact of the way in which these concepts are measured
Social selection	Health determines social class through the process of health-related social mobility
Behavioural and cultural	Social class determines health through social class differences in health-damaging or health-promoting behaviours
Materialistic	Social class determines health through differences in the material circumstances of life (e.g. Whitehall studies) This may manifest directly (e.g. accidents), indirectly (e.g. social capital) or through psychosocial mechanisms

An example of the relationship between occupation and health is shown in Box 4A.8.2.

Box 4A.8.2

Example: Occupation and health – Whitehall studies I and II

The first Whitehall cohort study examined mortality rates over 10 years among male British civil servants aged 20–64 in the 1960s and 1970s. This revealed differences in health between different employment grades:

- Men in the lowest grade (messengers, doorkeepers, etc.) had a threefold higher mortality rate than men in the highest grade (administrators)
- Blood pressure at work was associated with 'job stress', including 'lack of skill utilisation', 'tension' and 'lack of clarity' in tasks. The rise in blood pressure from the lowest to the highest job-stress score was much larger among low-grade men than among upper-grade men. Blood pressure at home, on the other hand, was not related to job-stress level

A second longitudinal study of British civil servants (Whitehall II) started in the 1980s and focused on occupational effects on health and disease. The study involved around 10 000 men and women aged 35–55 in the London offices of 20 civil service departments, and many people in the cohort are still being followed up. The study found that employment grade was strongly associated with work control. Lack of control in the job was related to long spells of absence and an increased risk of cardiovascular disease.

Reproduced from www.workhealth.org/projects/pwhiteabs.html.

4A.9 SOCIAL FACTORS IN THE AETIOLOGY OF DISEASE

Role of social, cultural, psychological and family relationship factors in the aetiology of illness and disease

See also Section 4A.8.

These factors are listed in Box 4A.9.1.

Box 4A.9.1

Social	Inequalities, environment, economic
Cultural	Lay health beliefs, explanatory models
Psychological	Health beliefs (perceived costs, risks and benefits of health behaviour) Locus of control – internal, external orientations
Family relationships	Benefits to health of stable relationships Lay referrals (where family members sanction each other to seek professional advice)

4A.10 SOCIAL CAPITAL AND SOCIAL EPIDEMIOLOGY

Although social capital and social epidemiology were both first described in the mid-twentieth century, academic interest in the two concepts only began in earnest in the early 1990s. The concepts are now widely regarded as being important and meaningful across the social sciences.

SOCIAL CAPITAL

Social capital gives a value to the social networks in which individuals live. These networks provide **norms** (i.e. defined limits of acceptable behaviour) and **sanctions** when these bounds are crossed (e.g. social exclusion, gossip, stigma). Social capital functions in two ways:

- **Bonding**: social capital strengthens the links between members of families and tight-knit communities and provides social support
- **Bridging**: social capital strengthens the links with members outside the group, i.e. 'networking'.

Social capital can be considered at the **micro**- (individual), **meso**- (community) and **macro**- (national) levels. At all three levels it is demonstrably correlated with economic affluence, low crime, educational attainment and health. The WHO regards social networks as a determinant of health, and it advocates the provision of social support to improve health outcomes through increased social capital.

SOCIAL EPIDEMIOLOGY

Social epidemiology is the study of the social determinants of the distribution of disease within a population. Multi-level analysis is used to determine to what extent an individual's health is shaped by **micro** (individual), **meso** (household/small area) and **macro** (large area) characteristics.

Macro-level epidemiological factors may be **compositional** (e.g. prevalence of childhood poverty) or **contextual** (e.g. population density) – with the latter being irreducible to the individual.

Social determinants of health are listed in Box 4A.10.1.

Box 4A.10.1

Social determinants of health

- Social gradient
- Stress
- Early life
- Social exclusion
- Work
- Unemployment
- Social support
- Addiction
- Food
- Transport

Reproduced from Wilkinson and Marmot (2003).

4B
Health Care

Most of the discourse concerning health care focuses on formal health systems and the role of professionals as health-care providers. In **different approaches to health care,** the topic is widened to encompass other kinds of care, including self-help and complementary practices. This chapter also considers the social roles and characteristics of health-care providers. In **hospitals as social institutions**, the archetypal providers of health care are considered in terms of their functions in society other than health care and their potential to constrain the actions of individuals. **Professions** looks at how professional status was created and is maintained, and examines the situations where conflicts due to professional roles can arise.

4B.1 ALTERNATIVE SOURCES OF HEALTH CARE

Different approaches to health care (including self-care, family care, community care, self-help groups)

The decisions of whether or not to access health care and which type of health care to access are influenced by a range of factors including:

- Characteristics of the person (age, gender, ethnicity, previous experiences with health care)
- Nature and duration of symptoms
- Accessibility of formal health care (cost, convenience, welcoming attitude).

The term **clinical iceberg** was used by Last (1963) to describe the finding that professional health services treat only a small fraction of the total burden of ill health. It has been estimated that a typical adult experiences some sort of somatic symptom once every 3–6 days (Barsky and Borus 1995). Assuming that some of these symptoms will require health care, it is probable that the bulk of health care occurs in the so-called **informal sector** because patients in Britain visit their GP only four or five times a year (see www.statistics.gov.uk).

SELF-CARE

Self-care (taking care of oneself without professional assistance or oversight) is the commonest form of health care. For a new symptom, such as a cough, a person may typically instigate a management plan such as waiting to see what happens, unless a:

- Symptom changes (e.g. the cough becomes productive) or
- Symptom persists beyond a time limit (e.g. the end of the week).

In these circumstances, the person may choose to consult another person (e.g. a family member or a GP) or else try an over-the-counter remedy (e.g. cough mixture).

Self-care remedies can be orthodox or complementary/alternative. Orthodox medicines in the UK can be accessed without consulting a clinician for a prescription if they are listed on the pharmacy (P) or general sales list (GSL). Restrictions apply both to which items may be sold and to the quantities that may be purchased (e.g. 32 tablets of paracetamol on the P list and 16 on the GSL list). See Table 4B.1.1.

UK Table 4B.1.1 Classes of medications and their restrictions

Class of medication	Abbreviation	Restriction	Example
Controlled drug	CD	Can be dispensed only with a detailed prescription. Stored in locked cupboard; closely monitored	Diamorphine
Prescription-only medication	POM	Can be dispensed only with a prescription	Flucloxacillin
Pharmacy list	P	Can be sold only under the supervision of a pharmacist	Ranitidine
General sales list	GSL	Available on open shelves, e.g. in supermarkets	Ibuprofen

LONG-TERM CONDITIONS

Self-care accounts for the vast majority of treatment of any long-term medical condition. For example, the Department of Health has calculated that a typical patient with diabetes has 3 h of contact time with professionals per year, and will therefore self-care for the remaining 8757 h (see www.dh.gov.uk/assetRoot/04/10/17/02/04101702.pdf).

The population of people with a long-term condition is often represented as a triangle (called the *Kaiser pyramid*) (Figure 4B.1.1), which depicts the large number of patients at the base of the triangle who have straightforward conditions, rising up through the triangle to the small numbers at the top with complex disease.

The proportion of self-care undertaken by patients varies with the complexity of their condition: see Figure 4B.1.2.

COMPLEMENTARY OR ALTERNATIVE THERAPY

Complementary/alternative treatments can be delivered as self-care or by practitioners. Increasingly, complementary practitioners share many of the attributes of health-care professionals (see Section 4B.3). Complementary therapies range in complexity and the degree to which they can be self-administered: see Box 4B.1.1.

The use of complementary therapies can differ somewhat from conventional medicine. Their uses are not restricted to the treatment of specific health complaints, but are also employed by the healthy for the maintenance and promotion of wellbeing.

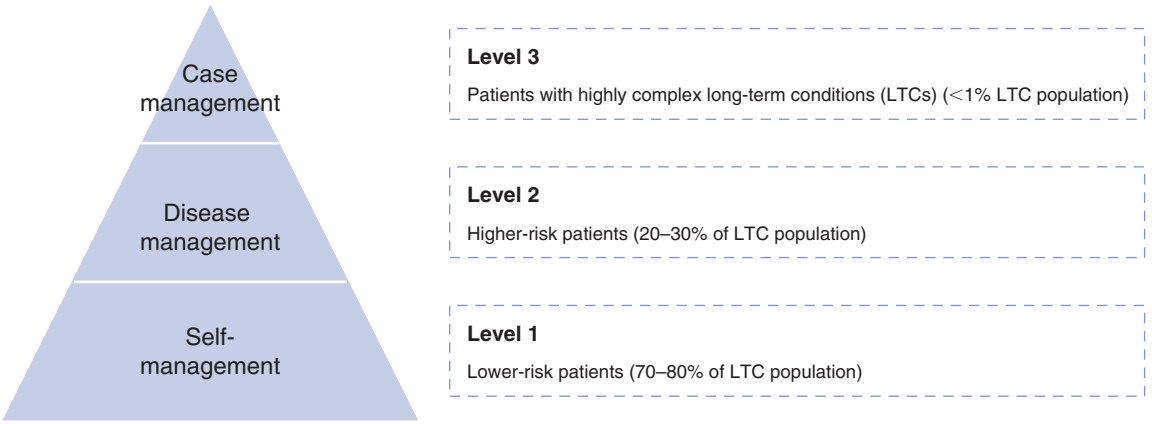


Figure 4B.1.1 Kaiser pyramid. *Reproduced from www.natpact.nhs.uk/cms/336*

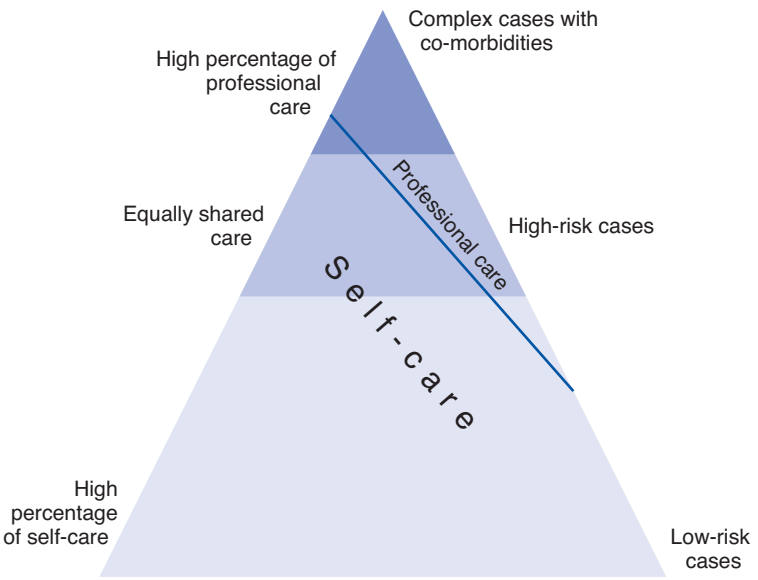


Figure 4B.1.2 Pyramid showing proportion of patients who self-care, under professional care and mixed according to complexity of their condition. *Reproduced from www.dh.gov.uk/assetRoot/04/10/17/02/04101702.pdf*

Eng The Health Survey for England asks about the use of complementary and alternative medicines among adults (see www.ic.nhs.uk). Data from 2004 indicate that:

- 40% of the general population have used some form of complementary and alternative medicine, though women (51%) are more likely to use them than men (21%)
- the most commonly used forms were **osteopathy** (11%), **massage** (10%) and **acupuncture** (8%)
- over 90% of osteopathy and acupuncture, and 70% of massage therapies, were delivered by a practitioner. In contrast, less than half of all those surveyed who used herbal medicine, aromatherapy and meditation visited a practitioner for this therapy.

Box 4B.1.1

Examples: Complementary therapies

Homemade:

- Honey and lemon in hot water for a cold
- Avocado pulp as a skin moisturiser

Commercial:

- Ginseng for energy
- St John's wort to combat depression

Self-administered practices:

- Meditation

Therapies delivered by a practitioner:

- Massage
- Acupuncture

FAMILY CARE

Friedson (1959) described the way in which people tend to discuss medical issues with family, friends or colleagues before seeking professional advice. This **lay referral system** is used for:

- Interpretation of symptoms
- Reassurance
- Advice about a remedy
- Advice about referral to another lay person or professional.

Where lay culture and professional culture differ, a ladder of consultations begins with the nuclear family, through progressively more distant and authoritative lay people, until the professional is reached. In contrast, where lay and professional cultures are more alike, patients typically take a great deal of time trying to treat themselves, and then go directly from self-treatment to consulting a doctor.

COMMUNITY CARE

As well as forming an important part of the lay referral system referred to above, the term *community care* has taken specific meanings in certain countries:

- UK In Britain, the National Health Service and Community Care Act 1990 led to a large-scale relocation of people with mental illness from large psychiatric hospitals into small local hostels or sheltered housing.
- Aus In Australia, the term relates to a joint Commonwealth, State and Territory initiative in which frail older people and people with disabilities are given funding to support themselves in continuing to live in the community.
- SA In South Africa, the term is typically applied to mean home-based care and support for people living with HIV/AIDS.
- Ire In Ireland, the term includes public health nursing, home help, and out-of-hospital physiotherapy, occupational therapy, chiropody service, day care, respite care service, etc.

SELF-HELP GROUPS

Self-help groups now exist for almost every conceivable medical condition – literally from achondroplasia to XXY syndrome. Typically these societies exist to:

- Provide information to patients and their carers through leaflets, helplines and websites
- Put people in touch with others who are affected with that condition
- Raise funds in order to commission research into the condition
- Lobby government and clinicians.

The rise in self-help groups in recent times is attributed to a general desire by patients to take more responsibility for their own health. As treatment options become more complex, and time with clinicians more limited, self-help groups have filled the gap between professionals' availability and the demand for disease-related information and support. Their expansion has been facilitated by the rise of the internet and email.

4B.2 HOSPITALS AS SOCIAL INSTITUTIONS

There are various ways that hospitals can act as a social institution, shaping cultures and practices in their community. Hospitals are not just a centre for treatment. They also function positively in many other ways in society, as:

- **Employers**: for clinical staff; and non-clinical, including technicians, caterers, porters, builders, secretaries. Health services are often one of the largest employers in a local area
- **Purchasers**: health care-related products, such as drugs and medical equipment, and other goods, e.g. food and drink for patients, visitors and staff
- **Community resource:** hospital resources can be important for the local community, where rooms in the hospital building are used for public meetings, hospital sports, catering, arts facilities.

There are negative ways in which a hospital can function, too:

- **Polluter**: travel to and from hospital by staff, patients and visitors adds significantly to congestion; hospital waste makes an impact on carbon emissions
- Means of **isolation or exclusion**: at its extreme, there is an intentional element to separating patients with infectious diseases from others. However, in general hospitals serve as a means of separating people from the rest of society until they are healthy.

Box 4B.2.1

Example: Eng NHS as a corporate citizen

The Sustainable Development Commission is working with the NHS to improve its performance as a corporate citizen. It highlights the market power of the NHS as a purchaser and its potential to exercise significant leverage to influence practice. As a consumer the NHS spends around £17 billion a year. Every year it buys:

- 1.3m chicken legs
- 12.3m loaves of bread
- 13.5m kg of potatoes
- 250 000 l of orange juice

Reproduced from www.corporatecitizen.nhs.uk.

As a responsible corporate citizen (see Box 4B.2.1), the NHS can engage in a range of processes with wide-ranging potential benefits, covering transport, procurement, facilities management, employment and skills, community engagement and new buildings.

Goffman (1963) defined **total institutions** as places where people are isolated from society over a period of time and lead life in an enclosed and formally administered way. He went on to consider the effects of institutionalisation on social relationships in the outside world, the ways in which people adapt to and become attached to the institution, and the complicit role of medicine in the process. See Box 4B.2.2.

Box 4B.2.2

Erving Goffman: asylums

Goffman used **participant observation** to study the inner workings of an American psychiatric hospital. Working as a hospital porter, he observed the effects of institutionalisation – notably the staid behaviour patterns of both staff and patients.

Goffman found that institutionalised patients were apathetic and became progressively less able to make decisions and care for themselves. Ways have subsequently been found for avoiding these negative effects, including:

- Providing information to patients prior to treatment
- Promoting mobility and self-care while in hospital
- Pre-discharge planning and education
- Reducing length of stay
- Community care

4B.3 PROFESSIONS

Professions, professionalisation and professional conflicts

See also Section 5A.11.

The three original professions – clergy, medicine and law – have been characterised by the following traits:

- Specialist area of knowledge
- A professional association
- Ethical code
- Control over certification or licensing.

Medicine is often used as a model for the study of professions. In contrast, the status of nursing as a profession has been controversial and taken much longer to establish.

PROFESSIONS

Talcott Parsons in the 1950s described various aspects of the medical profession and its effects on the relationship with patients:

- A clearly defined **knowledge base** that is highly developed, theoretical and specialised
- Patients expected to defer to doctors' **authority**
- Self-governing and **self-policing**
- Potential to **exploit power** over patients for financial gain
- A commitment to **public service** and ethics
- **Protection** for patients against exploitation.

In the 1960s, nursing was often described as a 'semi-profession' because it lacked the powers of self-regulation and a specialised body of knowledge.

PROFESSIONALISATION

Sociologists have studied the emergence of the professional status in medicine. Professionalisation tends to involve establishing:

- Acceptable qualifications (e.g. in medicine, Larson [1977] argues that the introduction of a university-based medical degree improved the credibility of the profession)

- A professional body (e.g. General Medical Council) to oversee the conduct of members of the profession
- Occupational closure (where there is no entry from outsiders, amateurs or the unqualified).

Successful professionalisation protects members of an occupation from external control, and acquires specialised knowledge, monopoly and **autonomy** for the profession which is guided by a code of professional ethics.

PROFESSIONAL CONFLICTS

Parsons' functionalist account of how medicine works as a profession has since been challenged. Sociologists started to question whether the espoused principles, e.g. of altruism, coincided with what doctors actually *did*. Feminist critiques of the profession focused on the gendered nature of the profession:

- In the 1970s, professions (such as medicine) were largely **populated by men**
- Professions seem to espouse '**traditional masculine' values** (technical expertise, rationality)
- Entrance to a profession was geared more towards the **opportunities** offered to men rather than women.

The medical profession is currently prone to conflicts in a number of areas:

- With **nursing and allied health professions** regarding disputes over professional boundaries
- With **complementary practitioners** over recognition
- With **management** regarding issues of professional autonomy and self-control
- With **patient groups** regarding issues of consumerism and paternalism
- With **government** over terms and conditions of employment, and health service policy.

4B.4 CLINICAL AUTONOMY

Role of clinical autonomy in the provision of health care

As discussed in Section 4B.3, autonomy (control over terms and conditions of work – clinical, financial, organisational) is a key attribute of professions. Clinical autonomy refers specifically to the control over content and delivery of health care.

In the 1960s, 1970s and 1980s, sociology focused largely on how the medical profession had achieved its autonomy. Since the 1990s, the focus has shifted towards threats to the medical profession's autonomy. Threats to clinical autonomy include:

- Increasing **role of management** in health care, in setting clinical priorities, monitoring standards of care
- **Cost containment**: since all health-care decisions are spending decisions, cost-containment is necessary – explicitly or implicitly – to safeguard resources. Explicit cost containment rules can limit clinicians' capacity to make decisions based on the characteristics of the individual
- **Guidelines and protocols**: some clinicians argue that the existence of guidelines and protocols removed the role of clinical judgement from medicine. Alternatively, David Armstrong (2002) views **evidence-based medicine** as the medical profession's response to falling public trust in clinical practice
- Both **external assessment** and **revalidation** open doctors' practice to greater scrutiny and to scrutiny from those outside the profession (see also 5A.11)
- **Market forces**: in a private health-care system, in particular, consumer satisfaction is required in order to remain a viable business.

The structure of health-care systems is often balanced so there is a trade-off between clinical autonomy, financial autonomy and control over terms and conditions of work: see Box 4B.4.1.

Eng Box 4B.4.1

Example: A new contract for GPs

Since its inception in 1948, the NHS has not employed GPs. Instead, they work as small businesses contracted to provide medical services, and hence have had a large degree of autonomy.

In 2003, the majority of GPs voted to accept a new contract with the NHS for their medical services. The contract in general led to a substantial pay rise for GPs, but the autonomy of the GP has, to some extent, been traded for lesser clinical autonomy and a stronger role for all employed by the practice.

	Earlier GP contract	2003 contract
Clinical autonomy	Contract held between individual GP and Secretary of State	Contract held between **practice** (not just individual GP) and local NHS organisation
Clinical Practice	Contract not dependent on clinical practices	Quality and Outcomes Framework rewards practices for certain clinical behaviours
Terms and conditions	GP responsible for patients 24 h a day	GP can opt out of out-of-hours care
Financial autonomy	GP pay not linked to service provision	Pay related to services provided

Reproduced from www.dh.gov.uk; bbc.co.uk.

4B.5 ILLNESS BEHAVIOUR

Behaviour in response to illness and treatments

Definitions of '*health*' and '*illness*' vary across cultures, communities and households. Sociologists study illness behaviour to learn why people seek or decline professional help.

Mechanic (1968) identified 10 variables that influence illness behaviour:

- Visibility of symptoms and signs
- Perceived seriousness (by the patient) of the symptoms, for present and future probabilities of danger
- Amount of disruption caused by the symptom to work, family, etc.
- Frequency and persistence or recurrence of signs or symptoms
- Tolerance threshold of person exposed to symptoms
- Knowledge, information and assumptions of the evaluator
- Basic needs leading to denial
- Needs leading to competition with illness
- Competing interpretations assigned to symptoms once recognised
- Availability of treatment: access, cost (not only money but also emotional, e.g. stigma).

CULTURAL VARIATION

Pilowski and Spence (1977) noted marked cultural differences between Anglo-Saxon (stoical, withdrawn) and Mediterranean groups (experience) in their interpretation of a response to symptoms and signs. Zborowski (1952) found that Americans of Irish origin had a matter of fact attitude towards pain, whereas people with an Italian or Jewish background were more demanding and dependent on medical help.

PHENOMENOLOGY OF SYMPTOMS

Diseases that present with striking symptoms (e.g. severe pain, jaundice) are more likely to receive prompt medical attention than those that are less dramatic.

LAY REFERRAL AND INTERVENTION

See Section 4B.1.1, *Family care*, a lay person may also intervene to initiate medical consultation, e.g. on behalf of a child, or by calling an ambulance for someone who is having an epileptic fit or chest pain.

4B.6 PSYCHOLOGY OF DECISION-MAKING IN HEALTH BEHAVIOUR

The success of medical practice can be linked to patient behaviour. Understanding the issues that influence patients' decisions when they choose not to follow preventive or therapeutic recommendations is instrumental to improving concordance and, ultimately, to improving health outcomes.

There are a number of models and theories that attempt to offer explanations for why people behave as they do with regard to health: see Section 2H.7.

4C
Equality, Equity and Policy

The NHS in the UK was established to provide equal access for equal need. At first glance, this does not appear to be a contentious or ambiguous aim. As will be seen in the beginning of this chapter, however, the concepts of access, need, equality and equity can be complex and they may conflict with other priorities such as efficiency. The rest of the chapter covers the extent to which equity and equality are present in society and policy, and how they fare in competition with other priorities.

4C.1 NEED AND SOCIAL JUSTICE

Concepts of need and social justice

The concept of **need** relates to the **capacity to benefit**.

Social justice is characterised by the **availability** of **equal** rights, opportunities, obligations and benefits available to all members of the population.

NEED

Given that the resources available to any health service are finite, the issue of determining the relative needs of patients will always be crucially important. In terms of allocating health-care resources, Bradshaw (1972) differentiated need into four categories: see Table 4C.1.1.

Table 4C.1.1 Bradshaw's categories of need

Type of need	Description	Measurement
Felt need	This is also known as a **need for health** and relates to an individual's subjective experience of feeling unwell. Felt need does not necessarily relate to health service use. For example, someone with a headache may well report that she is in pain but may seek no health service intervention	Felt need can be measured through **surveys**, e.g. in the UK, the **census** asks about experience of illness
Expressed need	This is synonymous with **demand**. Expressed need occurs when a patient seeks health care for a felt need	**Waiting lists** can be used as a proxy measure of expressed need
Normative need	This is also known as a **need for health care** and relates to a professional's judgement of an individual's health state. A professional's assessment of whether an individual has a normative need depends not only on the experience of symptoms but also on a range of other factors, such as whether an effective treatment exists, whether the treatment is available and whether the patient is in a suitable state to benefit from the treatment. Normative need is therefore dependent upon the health system in place and the technology available. In public health, practitioners may identify a normative need where individuals may have no felt need (e.g. recommendation for weight loss in someone with a high BMI who has no wish to change their size and lifestyle)	**Health needs assessment**
Comparative need	This is also known as **relative need** and describes the process of comparing the services available for areas with similar prevalence of disease and epidemiology. If one has a greater service provision, the other area could be said to be in relative need of health services. Comparator areas could be neighbouring boroughs, or the country as a whole	**Deprivation** and **mortality** measures may be useful for indicating a need for health but may not reflect a need for health care

In certain circumstances these needs may overlap, as is illustrated in the Venn diagram in Figure 4C.1.1.

SOCIAL JUSTICE

Humans come together as societies for their mutual benefit. However, for societies to function, they must curtail the freedoms of each individual for the purpose of the greater good. Theories of social justice address:

- The extent to which individuals' freedoms are (or should be) restricted
- How decisions are made in a society (and by whom)
- What constitutes 'the greater good'.

Three major theories of social justice are **utilitarianism**, **distributive justice** and **procedural justice**.

UTILITARIANISM

Utilitarian philosophy as expounded by the eighteenth-century philosopher, Jeremy Bentham, stipulates that individuals and societies should make their choices with the aim of achieving the **greatest good for the greatest number**. Utilitarianism forms the basis for many economic and efficiency arguments within public health – but it dismisses the principles of distributive justice and equity.

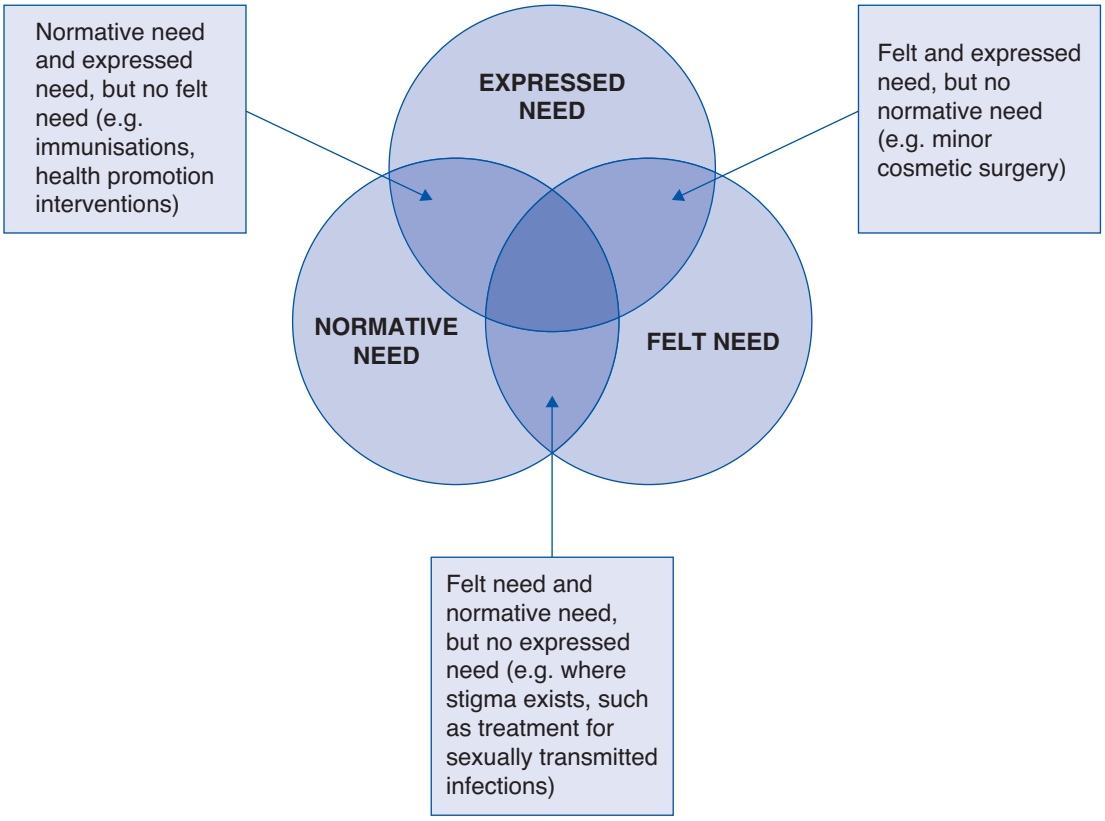


Figure 4C.1.1 Venn diagram illustrating Bradshaw's overlapping categories of need. *Adapted from Bradshaw (1972)*

DISTRIBUTIVE JUSTICE

The twentieth-century philosopher, John Rawls, set out a theory of *'Justice as Fairness'* that outlined two principles for achieving a fair society: see Box 4C.1.1.

Box 4C.1.1

Basic liberties are a right for everyone	Restricting the individual liberties of some members of society is not justified even if it could lead to the greater good of society
Difference principle	Resources need not be distributed equally, but social and economic inequalities should benefit those who are most disadvantaged. Society can agree to pay a doctor, for example, more than a cleaner because the doctor has the potential to save people's lives

Rawls proposed a model for societal decision-making named the **veil of ignorance**. In this model, those who make resource allocations should do so from a stance of ignorance regarding their personal future position in society. A health-care system supported by progressive taxes is an example of justice as fairness because, in general, individuals pay tax with no knowledge of what health-care resources they will need in the future. On this basis, society has deemed that those currently earning more should pay more tax in order to provide resources for those who are currently unable to pay.

PROCEDURAL JUSTICE

While distributive justice is concerned with the **outcomes** of society in distributing rights and resources, procedural justice also examines the **ways** in which those outcomes are reached. Three models of procedural justice are based on **outcomes**, **balancing** costs and outcomes, or **participation**: see Box 4C.1.2.

Box 4C.1.2

Outcomes	Society should adopt procedures that produce the **fairest outcomes**
Balancing	Society should adopt procedures that **balance** the **costs** and the **benefits** of the procedure. For example, an insurance system that provides rapid settlement of claims but does not seek to investigate the validity of those claims may pay out for invalid cases. However, it will save resources through not carrying out laborious investigations
Participation	Society should **ask those** who will be **affected** by a decision to choose which procedures to adopt. This model holds that those who will be affected by a decision should be involved in determining its outcome, whatever that outcome may be. This model is particularly relevant to health systems that enshrine public participation as a central part of their service development

4C.2 PRIORITIES AND RATIONING

Rationing in the context of health care can be defined as the process of **allocating finite resources**. The term is often used synonymously with prioritisation.

RATIONING

'*There are two certainties in life: death and scarcity.*' (Maynard 2001)

In economics, the term **scarcity** refers to finite resources, which may be monetary, but also include materials, equipment and human resources. Since in the field of health and health care there will always be finite resources but infinite demand, a system of decision-making is required that will decide which services should be provided and which should not.

Although explicit rationing systems (where transparent, consistent criteria are used to make entitlement decisions) are contentious, rationing decisions are still made every day and at every level of health care. These decisions range from:

- A GP receptionist's judgement of whether a patient should be offered an emergency appointment (based on an implicit decision of whether this is justified by the severity of symptoms), through to
- National guidance recommending or limiting the use of a new health-care technology (based on an evaluation of its cost-effectiveness).

NATIONAL, ORGANISATIONAL AND INDIVIDUAL RATIONING DECISIONS

See Table 4C.2.1.

Table 4C.2.1 Levels of rationing

Level	Agency	Advantages	Disadvantages
Macro	Government or state level (e.g. NICE guidance for health technologies in the UK)	Saves duplication of work at local level	No scope for local decision-making or autonomy
Meso	Health-care organisation (e.g. local exceptional treatments panel to decide whether interventions sought by patients should be funded by a health authority)	Can respond to local circumstances	Little scope to address individual circumstances Risk of unnecessary duplication if same activity happening in neighbouring areas
Micro	Individual clinician (e.g. consultation duration and frequency; intensity and scope of interventions provided)	Can respond to individual circumstances	Vulnerable to inconsistencies and lack of accountability in decision-making

STAKEHOLDER INVOLVEMENT IN RATIONING

Rationing decisions are greatly affected by the views and experiences of decision-makers: see Table 4C.2.2. Therefore, a wide range of stakeholders is to be encouraged. These may include:

- Voters
- Clinicians
- Patients
- Health-care management
- Pharmaceutical industry.

Table 4C.2.2 Ways in which the views and experiences of decision-makers can influence rationing

Perception of outcomes	For example, if large breasts were seen as a necessity rather than an aesthetic choice, then plastic surgery for breast augmentation would be more likely to be funded
Acquisition of opinions	Willingness-to-pay methods for valuing benefits will reach different conclusions according to: • An individual's **ability to pay** • An individual's **likelihood of requiring** the service • An individual's **likelihood of being required to pay** for the service
Position of groups in society	For example, attitudes towards older people may influence the degree of funding provided for social care and dementia treatments, etc.
Social norms	For example, given four hypothetical candidates for a heart transplant which would be chosen from the following? • A person who smokes • A person who drinks excessive alcohol • A person who is overweight • A person who is in prison

RATIONING CRITERIA

Rationing decisions in health care generally take into account the three criteria shown in Box 4C.2.1.

Box 4C.2.1

Necessity for an intervention	This can apply at two levels: • **Individual** (see Section 4C.1) (e.g. someone with severe atherosclerosis is more likely to have a myocardial infarction than another person with less severe atherosclerosis) • **Societal** (e.g. societies with a declining birth rates may be more likely to invest in IVF treatment for infertility than those with a rapidly growing population)
Cost-effectiveness	By definition, this includes only effective treatments. NICE uses cost per QALY as a measure of the incremental **cost-effectiveness** ratio – ICER (see Section 4D.8)
Fairness	Individuals in similar circumstances should have equal access to care, but the process for allocating resources should also be fair (i.e. it should consider concepts of **procedural justice** – see Section 4C.1)

However, the relative weight applied to each criterion will vary. As a result, even if the criteria for allocating resources are agreed, there can still be disagreement regarding the resultant decision. See Box 4.2.2.

USA Box 4C.2.2

Example: Setting health-care priorities in Oregon

In the late 1980s, the American state of Oregon chose to expand its publicly funded medical system – *Medicaid* – so as to increase the proportion of the population that was eligible for state-funded treatment. The ambition was for Medicaid to cover people on low incomes as well as those in absolute poverty. However, in order to remain within budget, the range of services funded by Medicaid would need to be reduced.

The rationing process

Oregon undertook an extensive research programme to assess the cost-effectiveness of a range of treatments. The outcome was a list of interventions ranked according to cost-effectiveness. The methods used initially produced some counterintuitive rankings, e.g. tooth capping was rated similar to appendicectomy.

The list went through several iterations to ensure that it was both politically acceptable and methodologically sound; then it was used to select which interventions Medicaid would fund.

'Priority lists' versus 'rule of rescue'

While Oregon's revised system provided access for more people on low incomes to certain interventions, those individuals in the lowest income bracket initially saw their access to health care narrowed. For example, Oregon's initial priority list resulted in denial of bone marrow transplantation to a child with leukaemia who subsequently died. The ensuing public outrage in Oregon typifies the inherent controversy of rationing processes. Application of the rules leads to cases where a named individual, who intuitively should receive treatment but does not fit the criteria, is denied a potentially life-saving intervention. This difficulty for society is termed the **rule of rescue**, i.e. society feels obliged to 'rescue' the named individual. The process of rationing denies services to some people and not others, and explicit rationing processes make this fact uncomfortably apparent.

Reproduced from Hadorn (1991).

4C.3 BALANCING EQUITY AND EFFICIENCY

It is generally accepted in public health that the allocation of health-care resources should be guided by the principles of **efficiency** and **equity**. However, the nature of these concepts depends on the exact definitions used. Furthermore, a system that distributes resources equitably requires some **sacrifice** of efficiency and vice versa. See Box 4C.3.1.

Box 4C.3.1

Concept	Definition
Equity	Fairness
Equality	Same for all
Efficiency	Greatest benefit achievable from a given resource

CONCEPTS OF EFFICIENCY

Definitions of efficiency typically relate to a utilitarian philosophical position, i.e. to achieve the greatest aggregated good across the greatest number in the whole community. For example, Donaldson et al (2004) defines efficiency in health care as securing the:

'… greatest improvements in wellbeing from available resources.'

Therefore, the concept of **efficiency** within health care depends on the definition of **wellbeing**. This is problematic because an improvement in health status may not necessarily improve wellbeing. For example, offering people a choice of how they wish to be treated may actually provide more wellbeing than would treating all individuals with the same treatment known to have the best evidence of success.

A state of ultimate efficiency is described by economists as being a **Pareto optimal** state (not to be confused with the Pareto principle [also known as the 80:20 rule] which states that, for many phenomena, 80% of the consequences result from 20% of the causes). In these circumstances, efficiency has reached the point where no further improvements can be made in one part of the system without disadvantaging others. In a health-care system that is not Pareto optimal, providing more resources for one part of a health system may not disadvantage others, and so this change is regarded as being an efficient choice.

CONCEPTS OF EQUITY

The two principal types of equity, **vertical** and **horizontal,** are described in Section 1C.10.

CONFLICT BETWEEN EFFICIENCY AND EQUITY

An intervention is efficient if it produces the **greatest net health gains** in a population for a given budget, whereas to be **equitable** it should be **distributed fairly** within the population.

Overall improvements in health status may mask widening differences in access to health care between constituent groups of the population. Even when access to services is similar, some groups within the population will require extra resources in order to achieve the same health gain.

ALIGNMENT BETWEEN EQUITY AND EFFICIENCY

Conflicts between equity and efficiency can sometimes be reconciled by adjusting the definition of wellbeing, e.g. by incorporating the concepts of **distributive justice** or **externalities**.

DISTRIBUTIVE JUSTICE

Distributive justice states that, for a system to be just, its benefits should be shared fairly. Justice then depends not only on the absolute **amount** of benefit received by individuals but also its **distribution** among them. A Pareto efficient system which is inequitable (i.e. the system benefits everyone but some people receive unfair amounts of benefit) is distributively unjust. By defining the wellbeing of individuals as being dependent both on their own health status and on distributive justice, the intervention will be seen as inefficient.

EXTERNALITIES

An externality is a by-product of the production or consumption of goods that are enjoyed by society in general. Equitable care provides positive (i.e. beneficial) externalities, and therefore improves the wellbeing of society. See box 4C.3.2.

Box 4C.3.2

Example: A hypothetical equity/efficiency alignment in the provision of universal immunisations

In order to achieve herd immunity to measles, 95% of the population need to be vaccinated. Uptake of vaccinations using a standard call–recall system with injections given at general practices is relatively cheap and easy to use, and achieves 80% coverage of the population.

For a variety of reasons, the remaining 20% do not attend for immunisations offered in the standard way. Additional resources will be required to encourage these groups to be vaccinated (e.g. mobile vaccination clinics, targeted health promotion, individual calls and visits). Although the unit cost of vaccinating these hard-to-reach people is much higher than the standard unit cost, it may actually be more efficient to spend this money because the 95% coverage achieved will offer herd immunity to the whole population.

4C.4 CONSUMERISM AND COMMUNITY PARTICIPATION

In health and health care, consumerism transforms patients from being **passive recipients** of care into **customers** who have the right and the capacity to choose whether and where to seek health or health care.

CONSUMER RIGHTS IN HEALTH CARE

In the UK health system, the concepts of the *'patient as a consumer'* and *'health care as a product'* were introduced in the 1980s. In 1991, the Department of Health produced *'The Patients' Charter'* which set out a list of rights that individuals could expect from health-care services. In common with many private companies, the NHS now uses satisfaction surveys as a key tool to assess the quality of its provision.

PATIENT AND CONSUMER CHOICE

Consumers of products have the choice of whether to consume services and, if so, where. In England, several policies to modernise the NHS have focused on providing more choice for patients. They are intended to:

- Introduce **competition** between health-care providers
- Improve the **responsiveness** of the health-care system to consumers' preferences
- Enhance the **efficiency** of health-care services.

None the less, choice within health care can be constrained: see Table 4C.4.1 and an example illustrates this in Box 4C.4.1.

Table 4C.4.1 Constraints in choice within health care

Health-care gatekeeping	Where patients are obliged to see a GP (or other **gatekeeper**) in order to access specialist health-care services
Urgency of treatment	In circumstances of acute health-care need (e.g. following a road traffic accident), individuals do not have the choice of *whether* to accept health care and may **lack the capacity** to decide where to seek services
Disempowerment	Individuals who are **unwell** in hospital may not feel as empowered to question decisions that they would have challenged when well or on familiar territory. In a small minority of circumstances (e.g. under a section of the Mental Health Act), this freedom to challenge decisions is suspended
Information asymmetry	Patients may be dependent upon clinicians to advise them about the **complexities** of the management of their conditions. While growing access to the internet has meant that some consumers have an abundance of information available to them, this potentially widens inequalities. One reason for this is that older people and people on low incomes are the least likely to have internet access

Eng Box 4C.4.1

Example: *Choosing health: making healthy choices easier*

The language and policy recommendations in the government's public health White Paper for England (DH, 2004) typify the high profile of choice in the country's health policy in the 2000s. The strategy explicitly recognises individuals as consumers: e.g. its chapter on health in a consumer society notes that, in a market economy, it is not for *'Government to dictate … what [people] can and cannot consume.'*

Recommendations focus on enabling individuals to choose health by means of:

- The **provision of information** to make an informed choice about whether to follow advice to promote or protect health. Therefore, many of the recommendations in the paper centre on the provision of effective, accurate and clear information. For example, there are recommendations for working with food manufacturers, on providing information on the fat, salt and sugar content of products
- **Availability of health products** for consumers to choose, e.g. through increasing the availability of foods with lower fat, salt and sugar content or in smaller portion sizes
- **Creating demand for health**, through the use of social marketing of health

Reproduced from Department of Health (2004).

CONSUMERISM AND COMMUNITY PARTICIPATION

Health care has traditionally had connotations of a passive role for patients, who were cared for by expert health professionals. However, other models of health-care provision involving self-help and community involvement have developed – both formally and informally. Formal examples include **expert patient programmes** and the provision of health care through not-for-profit organisations and social enterprises: see Sections 2I.4, 3 and 4B.1, and Box 4C.4.2.

Eng **Box 4C.4.2**

Example: NHS foundation hospitals – community participation in running health care

Foundation trusts have different regulatory and financial arrangements from other NHS trusts, and were first introduced in England in April 2004. It is of note that they were established as **public benefit corporations**, meaning that they have a **board of governors** (comprising patients, staff and members of the public) who have responsibilities regarding the appointment of the trust's chair, non-executive directors and chief executive, and the right to be consulted on the trust's strategic direction. Foundation trusts are also required to **recruit members**, i.e. people in their local community who want to be informed about the management of the trust. In this respect, they have the potential to bring true community participation to the hospital system.

In its early evaluation of foundation trusts, the King's Fund found that they had attracted a large membership (with over half a million people being members of the first wave of foundation trusts), but it was unclear to what extent these member were representative of local populations. With a few notable exceptions, the evaluation concluded that, *'so far there is little compelling evidence that members or governors have made a significant impact on the management of foundation trusts'*.

Reproduced from Lewis et al (2006) and Foundation Trust Network (2005).

4C.5 PUBLIC ACCESS TO INFORMATION

Providing the public with access to health-care information has the potential to:

- Improve individuals' **understanding** of health and disease, so as to inform their decisions about whether to seek care and how to prevent disease
- Make services more **accountable**, through the provision of information on patients' rights, the availability and quality of services. This enables patients to decide where to seek care if they are given a choice
- Keep patients **informed** about their own health care. In recent years, health services have made moves to improve access to information for patients about their own health. For example, in the NHS, patients now have a right to see information in their medical records

UNDERSTANDING OF HEALTH AND DISEASE

The **internet** has transformed access to health information for many people, and according to the National Consumer Council it has:

> *'The potential to result in a much more **active patient** and more **balanced relationship** with health professionals.'* (Sihota and Lennard 2004)

However, alongside this explosion of information, the following issues should be borne in mind:

- Variable **quality** of information available
- **Applicability** of information to the individual patient
- Health **literacy** regarding technical terminology
- Potential to worsen health **inequalities** because of varying access to the internet.

FREEDOM OF INFORMATION ACT

UK The Freedom of Information Act 2000 applies in England, Wales and Northern Ireland, and there is a similar act in Scotland. The Act applies to all *'public authorities'*, including:

- Central and local government
- NHS

- Schools, colleges and universities
- Police
- Many other non-departmental public bodies, committees and advisory bodies.

The Act requires public authorities to specify the kinds of information that they publish, how it is made available, and whether it is available free of charge or upon payment. In addition, the Act gives any person the legal right to ask for, and be given, any information held by a public authority – although certain exceptions apply. Public authorities must provide the information requested within **20 working days**. Information can be withheld to protect various interests which are allowed for by the Act, in which case an explanation that the information was withheld and the reasons why should be provided.

Those requesting information can be asked to pay a small amount for making photocopies or postage. If the public authority thinks that it will cost it more than a set amount (around £500 in 2007) to find the information and prepare it for release, then it can turn down the request.

When people ask for information that a public body holds about them, the request is handled under the Data Protection Act rather than under the Freedom of Information Act. There are slightly different rights under the two Acts, with different fees and timeframes being applicable.

NZ Under the Official Information Act 1982, any person can request government agencies (ministers, departments, local authorities) to release information which must be made available unless there is a good reason for withholding it.

4C.6 USER AND CARER INVOLVEMENT

User and carer involvement in service planning

Since the 1980s, governments throughout western Europe and North America have encouraged patients to contribute to the planning and development of health services. In England and Wales the involvement of patients is a key component of current strategies to improve the quality of health care. Underlying these changes is the belief that involving patients improves their health and quality of life. In particular, it leads to more **accessible** and **acceptable** health services.

POLICY INITIATIVES TO PROMOTE INVOLVEMENT

UK The UK has a long history of patient and public involvement in policy initiatives (Department of Health 1999) – in both health and social care. Key developments in this history include:

- Creation of **community health councils** that provided an early forum for public voices
- Development of a **needs-based approach** to the planning of health and social services
- Raising of expectations about patient rights and responsibilities through the **Patients' Charter**
- Recognition of the skills of patients and carers in **managing long-term conditions**
- Focus on the **patient experience** as the driver of NHS modernisation
- Pursuit of closer **partnerships** among health services, local authorities and the third sector
- Establishment of public involvement as a **statutory duty** of NHS organisations
- **Patient Advice and Liaison Services** (**PALS**).

4C.7 PROBLEMS OF POLICY IMPLEMENTATION

Policy implementation involves three principal activities: see Box 4C.7.1.

Box 4C.7.1

Activity	Description
Interpretation	The intent of policy-makers is discerned, and details are added
Organisation	A strategy for accomplishing policy goals is decided Administrative units are defined Methods of service delivery are chosen
Application	Policy is put into action

Typical barriers to implementation include:

- Overambitious **time scales**
- Lack of **skills** or training
- **Ill-defined roles** and responsibilities
- Poor **project management**
- Inadequate **contingency planning**.

The UK's National Audit Office lists the following six adverse consequences that can result from poor implementation of policy:

- Users' **expectations** not met
- Poor **quality** public services
- Adverse effects on economic **competitiveness**
- Adverse social or **environmental consequences**
- Little or no benefit delivered or **not sustainable** in the longer term
- Sections of society **excluded** from benefits.

4C.8 FORMULATING POLICY

Principal approaches to policy formation

Policy formulation is subject to a number of disparate influences: see Table 4C.8.1. The relative strengths of these influences will vary according to the nature of the policy and the organisation in which it is made.

Table 4C.8.1 Types of influences involved in policy formation

Type of influence	Explanation
Incremental	Policy is influenced by chance events, learning from mistakes, experimentation and a range of other influences
Pluralist	Many actors and interest groups influence the policy process. There may be a range of mechanisms by which these different voices are heard
Policy narratives	Different stories evolve to describe events. Some gain more authority and have more influence on policy decisions than others
Actor networks	Certain individuals or institutions spread narratives through chains of persuasion and influence
Political	Both personal politics and party politics influence policy decisions
Practitioners	Front-line staff have a strong influence on policy

Reproduced from Anderson and Sotir Hussey (2001).

MODELS OF POLICY FORMATION

Traditionally, two models of decision-making have been proposed: **rationalist** and **incrementalist**. A third model – that of **mixed scanning** – has recently been added. See Table 4C.8.2.

Table 4C.8.2 Models of decision-making in policy formation

Model	Proponent	Type	Description	Criticisms
Rationalism	Simon (1947)	Prescriptive	The aims of the policy are stated explicitly and are analysed in the light of existing and potential problems A list of available options is drawn up, and each option is evaluated in terms of costs and benefits, and the option that maximises benefits relative to costs is chosen	In reality, decisions do not begin with explicit goals, they do not involve consecutive stages, and decision-makers have restricted time and resources available to them. In practice, policy-making is actually a process of '*muddling through*'
Incrementalism	Lindblom (1959)	Descriptive	This model describes the process of '*muddling through*' in which: • Explicit goals are not set • Only marginal changes are made to existing policy • Decision-making is iterative • Policies adopted are those enjoying the widest agreement	This model is criticised as being too conservative in that it neglects the potential for innovation
Mixed scanning	Etzioni (1967)	Prescriptive	This model divides policy decisions into **fundamental** and **regular** decisions: • Fundamental decisions set new policy directions and are to be attained using a rational approach • Once the parameters of policy are set, an incremental approach is used for regular decisions	• Which decisions count as 'fundamental' and which as 'regular'? • How much time and resources should be invested into making fundamental decisions? • How long should regular incremental changes be continued? Many small changes may alter the fundamental nature of policy

4C.9 POWER, INTERESTS AND IDEOLOGY

Appreciation of concepts of power, interests and ideology

Policy changes are heavily influenced by power relationships within and between groups of people.

POWER

Power is the capacity to make something happen and often involves making others do things that they would not otherwise do. French and Raven (1960) identified several different types of power: see Box 4C.9.1.

Box 4C.9.1

Resource	Also known as **reward power,** the person with this power has control of resources (e.g. budgets, people) and has the power to reward people with promotion or funds
Position	By virtue of holding a particular job within an organisation, the **post holder** is entitled to the rights and privileges of that role
Coercive	The type of power that comes from the ability to **punish** (e.g. through the withdrawal of privileges or the imposition of penalties)
Personal	Also known as **charisma**. This power resides in the personality
Expert	This is power vested in someone because of their acknowledged **expertise**. A public health practitioner within an organisation has a degree of expert power
Negative	This power is the capacity to **stop** things happening or from even being discussed

Identifying which individuals have which types of power can be helpful for identifying which people need to be influenced in order to secure support for a new policy. It can also be used to identify and counter negative sources of power. Finally, it can help identify which sources of power an individual is under-using.

INTERESTS

See Section 5B.2.

IDEOLOGY

A political ideology describes a belief of how power should be **allocated** and to what **ends** it should be used. It can be a construct of political thought, and it often defines political parties and their policies.

Box 4C.9.2 attempts to group common ideologies into themes, but note that one ideology may belong to several groups and that related ideologies often overlap. For this reason, modern political parties often subscribe to a combination of ideologies. Note also that that the meanings of political labels differ between countries.

Box 4C.9.2

Type of ideology			
Class struggle	**Collectivity**	**Ethnicity**	**Religion**
Socialism	Socialism	Nationalism	Christian-based ideologies
Communism	Religious socialism	Fascism	Christian anarchism
Marxism	Christian socialism	Nazism	Hindu-based ideologies
Leninism	Democratic socialism	Neo-Nazism	Hindu nationalism
Stalinism	Communism	Racism, racialism	Islam-based ideologies
Neo-Marxism	Religious communism		Islamism, Muslim fundamentalism
	Marxism		Jewish-based ideologies
			Religious Zionism

4C.10 HEALTH INEQUALITIES

Inequalities in the distribution of health and health care and its access, including inequalities relating to social class, gender, culture and ethnicity, and their causes

See Section 4A.8.

Health inequalities are differences in health that are attributable, among other things, to a range of factors such as social class, age, gender, ethnicity and geography.

SOCIAL CLASS

Social class is a strong predictor of health outcomes.

UK BLACK REPORT (1980)

This key report was commissioned by the Labour government in the late 1970s, but was published after the Conservative party came to power in 1979. The incoming government attempted to suppress the publication of the report (although the endeavour backfired) and did not support its findings. Black and colleagues noted a **social class gradient** in both **morbidity** and **mortality** that was:

- Present at every age
- Present for all major diseases
- Increasing over time.

The authors listed four possible explanations for the gradient that they observed (see also Box 4A.8.1, p 385), namely:

- Statistical artefact – the way that social class is measured
- Social selection – the ill become poorer
- Behavioural factors – direct impact of nutrition, etc.
- Indirect impact of deprivation – effect on physical health is mediated by psychology.

The implication of the Black Report is that income fundamentally influences health, and that this influence lies **outside** the scope of the health-care system. A number of studies have been published since the Black Report that investigated the issue in more detail and provided potential solutions: see Table 4C.10.1.

Table 4C.10.1 Solutions to inequality proposed by key studies

Benzeval (1995)	Acheson Inquiry (1998)	Saving Lives: Our Healthier Nation (1999)	Wanless reports
Improve **physical environment** Address socioeconomic factors: **income** and employment Promote healthier **lifestyles** Improve **access** to services	**Evaluate government policies** regarding health inequalities Focus on policies affecting health of families and children Reduce inequalities in income and improve **housing** standards	Prominence given to health inequalities and four priority areas: • Cardiovascular disease • Accidents • Cancer • Mental health	2002 – made the **economic** case for public health 2004 – securing good health for the **whole population** 2006 – funding of **social care** for older people

AGE

Older people report more illness, and use health services more frequently than younger adults. This is due to a combination of material, biological and social factors:

- Avoidable illness and death linked to poverty (e.g. cold houses)
- Biological factors, which may be more common in certain ethnic groups (e.g. hypertension, type 2 diabetes)
- Cultural treatment of older people (e.g. disengagement/stigma) and social isolation.

GENDER

Demographic statistics and surveys show gender differences in health-related behaviour, illness and mentality. Risk-taking is greater in men – they are more likely than women to smoke and drink heavily, while women are more likely to eat fresh fruit and vegetables. Women are more likely to be diagnosed with anxiety or depression, although suicide rates are greater in men. Life-expectancy is greater for women than men at all ages. In childhood boys are more likely to die from accidents, poisoning and injury, and in adult life, men die earlier for a range of common causes such as cardiovascular disease.

UK ETHNICITY

There are differences in prevalence of illness and health seeking between ethnic groups:

- Cardiovascular disease, type 2 diabetes and hypertension are more prevalent in people of South asian and African–Caribbean ethnicities
- Mental illness is commoner in the African–Caribbean population
- Lowest levels of health service usage are seen in Chinese population.

GEOGRAPHICAL AREA

Geography may have an independent effect on health through the following mechanisms:

- Physical environment
- Home and work environments
- Local services
- Sociocultural factors
- Reputation (psychological factors).

4C.11 MIGRATION AND HEALTH EFFECTS OF INTERNATIONAL TRADE

While migration that is properly managed can bring huge financial and cultural benefits to both individuals and countries, mismanaged migration and human trafficking represent significant health risks.

The UN's **International Organization for Migration** estimates that 3% of the world's population are migrants (the highest proportion in world history). It expects migration to continue growing because of the factors listed in Box 4C.11.1.

Box 4C.11.1

Economic liberalisation	Increases in trade and globalisation lead to increased demands for skilled labour – especially in the IT, financial and hospitality sectors
Economic decline	Counter to what might be expected, temporary economic recessions tend not to lead to a downturn in migration
Demographic changes	Most developed countries have populations that are expected to shrink and to become older over the course of the coming decades. The young, growing populations of developing countries may serve to counter this '*population time bomb*'

Transnational migration	These are migrants who shuttle between multiple homes and maintain links with more than one country. Such people are said to live in *'transnational migration space'*. This growing phenomenon is facilitated by dual citizenship and multiple properties and voting rights
Conflict	There is no end in sight for wars and political upheaval

A distinction is made between **voluntary** and **forced** migrants: see Box 4C.11.2.

Box 4C.11.2

Reasons for voluntary migration	Reasons for forced migration
Employment	War
Study	Industrial, environmental or natural disasters
Rejoin family members	Famine
Retirement	Development projects, e.g. dams

CLASSIFICATION OF MIGRANTS

Migration is the permanent relocation of people between one country and another: see Box 4C.11.3.

Box 4C.11.3

Term	Alternative name	Definition	Rate
Emigration	Out-migration	Process of people leaving one country on a permanent or semi-permanent basis	Number of people departing from a country per 1000 of its population per annum
Immigration	In-migration	Process of people entering a country to take up residence	Number of people arriving per 1000 of the population of the receiving country per annum

An **asylum seeker** is a person who requests sanctuary in a destination country on grounds of having escaped persecution in the country of origin. A **convention refugee** is a person recognised by the destination country as having a *'well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion, is outside the country of his nationality, and is unable to, or owing to such fear, is unwilling to avail himself of the protection of that country'* and is therefore accorded the full rights of the 1951 UN Convention on Refugees. In contrast, a **quota refugee** is a person who is granted limited refugee status by the destination country before leaving the country of origin, usually as part of an agreement by which the destination country agrees to take a finite group of refugees over a short period of time.

Migrants may be classified as **documented** or **undocumented,** depending on whether the state authorities in the host or transit country have authorised residence and employment. Undocumented migrants (sometimes inappropriately referred to as *'illegal immigrants'*) are people who have either entered a host country without legal authorisation or overstayed their period of temporary, authorised entry.

UK A Home Office study (Robinson and Segrott 2002) in which asylum seekers were asked why they came to the UK found that the main reason was to seek a place of safety. Those who were in a position to choose a destination country, selected the UK for the following reasons:

- Relatives or friends already in the UK
- Their belief that it is a safe, tolerant and democratic country
- Established links between the country of origin and the UK (including colonialism)
- Ability to speak English or desire to learn the language.

MIGRATION AND HEALTH

Health implications can be divided into those on the country of origin, the migrants themselves and on the destination country: see Table 4C.11.1.

Table 4C.11.1 Migration and health

Country of origin	The WHO is particularly concerned about the drain of health-care workers from developing countries to developed countries because it can cause serious deficiencies in local provision of services. Most migrants are highly skilled, and the loss of new graduates places financial and human resource strains on the countries of origin. This is by no means compensated for by **remittances** (the term for the portion of migrants' wages that is sent back to the country of origin), so it contributes to the widening gap in income between rich and poor countries
Migrants	During conflicts, health-care workers may be displaced from the populations that they serve Poor sanitation and overcrowding in refugee camps are ideal conditions for **epidemics** to develop Sexual violence in conflict situations and refugee camps is a cause of **sexually transmitted infections** **Post-traumatic stress disorder** is highly prevalent among asylum seekers and refugees People being trafficked across international borders suffer **death in transit** (e.g. asphyxiation of a group of Chinese migrants in a cargo container while crossing the English Channel in 2000) On arrival in the destination country the health outcomes of migrants are affected by poor **living conditions,** a **lack of social integration,** stigma and open **xenophobic hostility**
Destination country	Screening of migrants at the point of entry is a politically sensitive issue In many countries, temporary migrants (most fall under this category for a time) are only entitled to free **emergency** health care Undocumented migrants avoid health-care workers for fear that they have links to the **immigration authorities** Moreover, migrants experience impaired **access** to health care because of unfamiliarity with local health-care systems and conventions, as well as linguistic or cultural difficulties in communicating their symptoms

Adapted from WHO International Migration, Health and Human Rights (2003b).

First-generation migrants often retain patterns of disease from their country of origin. For example, stomach cancer (which has a high prevalence in Japan and China due to diets that are rich in smoked, salted and pickled foods) is common among first-generation Japanese and Chinese immigrants to the USA. However, the incidence is much lower among descendants who adopt the lifestyle practices of the destination country.

INTERNATIONAL TRADE

The effects may be **direct** (e.g. an infectious disease is transported within traded goods or by means of an infected tradesperson) or **indirect** (e.g. trade of health technologies). Equally, health regulations may have effects on trade (e.g. transport of food) (WHO 2005). See Box 4C.11.4.

SA Box 4C.11.4

Example: spread of HIV/AIDS along trade routes

In Africa, international trade is implicated in the spread of HIV/AIDS. As a representative of the South African Medical Research Council put it:

'If you wanted to spread a sexually transmitted disease, you'd take thousands of men away from their families, isolate them in single sex hostels and give them easy access to alcohol and commercial sex. Then to spread the disease, you'd send them home every once in a while to their wives and girlfriends.' (Mark Lurie, South African Medical Research Council)

4.C12 INTERNATIONAL INFLUENCES ON HEALTH AND SOCIAL POLICY

When the United Nations was formed in 1945, member states agreed to work together to promote the *'economic and social advancement of all peoples'*. More than 60 years later, the health of the world population is under threat from environmental crises, and income and health inequalities continue to rise between the richest and poorest nations.

HEALTH PROMOTION

Cooperation at an international level to promote health is of particular value in the following respects:

- Exchange of ideas and mutual learning
- Pooling of resources
- Joint action.

SMOKING IN ENCLOSED PUBLIC PLACES

In recent years, one of the most significant steps taken by developed countries to improve public health has been to prohibit smoking in enclosed public places. The rapid spread of smoke-free policies is, however, attributed more to *zeitgeist* than concerted international policy. That said, the European Commission does promote cooperation between EU Member States on public health (particularly regarding health protection) and it commissions **Eurobarometer** surveys to investigate major public health issues such as physical activity, smoking and mental health.

SOCIAL POLICY

The field of international social policy compares the welfare provision found in different countries, ranging from the Nordic **welfare model** of comprehensive provision at one extreme, through to the scantier provision available in other countries.

UK In Britain, the Welfare State was established following the **Beveridge Report** in 1942, which identified five *'giant evils'* in society: squalor, ignorance, want, idleness and disease. By creating the welfare state (a series of policies designed to support people with financial, health or social needs) the government acknowledged its responsibility to care for the population *'from the cradle to the grave'*.

International social policy also addresses the potential impact of **globalisation** on the welfare state, and the influential roles played by **international actors** such as the International Monetary Fund, the United Nations, the World Bank and the World Trade Organization.

4C.13 INVESTMENT IN HEALTH IMPROVEMENT

Critical analysis of investment in health improvement, and the part played by economic development and global organisations

Although health promotion interventions have the potential to reduce future health-care costs by avoiding disease, the evidence for this is equivocal. Even where a health promotion programme is demonstrably successful (e.g. stop-smoking programmes), the overall cost to the state from a population with higher longevity may be more.

It is therefore important to remember that the objective of health promotion is not to save money but to reduce morbidity and mortality. Health promotion programmes should therefore be assessed along with other health-care or social care programmes, using the methods of economic evaluation and cost-effectiveness analysis (see Section 4D).

Hale (2000) identifies a number of potential reasons why the use of economic analysis is problematic in the context of health promotion interventions, including:

- Health promotion programmes have multiple objectives
- Clients are essentially healthy at the time and the benefits are broader than simply gains in health status
- QALYs do not capture the full range of benefits from health promotion
- Randomised controlled trials are needed to confirm that utility benefits were caused by the intervention, but these are difficult to conduct in the field of health promotion because of the timeframes involved
- The conclusion of the economic evaluations of health promotion interventions are highly dependent on the rates of discounting applied.

An example from the UK is given in Box 4C.13.1. The roles in investment played by global organisations are outlined in Table 4C.13.1.

Eng Box 4C.13.1

Example: NICE economic evaluation of physical activity interventions

In 2005, the Department of Health asked NICE to evaluate four health promotion interventions aimed at increasing physical activity:

- Primary care brief interventions
- Exercise referral
- Pedometers
- Community walking and cycling programmes

The review found an absence of evidence for the cost-effectiveness of pedometers and of community walking and cycling programmes. It found limited evidence that brief interventions in primary care were *less* cost-effective than usual care. For exercise referral, it found some cost-effectiveness evidence to suggest that the intervention is both more effective and more costly than usual care. The review was unable to determine the ICER (incremental cost-effectiveness ratio), however, because of the different outcome measures employed in the studies that it reviewed.

Other public health interventions being reviewed by NICE include:

- Smoking and tobacco control
- Obesity, diet and nutrition
- Alcohol
- Sexual health
- Mental health
- Drug misuse
- Promoting the health of children and young people
- Preventing accidental injury

Table 4C.13.1 Roles played by global organisations in investment in health improvement

Public health	Child health	Emergency aid
World Health Organization World Bank Food and Agriculture Organization	UNICEF (UN Children's Emergency Fund)	World Food Programme International Committee of the Red Cross/Crescent Medécins Sans Frontières

4D
Health Economics

Economics is based on the notion that we operate in a world of **scarcity**. In other words, there are infinite demands but only finite resources. This chapter considers how these finite resources can meet the infinite demands for health and health care through the practice of **resource allocation**. It provides a reference resource of techniques used in health economics, in particular **economic evaluation**, together with a consideration of their role in policy-making and public health.

4D.1 HEALTH ECONOMICS

Principles of health economics (including the notions of scarcity, supply and demand, marginal analysis, distinctions between need and demand, opportunity cost, margins, efficiency and equity)

It is an axiom in health care that **demand exceeds supply**. This is because there is an infinite amount that could potentially be spent on health care, so the actual resources available will always be relatively **scarce**. This means that **choices** must be made regarding how to spend the health-care budget. Health economics is the science of making these decisions, i.e. how best to employ scarce resources that have alternative uses.

The science is divided into **positive** and **normative** economics: see Box 4D.1.1.

Box 4D.1.1

Positive economics	Study of how **markets** work and how interventions will affect outcomes
Normative economics	Study of determining **what** should be produced, what resources to use and how to distribute goods

SPECIAL FEATURES OF HEALTH CARE

Economics has become a core discipline within public health. However, health care is seen as being special in economic terms for the reasons shown in Box 4D.1.2 (these will be explored later in this chapter).

Box 4D.1.2

Supply and demand	Demand and supply are not truly independent in health care
Imperfect markets	All health-care systems (but especially those that are publicly funded) are imperfect markets
Immediacy	Life and death decisions often need to be made with very short timescales
Agency	The nature of health care that people need, especially when they are critically ill, is largely specified by health-care providers
Uncertainty	Illness is often unpredictable
Necessity	Health care is an unavoidable commodity

SCARCITY

Health care can be regarded as a production process that uses a number of **inputs** in a **process** that produces **outputs**. The inputs (or 'factors of production') are divided into four categories: see Box 4D.1.3.

Box 4D.1.3

Land	Physical resources of the **planet** including mineral deposits
Capital	Resources **created by humans** to aid production, such as tools, machinery and factories
Labour	Human resources in the sense of people as **workers**
Enterprise	Human resource of **organising** the other three factors to produce goods and services

It can be seen that none of these resources is infinite. The term '*scarcity*' is used where more of a resource is wanted than is available. In these circumstances, every choice made involves a **sacrifice**, since the same resource cannot subsequently be used for something else. This sacrifice is called the **opportunity cost** (i.e. the benefits that are forgone are effectively the value of the benefits that we enjoy) and this is a fundamental concept of health economics.

CONCEPTS OF SUPPLY AND DEMAND

Demand is what people request; **supply** is what is provided.

Note that **demand is different from need** (need is something that a person will benefit from receiving) because people do not benefit from everything that they demand (e.g. antibiotics for a common cold).

Note also that **demand is different from what is received** because people do not receive treatments that they demand when they do not meet the eligibility criteria for that treatment (e.g. breast enlargement surgery under the NHS in England).

RISING DEMAND

In developed countries, demand for health care is rising for the following reasons

- Demographics (ageing population)
- Innovation (technology)
- Lifestyle (abuse)
- Information (educated consumer)
- Standards of living (quality-of-life expectations).

It is a principle of economics that – in a perfect market – supply and demand are **determined independently**, i.e. producers determine supply and purchasers determine demand. The price of goods rises or falls until the amount supplied equals the amount demanded, i.e. equilibrium is reached.

Another fundamental principle of economics is that **demand will equal supply in a perfect market**. While health and health care are not perfect markets, certain aspects of supply and demand do remain applicable.

SUPPLY

The supply of health care is the care that is made available; it is the capacity of services to meet need. Health-care supply may be quantified in terms of:

- Staffing (e.g. whole-time-equivalent consultants, nurses, etc.)
- Beds
- Equipment
- Budget.

DEMAND

Demand is the **expressed need**. Note that there is a distinction between the demand for *health* (i.e. to feel good and participate in all areas of life) and the demand for *health care* (i.e. a service required to achieve health), the latter being a derived demand.

Demand for health care may be quantified in terms of:

- Bed occupancy
- Consultation rates
- Waiting lists.

According to economic theory, demand is determined by four factors: price, income, preferences and alternatives (Table 4D.1.1).

Table 4D.1.1 Four factors that determine demand under economic theory

Price of health/health care	When fees are introduced to health-care systems (e.g. a charge to see a doctor or to obtain medicines), then demand drops (known as *price elasticity*). In England certain people (e.g. children) are exempt from prescription charges in order to ensure that their demand for health is not affected by the price of health care
Individuals' income	Studies have shown that the introduction of user fees reduces demand disproportionately in those on lower incomes (i.e. their demand is more *income elastic* compared with people on higher incomes)
Tastes and preferences	Different people place different values on different lifestyle factors. People also place different values on the benefits of health care: some people are more likely to seek health care for particular symptoms than others. For example, an individual may choose to trade off the unhealthy effects of a takeaway meal against the convenience of not having to cook
Price and availability of complements and substitutes	Substitutes for health care could include other types of health care (e.g. complementary medicine). Some individuals who use these services may choose not to use conventional health care. Others (e.g. those who place a particularly high priority on health) may use both types of health care

AGENCY

As will be seen in the section on markets below, one of the characteristics of a perfect market is that each consumer has perfect knowledge about the products on offer. There are several reasons why health care is not a perfect

market: one is that consumers do not have perfect knowledge about the complexities of health care available. Instead, they rely on **agents** such as doctors to inform them about what services they need. For example, patients with angina only receive angioplasty if their cardiologist thinks that they will benefit from it.

Perfect agents (like perfect markets) do not exist, because a perfect agent would have to strike a perfect balance between all of the following conflicting priorities:

- **Health status** of an individual patient
- **Preferences** of an individual patient
- **Utility** to society.

Moreover, agents may sometimes be motivated by other factors (e.g. self-interest).

SUPPLIER-INDUCED DEMAND

Supplier-induced demand occurs when agents act in their own interests and thereby recommend more health care than is necessary, i.e. more than would a perfect agent. For example, a dentist might recommend a dental filling that was not strictly necessary in order to gain the fee for performing the procedure.

The phenomenon is difficult to identify on a macroscopic scale because only price equilibrium points can be observed (i.e. if costs of dentistry increase and demand decreases, then it is difficult to say if it has decreased to the extent that would be expected if there were no supplier-induced demand. However, should the cost of a dental consultation increase when more practitioners entered the market, then the market would be demonstrably abnormal and supplier-induced demand would have been detected).

SUPPLIER-REDUCED DEMAND

In systems where the supply of health care is particularly scarce, **supplier-reduced demand** is seen. In this case an agent may not recommend a particular health-care intervention, whereas a perfect agent would have done so in the same circumstances. To an observer it would appear that there is less demand for a service than is in fact the case.

Rather than seeking to demonstrate and punish supplier-reduced demand, policy-makers can design contracts to eradicate it (e.g. by linking payment to quality indicators).

LAWS OF SUPPLY AND DEMAND

The **law of the downward sloping demand curve** states that demand is affected by the following:

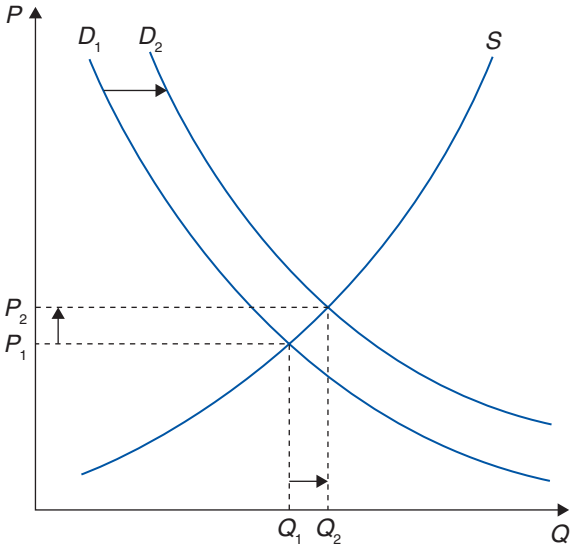
- Price
- Price of other products
- Income
- Consumer preferences.

SUPPLY AND DEMAND CURVES

See Figure 4D.1.1.

If a product sells at a low price, then producers will be disinclined to make large amounts of it. If the price rises, then producers will make more of the product and therefore the **supply curve** (S) **slopes upwards**. In contrast, if the price of a product is high, then consumers buy relatively little of it. If the price of the product drops, then consumers will be willing to buy more of the product, and thus the **demand curve** (D) **slopes downwards**.

Price reaches equilibrium at the intersection of the supply curve and the demand curve. If the price is set higher than this point (a move from price P_1 to price P_2), then producers will want to produce more of it (i.e. move from quantity Q_1 to Q_2). Since there is now more of the product available, customers will not be willing to pay as much – thereby returning the price back to the equilibrium point.



P = Price

Q = Quantity of goods

S = Supply

D = Demand

Figure 4D.1.1 Supply and demand curves

SHIFTS AND MOVEMENTS IN THE DEMAND CURVE

A change in price (*ceteris paribus* – all other things being equal) leads to a movement along the curve, whereas a change in other factors leads to a shift of the curve (i.e. the curve itself moves). See Box 4D.1.4.

Box 4D.1.4

Phenomenon	Causes	Change seen on the curve
Rise in price	More product being produced	Movement rightwards along the demand curve
Fall in price	Less product being produced	Movement leftwards along the demand curve
Purchasers not willing to pay as much	Falls in income	Leftward shift of the demand curve
Purchasers willing to pay more	Increased population Increased income Changes in taste (e.g. as a result of advertising)	Rightward shift of the demand curve
Producers find it easier to produce the product	More land More labour Technology available Price drops	Rightward shift of the supply curve

SUBSTITUTES AND COMPLEMENTS

Substitutes and complements are related goods where a change in price of one good affects the demand for the other: see Box 4D.1.5.

Box 4D.1.5

	Definition	Example
Substitutes	Products where an increase in price of one type of good causes an **increase** in demand for the other	Two different brands of the same vaccine
Complements	Products where an increase in price of one type of good causes a **decrease** in demand for the other	Needles and syringes

PRICE ELASTICITY

The price elasticity of demand is a measure of how sensitive is the demand for a particular good to changes in price.

PRICE ELASTICITY OF DEMAND

$$\text{Price elasticity of demand (PED)} = \frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in price}}$$

See Box 4D.1.6.

Box 4D.1.6

Absolute value* of PED	Definition	Characteristic
|PED| >1	Price elastic	Large response to price
|PED| = 1	Unit elasticity	Proportionate to price
|PED| <1	Price inelastic	Small response to price
|PED| = 0	Perfectly inelastic	No response to price

**The absolute value of PED is the magnitude with the + or – sign removed. It is symbolised using two vertical straight lines, e.g. |–0.5| = 0.5.*

PRICE ELASTICITY OF INCOME

$$\text{Income elasticity of demand (IED)} = \frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in income}}$$

See Box 4D.1.7.

Box 4D.1.7

Value of IED	Definition	Characteristic
IED >1	Luxury goods	Disproportionately large amounts demanded as incomes rise
IED >0	Normal goods	Amount of these goods demanded changes in line with income as would be expected from demand curve
IED <0	Inferior goods	Disproportionately large amounts demanded as incomes fall

MARKETS

A perfect market has the characteristics shown in Box 4D.1.8.

Box 4D.1.8

Atomicity	Many buyers and many sellers
Homogeneity	Identical products
Free entry	Sellers are free to join and leave the market
Equal access	Production technology is available equally to all sellers
Perfect information	All buyers and all sellers know the products and prices of all sellers
No externalities	An externality is a benefit (or disbenefit) to someone other than the purchaser, e.g. an externality of an immunisation programme is herd immunity

In a perfect market, the producers are **price-takers**, i.e. the market sets the price. In such circumstances, producers produce at the lowest possible cost in the long run, and they earn only **normal profits**. A normal profit is the same profit that could be achieved in the best alternative business (rather than an **economic profit**, which is revenue additional to this).

The **long run** is defined as the timeframe in which firms can enter or exit the market, and in which they can change their capital (e.g. build an extra operating theatre).

If producers do not operate in this way and, in particular, if they have a significant power to influence price or the total quantity being produced, then the market will fail.

CAUSES OF MARKET FAILURE IN HEALTH CARE

A market may fail if any of the conditions listed in Box 4D.1.8 are not met. The principal causes of market failure are listed in Box 4D.1.9.

Box 4D.1.9

Externality	This is a **side effect** of the product that is **not traded** on the market (e.g. herd immunity as a side effect of immunisation is an externality). Externalities may be beneficial (termed positive externalities) or harmful (negative externalities)
Public goods	These are *extreme* examples of an externality, such as a health promotion poster campaign, and are characterised by: • **'Non-rivalness'** (when one person reads the poster, other people do not suffer) • **Non-excludability** (it is impossible to stop a person from reading the poster)
Monopoly	**Monopoly** (single producer) **Monopsony** (single purchaser) **Oligopoly** (few producers)
Imperfect knowledge	**Uncertainty** (unable to predict demand, e.g. when trauma care needed) **Moral hazard** (no price to consumer at the time of use) **Adverse selection** (those at low risk of illness opt out of health insurance)
Merit goods	Belief that health services are in some way special

DISTINCTIONS BETWEEN NEED AND DEMAND

Economists define need as the **capacity to benefit** from health care (see Section 4C.1 for the differences between expressed, felt and normative need). In contrast, demand is a **request for health care**.

The key distinction, then, between need and demand is that the latter must be expressed by people by one of the following:

- **Attending** the place where the service is offered
- **Waiting** for the service
- **Paying** for the service.

Where need is not identified, it is not expressed, and it is therefore not a demand. Moreover, the health care that is demanded does not represent all health-care needs.

See Figure 4D.1.2.

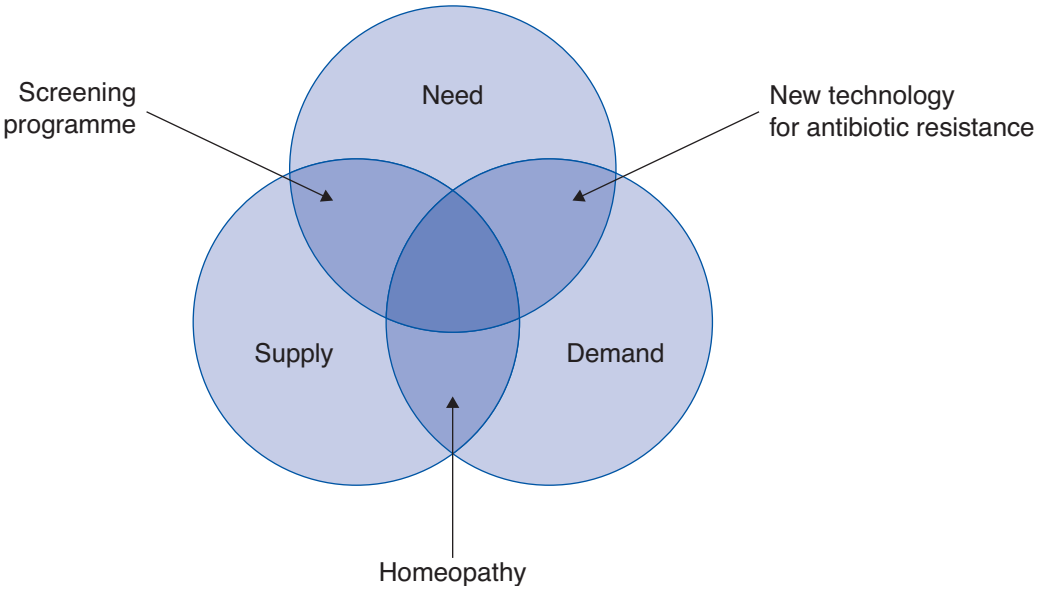


Figure 4D.1.2 Need, supply and demand

MARGINS

Choices often involve demanding a little more of one product or a little less of another. These are known as marginal changes. The margin is defined as the incremental variation in inputs that is required to have a corresponding variation on outputs. See Box 4D.1.10.

Box 4D.1.10

Marginal cost	Cost of producing one extra unit of service. This will reflect any **stepped costs** that are encountered (such as having to open an additional operating theatre because the capacity of the existing theatres has been exceeded)
Marginal benefit	Benefit derived from one extra produced unit

Marginal analysis examines the effect of small changes in the existing pattern of expenditure. It can identify:

- Where additional resources should be targeted

- Where reductions should be made if expenditure must be cut
- How resources can be reallocated to achieve an overall gain in benefit with no overall change in expenditure.

See Box 4D.1.11.

Box 4D.1.11

Example: Targeting a screening programme

A small-scale screening programme targeted at the highest risk groups may show a low cost per positive case detected. Continual expansion of the programme will entail screening progressively lower risk groups or screening more frequently. The number of screens required to detect each additional positive case will rise, increasing the cost per case detected.

Reproduced from Cohen (1994).

ECONOMIES AND DISECONOMIES OF SCALE

The average cost curve for the production of a good is U shaped: see Figure 4D.1.3.

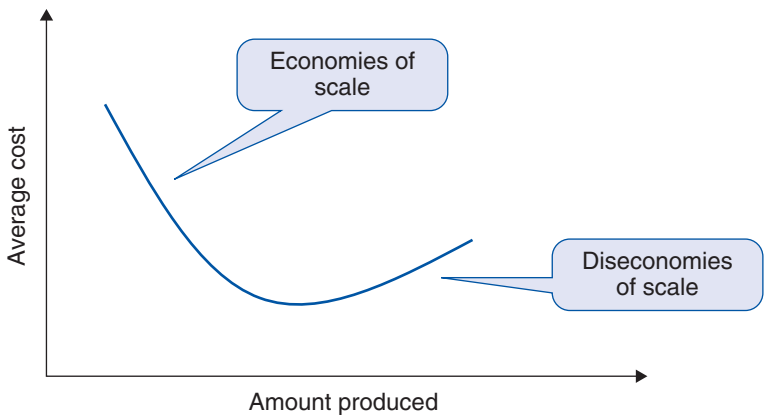


Figure 4D.1.3 Cost curve for the production of a good

As more of a type of goods is produced, so the average cost falls due to more efficient use of inputs. However, a point is reached where the diseconomies of scale begin to feature. These are caused by:

- Difficulties in managing an organisation that is so large that it is **unwieldy**
- Diseconomies of **scope** where it becomes increasingly costly to reach remote areas.

OPPORTUNITY COST

In health care, opportunity cost is quantified as the health benefits (life years saved, QALYs gained, etc) that could have been achieved had the money been spent on the **next best alternative** intervention or health-care programme. It can be calculated directly by means of cost-effectiveness or cost–utility studies (see Section 1C.14). Many studies attempt to compare particular interventions with existing practice, which itself may not be well defined. Failure to select an appropriate comparator can make the intervention appear more cost-effective than it should – leading to incorrect estimates of the opportunity cost.

EFFICIENCY

See Section 4C.3 for a comparison of equity, equality and efficiency.

Efficiency reflects how much health-care benefit is being achieved from the available resources. It can be considered in three ways: see Box 4D.1.12.

Box 4D.1.12

Technical efficiency	Maximum output for given inputs or minimal **inputs** needed for a given output
Economic efficiency	Maximum output for a given expenditure or minimal **costs** needed for a given output
Allocative efficiency	Efficient budget to produce the goods according to demand (i.e. produce what consumers value more than their cost) or set the level of production such that **marginal benefit > marginal cost**

TECHNICAL EFFICIENCY

Technical efficiency addresses the issue of using given resources to maximum advantage. An intervention is technically efficient if the same (or greater) outcome could not be produced with less of one type of input. For example, consider the treatment of osteoporosis using alendronate: if a 10 mg daily dose is as effective as a 20 mg dose, then the lower dose is more technically efficient.

ECONOMIC EFFICIENCY

Economic efficiency refers to the maximisation of health outcomes for a given cost, or the minimisation of costs for a given outcome, e.g. a policy of changing from maternal age screening to biochemical screening for Down's syndrome. If the sum of the costs of the new biochemical screening programme is smaller than or the same as the maternal age programme and outcomes are equal or better, then the biochemical programme is economically efficient in relation to the maternal age programme.

ALLOCATIVE EFFICIENCY

Allocative efficiency is the achievement of the best combination of health-care programmes to maximise the health of society. The concept of allocative efficiency takes account not only of the productive efficiency with which health-care resources are used to produce health outcomes, but also the efficiency with which these outcomes are **distributed** among the community. Allocative efficiency is achieved when resources are allocated so as to maximise the welfare of the community. Given that there will always be scarcity, a decision-making system is required that determines how much of which kinds of health care is provided. There are three possibilities: the free market, the command system and the mixed system (Box 4D.1.13).

Box 4D.1.13

Free market	Health-care resources are allocated according to consumers' purchasing behaviour
Command system	Planning is used to allocate health care according to some pre-determined criterion such as 'need'
Mixed system	Combines elements of the free market with elements of the command model

EQUITY

See Section 1C.10 for a comparison of vertical and horizontal equities and Section 4C.3 for a comparison of equity and efficiency. An example is given in Box 4D.1.14.

Box 4D.1.14

Example: Equity versus efficiency in cervical screening

The NHS policy on cervical cancer screening has been primarily aimed at maximising coverage by using economic incentives to GPs. However, there has been lower participation by high-risk women particularly those in disadvantaged socioeconomic groups. It has been calculated that the programme could have achieved the same cost-effectiveness in terms of cancers avoided with less extensive but more equitable coverage.

Reproduced from Sassi et al (2001).

DISCOUNTING

Discounting is a method used in economics to deal with the phenomenon of positive time preference, i.e. the human nature of preferring benefits to be realised now and for costs to be borne at a later date.

The reason for positive time preference is that the future is uncertain, so it is logical to want benefits earlier and costs later. However, many public health interventions (e.g. stop-smoking campaigns) are costly today but will not realise their benefits (e.g. reduction in lung cancer deaths) until many years into the future. In order to allow fair comparisons of costs and benefits to be made at different times, positive time preference can be compensated for by means of discounting. The strength of the time preference is reflected in the discount rate (which need not be the same for costs and benefits). Discount rates range widely between 0% and 6%, and the discount rate for benefits is particularly controversial.

Present cost = (Cost in year *n*) × (Discount factor)

where

$$\text{Discount factor} = \frac{1}{n(1 + r)}$$

and r = discount rate.

Because of the large effect of discounting on public health interventions, sensitivity analyses should always include a range of discount rates for both costs and benefits (see Section 4D.5).

4D.2 ASSESSING PERFORMANCE

See also Section 1C.9.

The WHO advises that performance assessment should encompass the aspects listed in Box 4D.2.1.

Box 4D.2.1

Social goals	Measuring the health system's contribution to socially desirable goals
Resource use	Measuring the health system and non-health system resources used to achieve these outcomes
Efficiency	Estimating the efficiency with which the resources are used to attain these outcomes
Review	Evaluating the way the functions of the system influence observed levels of attainment and efficiency
Feedback	Designing and implementing policies to improve attainment and efficiency and monitoring the effect

Economic analysis (see Section 4D.5) allows performance to be assessed in terms of the benefits derived per unit of expenditure. The performance of an allocation system is assessed in terms of efficiency (including Pareto efficiency) and equity (see Section 4C.3).

4D.3 FINANCIAL RESOURCE ALLOCATION

The objective of financial resource allocation is the transfer of funds appropriately from purchasers to providers in order to meet health-care objectives. The allocation process can be used to promote:

- Equity
- Changes in activity
- Efficiency (by means of incentive mechanisms).

Expenditure by the health service can broadly be classified as **recurrent** and **capital**: see Box 4D.3.1.

Box 4D.3.1

Recurrent	Staff, drugs, consumables
Capital	Buildings, equipment

NHS FUNDING

Eng The majority of funding comes from general taxation and National Insurance contributions, with the remainder coming from patient charges, e.g. prescription charges and dental charges. Recurrent revenue allocations to primary care trusts (PCTs) cover hospital and community health services, prescribing, primary medical services and HIV/AIDS. See Figure 4D.3.1.

Eng **Figure 4D.3.1** The allocation of funds (DH 2006)

ALLOCATION METHODS

Eng The Department of Health (DH 2005) makes allocations to PCTs according to four factors, including the national **weighted capitation formula**: see Box 4D.3.2. The DH is in the process of moving away from recurrent baselines towards the target of allocations based purely on weighted capitation.

NI Scot Wal In Northern Ireland, Scotland and Wales, the **Barnett formula** is used by the UK government to allocate central funding for services that are the responsibility of the devolved legislatures. This represents about 80% of public spending in these countries, and includes health care. The formula has been in use since the late 1970s although it has no statutory basis. The devolved assembly or parliament decides how to spend its financial block, including the relative proportions to be spent on health care and other services.

Scot The Scottish Executive health department allocates funds within Scotland on the basis of the Arbuthnott formula.

Wal The Welsh Assembly decided to move towards allocating resources to local health boards (LHBs) purely according to the **Welsh Townsend formula** (see Section 1C.8), rather than according to previous allocations. However, the pace of this transition is constrained by political considerations. LHBs are in turn expected to allocate their resources to reflect local need rather than demand.

Box 4D.3.2

Weighted capitation	PCTs receive their share of resources calculated according to: • Size and age distribution of the population • Additional need • Unavoidable geographical variations in the cost of providing services (called the **market forces factor**)
Recurrent baseline	This is the previous year's actual allocation, plus any adjustments made in-year
Distance from target	This is the difference between the weighted capitation and the recurrent baseline
Pace of change policy	This determines the level of extra resources that are allocated to PCTs that are below their weighted capitation target. The pace of change policy is decided by ministers for each allocation funding round

RATIONING

See Section 4C.2.

4D.4 HEALTH-CARE SYSTEMS AND INCENTIVES

Systems of health and social care and the role of incentives to achieve desired endpoints

Health-care systems can be considered in terms of a **revenue collector**, a **payer**, a **purchaser** and a **provider**.

Eng In England these can be thought of as follows (although the reality is more complex with each element of the financial system being affected by the others): see Box 4D.4.1.

Box 4D.4.1

Revenue collector	HM Treasury
Payer	Department of Health
Purchaser	Primary care trust
Providers	General practice, hospital trust

Social care covers a wide range of services that facilitate people to carry on in their daily lives, and particularly focuses on the groups shown in Box 4D.4.2.

Box 4D.4.2

Elderly people	Through residential care homes, nursing homes, home carers, meals on wheels, day centres, lunch clubs
Disabilities	People with physical disabilities or learning disabilities
Mental health	Ranging from support for those with mild mental illness, up to exercising legal powers for compulsory admission to psychiatric hospitals of potentially dangerous people
Ex-offenders	Those leaving prison may need help with resettlement, especially those with drug or alcohol problems
Families	Particularly where children have special needs such as a disability
Child protection	Including monitoring of children at risk
Children in care	Through fostering, accommodation in children's homes and adoption

Eng Around 70% of the Social Services budget for England is spent on adult community care services – in particular older people, people with physical or learning disabilities and mentally ill people. Some of these people pay for their own social care, with the remainder receiving financial assistance from the state (either through welfare benefits or through Social Services funding). Some people who require social care are given **direct payments** that provide them with the freedom to purchase their own care rather than having this arranged by the local authority.

INTEGRATED CARE

Integrated care is the provision of both health and social care services **in combination**, and it ensures that individuals receive the medical treatment and social care that they need (which frequently overlap). Integrated care requires health and social care organisations to work together to deliver flexible services that are tailored to allow people to live independent lives.

Examples of integrated care include cross-organisational services for **drug users** who have a range of other difficulties in their lives such as housing and education. Clients and staff decide together what medical and social care support is needed.

For frontline staff, the provision of integrated care requires:

- Working with individuals to identify their whole **range of needs**
- Knowing what other services are available
- Working alongside **other professional groups**
- Taking **responsibility** for arranging the right care or service.

MONITORING PERFORMANCE

Eng The **Commission for Social Care Inspection** (CSCI) provides an independent assessment of the quality and performance of the social care sector on behalf of the government and the public. It assesses how well councils are providing Social Services against objectives set by ministers from the Department of Health, and supports councils in improving their performance.

Eng Wal The performance of the NHS and independent health care is monitored by the **Healthcare Commission** (HC), which publishes the *Annual Healthcheck*: an annual performance assessment for each health-care organisation. The HC monitors the quality of the NHS and independent health care. It also evaluates the value of additional investment in the NHS and thereby judges how performance is improving. Other roles of the HC are to identify how the quality of health services may be improved, and whether appropriate arrangements are in place to promote the public health needs of the local population.

Eng NHS foundation hospitals are regulated by **Monitor** rather than by the HC.

FINANCIAL INCENTIVES

Financial factors may be used to encourage or discourage the provision of particular services, and may also encourage or inhibit the use of these services by patients. For example, there is evidence that prescription charges are negatively associated with the uptake of prescription medicines.

UK As part of the UK government's public service modernisation agenda, explicit incentives have been introduced with the aim of improving the efficiency of the health service: see Table 4D.4.1.

UK **Table 4D.4.1** Incentives to improve the efficiency of health care

Quality and Outcomes Framework (QOF)	In England, PCTs and general practices have increasing incentives to be cost-conscious in their decision-making. The *'Quality and Outcomes Framework'* is a voluntary system of **financial incentives** aimed at improving quality within the General Medical Services (GMS) contract for GPs
Delayed discharge	*Delivering the NHS Plan* requires local authorities to use some of their additional resources to reduce the number of people who remain in hospital after they have been deemed fit to be discharged. Failure to make appropriate alternative provision available to patients will result in **hospitals charging Social Services** departments for the costs incurred in keeping older people in hospital unnecessarily
Readmissions	Incentives also apply to NHS trusts, which are held accountable for the cost of emergency **hospital re-admissions** following a recent discharge. This is aimed at ensuring that patients are not transferred prematurely
Vaccinations	**Fee for service** is used to fund certain items of service provided to patients. In these cases, providers have an additional incentive to provide the services as it attracts more fees, e.g. fees received by GPs for administering vaccinations
Case mix	**Case-mix** payments are made according to measures of illness severity. For example, the UK health economy uses the Healthcare Resource Group (HRG) for units of charging (e.g. inguinal hernia repair without complication). HRGs are based on the average cost of a patient treated with that diagnosis; therefore, providers have an incentive to deliver care costing no more than the fixed payment of that HRG

4D.5 ECONOMIC APPRAISAL

Techniques of economic appraisal including cost-effectiveness analysis and modelling, cost–utility analysis, option appraisal and cost–benefit analysis, the measurement of health benefits in terms of QALYs and related measures

Economic evaluation is the **comparative** analysis of **costs** and **consequences** between two or more alternative interventions. For example, it may be used to compare the costs and benefits of switching from one vaccination programme to another. The aim of economic evaluation is to **improve efficiency** in the context of **scarce resources**, while also considering the impact on **equity**. Three types of efficiency are considered in health economics: technical, economic and allocative (see Section 4D.1).

COSTS

Evaluation of costs begins by defining:

1. The **perspective** (i.e. the viewpoint from which the costs and benefits are regarded). Commonly used perspectives are those of the patient, the hospital, the health authority or wider society
2. The **timeframe** for the evaluation.

The costs of the alternative interventions are then calculated. Depending on the perspective and timeframe chosen, different costs may or may not need to be included. For example:

- Costs to other government agencies (e.g. Social Services)
- Costs from loss of productivity
- Costs to the patient's family
- Out-of-pocket costs to the patient
- Future costs.

Costs may be classified as shown in Box 4D.5.1.

Box 4D.5.1

Direct	Salaries, drugs, diagnostic tests, patient transport, etc.
Indirect	Costs associated with reduced productivity due to illness, disability and death
Tangible	Costs that can readily be measured in currency terms and with certainty
Intangible	Psychological costs associated with illness or treatment, e.g. pain and suffering
Fixed	Costs that do not vary with the volume level of activity
Variable	Costs that vary in proportion to the quantity produced

Resource use is evaluated by using micro (bottom-up) or macro (top-down) costing, using either an RCT where the interventions represent different arms of the trial, or using some form of economic modelling (see below). The analysis should take account of opportunity costs, take a long-run economic perspective and be adjusted for discounting: see Box 4D.5.2.

Box 4D.5.2

Opportunity costs	It is the **opportunity cost** that should always be considered, i.e. the forgone value of the next-best use of the resources
Long-run costs	Costs should be assessed from the **long-run** perspective, i.e. where all inputs (including capital inputs such as buildings) can be altered freely. In the long run, the average cost is equal to the marginal cost
Discounted costs	Costs should be adjusted for human time preference by means of **discounting** (see Section 4D.1)
Marginal cost	Cost of producing one extra unit of service, reflecting any **stepped costs** that are encountered

Finally, a **sensitivity analysis** should be performed. This takes account of uncertainties in the economic analysis by establishing confidence intervals around the mean costs. One such uncertainty is the discounting rate that should be applied; therefore, the sensitivity analysis will include adjustments for a range of plausible discount rates. Note that because costs are often not normally distributed, **bootstrapping** (i.e. iterative) techniques may need to be used.

BENEFITS

Benefits may be expressed in a number of ways: see Box 4D.5.3.

Box 4D.5.3

Type of benefit	Definition	Example
Fixed benefit	Benefit is pre-determined such as a set number of operations	100 hip replacements
Clinical benefit	Benefit is a clinical endpoint	20 mmHg drop in blood pressure
Utility benefit	Benefit is expressed in terms of utility	Gain of 2.5 QALYs
Cost benefit	Benefits are translated into monetary values	Gain of £60 000

Approaches for the monetary valuation of life are shown in Box 4D.5.4.

Box 4D.5.4

	Approach	Advantages	Limitations
Human capital	The expected value of the individual's productivity is calculated (both market and household 'non-market' productivity), then adjusted for the individual's life-expectancy	• Easy to calculate • Easy to define	• Value of children, elderly and unemployed people appears less than that of males of working age • Does not reflect how much society is willing to pay for treatment
Contingent valuation	Surveys • *Willingness to pay* (WTP), i.e. how much a person is willing to pay for a health benefit or to avoid harmful risks • *Willingness to accept* (WTA), i.e. the minimum amount a person is willing to accept as compensation for a loss, or a reduction in a health-care service	• Does not rely on markets or observed behaviour • Can be applied to any good or service	• Requires large and costly surveys • Relies on hypothetical scenarios that may not reflect reality • Susceptible to bias: people may state different preferences from those that they actually hold • WTP and WTA can be affected by an individual's income
Hedonic wage/ revealed preference	Individuals' preferences regarding the value of health risk or benefit gain are traded against income (e.g. British soldiers' pay is greater than other comparable public sector posts because soldiers face risk during their service life)	• Based on actual consumer choices (indicates actual willingness to pay for items such as airbags, smoke alarms, etc.) • Gives insight into an individual's valuation of their own life	• Focused on immediate accidental deaths as opposed to deaths due to chronic exposures (e.g. asbestos) • Biased towards males of working age • Ignores imperfect labour markets (i.e. no knowledge of risks, may not have plenty of job choices) • Limited generalisability: people who undertake risky jobs are unlikely to be representative of the general population

ECONOMIC ANALYSIS

Having considered costs and benefits, the outcome of an economic analysis combines the two into a unified measure: see Box 4D.5.5 with further detail in Boxes 4D.5.6, 4D.5.7, 4D.5.8 and 4D.5.9.

Box 4D.5.5

Type of analysis	Type of benefit	Example	Advantages	Disadvantages
Cost-**minimisation**	Fixed outcome	£ per hip replacement	Simple to conduct	Restricted to one technology
Cost-**effectiveness**	Clinical outcomes	£ per 20 mmHg reduction in blood pressure	Straightforward outcome measures (life gained, blood pressure reduced)	One-dimensional Limited comparability
Cost-**utility**	QALY	£ per QALY	Allows comparison across health-care field	Practical problems in valuing utility
Cost-**benefit**	Currency (£/$, etc.)	£	Allows comparison across sectors (e.g. road building, schools)	Places monetary value on life which is considered priceless Practical problems in valuing health

COST-MINIMISATION ANALYSIS

This type of economic evaluation is used to find the least expensive method of achieving a single outcome: it assumes that interventions have an equivalent effect and compares programmes solely on the criterion of cost.

For example, treatments A and B both prevent 100 strokes per year. Under cost-minimisation analysis, the cheaper option would be chosen.

The main advantage and disadvantage of cost-minimisation analysis are shown in Box 4D.5.5. Additional advantages and disadvantages of cost-minimisation analysis are listed in Box 4D.5.6.

Box 4D.5.6

Advantage	Disadvantages
Simple (focuses on cost alone)	Assumes that equivalence of benefits has been proved unambiguously: much research effort would be needed to demonstrate this Few health-care interventions produce identical benefits

COST-EFFECTIVENESS ANALYSIS

This compares alternative treatments where both the costs and the benefits vary. Benefits are measured in **natural units** (e.g. years of life gained; fits of coughing prevented; mmHg drop in blood pressure; mmol/l drop in serum cholesterol). Since costs and benefits are measured in non-comparable units, the results are often expressed as **cost-effectiveness ratios**, i.e. the cost per unit of benefit for intervention (A) versus the cost per unit of benefit for intervention (B). Relative cost-efficiency can then be assessed, with the preferred option being that with the lower ratio.

If an intervention is both more expensive *and* more effective than an alternative, then the criterion for efficiency becomes the ratio of the **net increase in costs** to the **net increase in effectiveness**, i.e. the **incremental cost-effectiveness ratio** (**ICER**). The additional expense of the new intervention requires resources to be redirected from elsewhere. An economic evaluation assesses whether or not the additional benefits generated by the new intervention are greater than the loss of benefits from the reduction of other programmes (i.e. whether the re-allocation is efficient).

The main advantage and disadvantage of cost-effectiveness analysis are shown in Box 4D.5.5. Further advantages and disadvantages are listed in Box 4D.5.7.

Box 4D.5.7

Advantage	Disadvantage
Frequently incorporated within RCTs – hence cost-effectiveness analyses often provide the least biased estimates of effectiveness	Inability to compare interventions with differing natural effects. For example, interventions aimed at increasing life-years gained cannot be directly compared with those that improve physical functioning. Cost-effectiveness analysis therefore cannot directly address allocative efficiency

COST–UTILITY ANALYSIS

Cost–utility analysis is a form of economic evaluation that measures the effect of an intervention on both **morbidity and mortality**. By using a utility-based unit, such as QALYs, to measure benefits, cost–utility analysis is able to compare alternative health interventions that have completely different types of benefit.

Decision-makers can then be presented with **league tables** that **rank the incremental cost–utility ratios** of different interventions in order that they may select those interventions with the lowest ratios (i.e. best value) until the budget is expended.

The lower the incremental ratio for an intervention, the higher its priority should be in terms of maximising health benefits derived from a given level of expenditure. The point at which resources are exhausted defines a maximum price for a unit of effectiveness, e.g. £20 000 per QALY might be the upper limit of affordability within the budget.

Eliminating interventions with an incremental cost above this threshold price in favour of those below the threshold is a means of improving **allocative efficiency**. As with cost-effectiveness analysis, relative efficiency is assessed using an incremental ratio – here a cost utility ratio which takes as its units the **cost per QALY**:

$$\text{Incremental cost–utility ratio} = \frac{\text{Marginal cost (£)}}{\text{Marginal effect (QALY)}}$$

An intervention is deemed economically efficient, relative to an alternative, if it results in equal or higher benefits at lower cost.

For example, should £1m be spent on primary stroke prevention through antihypertensive treatment, or should £1.5m be spent on expanding a 'Children's Hospital at Home' service? A cost–utility analysis will deem that the money should be spent on whichever intervention produces more health gain (i.e. QALYs) per £ spent.

The main advantage and disadvantage of cost–utility analysis are in Box 4D.5. Further advantages and disadvantages are listed in Box 4D.5.8.

Box 4D.5.8

Advantage	Disadvantages
The use of a single measure of health benefit enables diverse health-care interventions to be compared, so cost–utility analysis can address both productive efficiency and allocative efficiency	Complex to conduct Debatable comparability of utilities arising from different measurement instruments, health problems and interventions

COST–BENEFIT ANALYSIS (CBA)

Finally, this type of economic evaluation involves measuring both costs and benefits in monetary terms. Benefits may be converted into monetary terms using willingness-to-pay exercises (revealed preference, contingent valuation). CBA compares two or more interventions from more than one sector (e.g. transport and education) according to their net effects on welfare.

For example, should £1m be spent on primary stroke prevention or on buying more books for local primary schools? CBA would quantify the monetary value of the health benefits and the educational benefits, and would recommend the programme in which the monetary benefit is the greater.

The main advantage and disadvantages of CBA are in Box 4D.5.5. Further advantages and disadvantages are listed in Box 4D.5.9.

Box 4D.5.9

Advantages	Disadvantage
Comprehensive: *all* costs and *all* benefits are considered in monetary units	Practical difficulty of assessing benefits in monetary terms
Increased use of interventions with the greatest net gain will increase efficiency	
By valuing all costs and benefits in the same units, cost–benefit analysis compares diverse interventions using the net benefit criterion	
Cost–benefit analysis thus simultaneously addresses issues of productive and allocative efficiency	

ECONOMIC MODELLING

In circumstances where a real experiment would be impractical or too time-consuming, economic models can be used instead to simulate the trial. Economic evaluation is particularly useful in the evaluation of interventions with benefits that will not be observed for many years (e.g. stop-smoking clinics).

While models are explicit and transparent, their validity remains questionable because they could easily be manipulated by changing one parameter slightly. This problem of parameter uncertainty can be addressed by conducting a sensitivity analysis in which parameters are varied across their range of plausible values to assess the effect on the ICER. Parameters can be altered one at a time (one-way sensitivity analysis) or simultaneously (multi-way sensitivity analysis and probabilistic sensitivity analysis).

See Figures 4D5.1–4D5.3.

OPTION APPRAISAL

This is the appraisal of various options chosen to achieve specific objectives. The relative advantages and disadvantages of the options are examined before resources are committed.

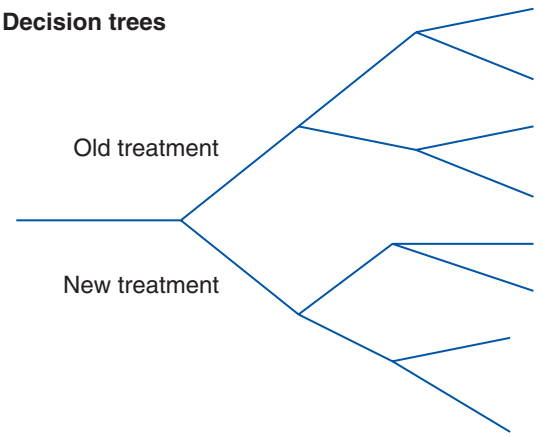
QUALITY-ADJUSTED LIFE YEARS

Economists use utility as an expression of an individual's preference for a particular health status or health outcome. A commonly used unit of utility is the QALY, which combines the **quality** and **duration** of life gained from an intervention as a single measure. For example, the UK's National Institute for Health and Clinical Excellence (NICE) collects evidence on the cost per QALY produced by the treatments that it appraises.

QALYs measure the years spent by an individual in different health states. Each year is weighted according to a scale between 0 (dead) and 1 (perfect health). Negatively weighted values may also be used for intractable symptoms.

Examples are shown in Box 4D.5.10, and advantages and disadvantages in Box 4D.5.11.

Decision trees



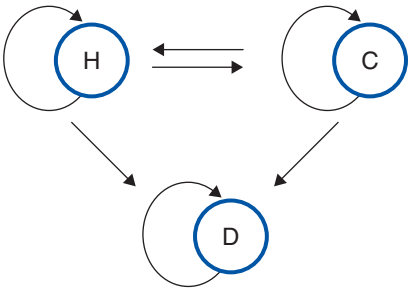
The first division of the tree represents the old and the new treatments. The next branch shows the potential outcomes for each treatment (either unproblematic or adverse events). Subsequent branches show potential consequences of those side effects (such as recovery or death)

A probability is assigned for each branching point (except for the initial division into old and new treatments)

Each final branch is assigned a cost (in £) and a benefit (in QALY). For the old treatment, the sum of the costs and benefits are added together, multiplied by the probability of each branch. The same is done for all the branches of the new treatment. The ICER then represents the incremental costs and benefits of moving from the old to the new treatment

Figure 4D.5.1 Decision trees

Markov modelling



Markov models are useful for analysing **chronic diseases**

Each circle represents a health state, e.g. healthy (H), chronic disease (C), dead (D). Using the best available evidence, each state is assigned a quality-of-life value and a cost

Straight arrows represent the probability, by the year-end, that a patient moves from one state to another. A new treatment will affect the probabilities associated with these arrows

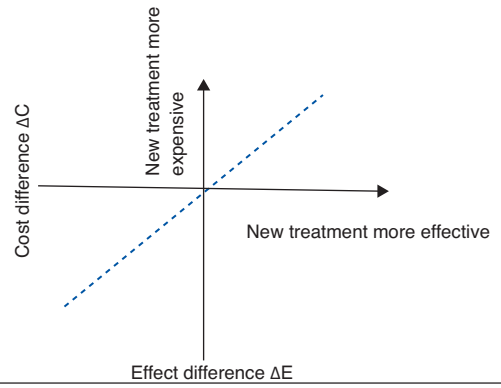
Circular arrows represent probability, at the year-end, that a patient remains in the same state

A hypothetical experiment is run where two cohorts of, say, 1000 patients each are followed over 25 years. One cohort receives the new treatment and the other receives standard care. Patients move between health states or stay in the same health state according to the probabilities defined above

Total costs (£) and total consequences (QALYs) over the 25 years are calculated for both cohorts. Adjustments are made to take account of discounting. An ICER for the new treatment can then be calculated

Figure 4D.5.2 Markov modelling

Reporting results



The dotted line represents the **cost-effectiveness plane**, i.e. the maximum acceptable ICER. For the UK this is generally accepted as being approximately £30 000 per QALY

Treatments below and to the right of the plane are likely to be accepted, those above and to the left rejected

Figure 4D.5.3 Reporting results: cost effectiveness plane

Box 4D.5.10

Examples: QALYs

5 years of perfect health = 5 QALYs

3 years of perfect health followed by 2 years in a state calculated to be 0.8 of perfect health = 4.6 QALYs

Box 4D.5.11

Advantages	Disadvantages
Single measure combining quality and quantity-of-life measures	Theoretically problematical
Allow comparisons of health-care outcomes across medical specialties	Difficulty of estimating consequences while healthy
Large studies using standard quality-of-life measuring instruments (e.g. SF-36 or EQ5D) lead to valid, reliable results	Moral questions about negatively weighted years (i.e life not worth living)
	Apparent discrimination in favour of paediatric treatments and against those for older people
	Do not capture externalities (e.g. benefits to the patient's family and friends); therefore, are likely to undervalue health care
	Calculation is dependent upon *who* and *how* asked
	No weighting for age (compare DALYs, below)
	No discounting

DERIVATION OF QALYs

QALY = Value of preference in a particular state x Length of time in that state.

Two key issues in determining value preferences are:

- **Who** is consulted
- **How** they are asked.

See Box 4D.5.12.

Methods for calculating QALYs are shown in Table 4D.5.1.

Box 4D.5.12

Who	There is evidence that the value of a QALY changes radically depending on who makes the value judgement. However, it remains a **moot point** whether the people consulted should be health professionals, the general public or patients who have experience of the particular medical condition or treatment
How	Five main approaches are used to derive quality-of-life weightings (see below). Note, however, that the **chronicity** of the hypothetical illness influences valuations, as does the **way** in which questions are asked. Since responses are given to imagined situations, they may not reflect real life accurately

Table 4D.5.1 Methods for calculating QALYs

Method	Description	Problem
Time trade-off	Respondents are asked how many years of life with the disease they are willing to give up for one year of full health	Influenced by time preference (compare discounting)
Standard gamble	Respondents imagine that they have the disease and are asked to gamble on taking a hypothetical cure which will either fully cure the disease or kill them. The treatment has a probability (p) of full cure and a probability ($1 - p$) of death, and p is varied between 0 and 1	Biased by respondents' attitude to risk
Rating scales	Subjects are asked to attribute values between 0 and 1 to a series of health states. With visual analogue scales, respondents place a mark on a 10 cm line that represents death at one end and full health at the other	Doubtful interval properties
Person trade-off	Subjects are asked to imagine groups of patients with a severe disease (X) and a mild disease (Y). Imagine there is a cure for both diseases and sufficient funds to pay for one patient with X. How many patients with disease Y would have to be cured for you to be indifferent? Subjects are asked whether they would cure one person in health state X or (n) people in health state Y. The value of (n) is varied until the subject is indifferent between the two alternatives	May be affected by factors other than the health state alone
Quality-of-life questionnaires	Quality-of-life questionnaires (e.g. EuroQol, Nottingham Health Profile) can be used to derive utility values for certain disease states	Scales are not appropriate for some disease states

DISABILITY-ADJUSTED LIFE YEARS (DALYs)

In contrast to the QALY, which measures the effect of disease on an **individual**, the DALY is a measure of the burden of disease in a **population**. It quantifies the impact on the population of **premature death** and **disability** by combining the two into a single measure.

DALY = Years of life lost to death (mortality) + Years of life lost to disability (injury and illness).

Disability is weighted between 1 and 0 (by using estimation methods similar to those for QALYs). DALYs are commonly used in low-income country settings.

DALYs are compared with QALYs in Box 4D.5.13.

Box 4D.5.13

Similarities with QALYs	Differences from QALYs
Combine morbidity and mortality into a single dimension Similar methods for calculating morbidity weightings (e.g. standard gamble)	Measure **disease burden** (rather than quality of life) Are **age weighted**, i.e. more weight is afforded to productive years (childhood is valued less than early adult life) Are **discounted** at 3%

SENSITIVITY ANALYSIS

A sensitivity analysis should be included in every economic evaluation. It involves varying the assumptions and estimates that underlie the study, e.g. by changing the discount rate to 2% from 6%, or by including a range of intangible costs. Since costs and other assumptions may not be normally distributed, iterative methods (e.g. bootstrapping) are often used.

Sensitivity analyses test the robustness of economic evaluations by:

- Making explicit any uncertainty, imprecision or methodological controversy within the evaluation
- Highlighting any areas where extra data are needed
- Allowing decision-makers to gauge how much confidence they should place in the study results.

4D.6 MARGINAL ANALYSIS

See Section 4D.1.

4D.7 DECISION ANALYSIS

Decision analysis is a process that involves:

- Division of a problem into simpler, manageable components
- Detailed examination of each component
- Formation of components into a logical sequence so as to identify the best solution.

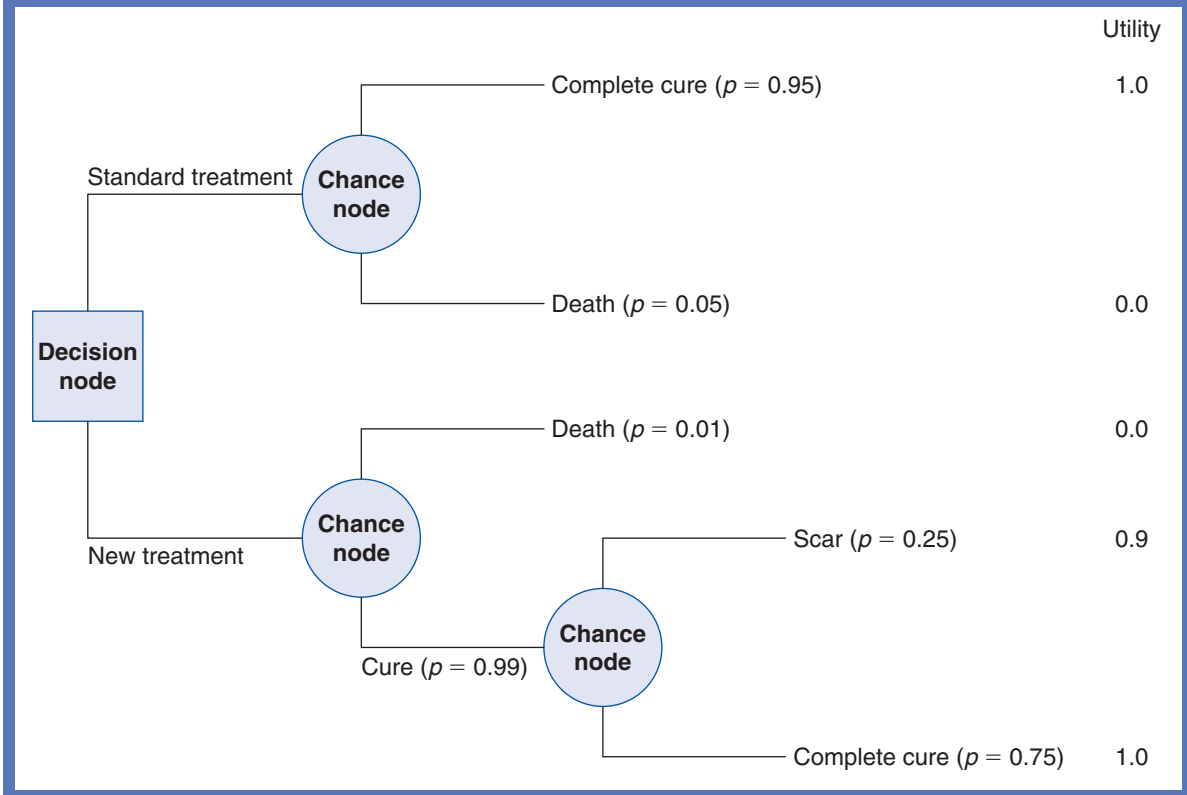
All options are identified and listed, together with the **probability** and **utility** of each possible outcome.

See Box 4D.7.1.

Box 4D.7.1

Example: decision analysis

For a particular disease there is a standard treatment (that results in either cure or death) and a new treatment (that may result in cure or death or disability). A decision tree is constructed to illiustrate these possibilities. Note that the branches emanating from each node are mutually exclusive, and that the probabilities at each branch must therefore add up to 1.



Reproduced from Jefferson et al (2000).

The probabilities of each outcome in a chance node are included, and utilities are assigned to every possible outcome using a common scale (a value between 0 and 1).

For each treatment, the sum of the utilities for each possible outcome multiplied by its probability is calculated.

Standard treatment = $(0.95 \times 1.0) + (0.05 \times 0.0) =$ **0.95**

New treatment = $(0.01 \times 0.0) + (0.99 \times 0.25 \times 0.9) + (0.99 \times 0.75 \times 1.0) =$ **0.966**

The option with the higher overall utility is chosen – in this case the new treatment. Finally, a sensitivity analysis is conducted (in which each probability and utility varied within a confidence interval) so as to test the robustness of conclusions.

4D.8 ECONOMIC EVALUATION AND PRIORITY SETTING

Role of economic evaluation and priority setting in health-care decision-making, including the cost-effectiveness of public health, and public health interventions and involvement

Wanless's first report (2002) argued that there would be financial benefits to the UK from investing more resources in public health. His second report (2004) called for a single framework for evaluating the cost-effectiveness of both **health care** and **public health** – the implication being that investment in the latter would represent markedly better value for money.

Both cost–utility analyses and cost–benefit analyses allow interventions with disparate outcomes to be compared on the same scale. These include public health interventions that tackle the risk factors for multiple diseases (e.g. smoking and obesity) and the wider determinants of health. Since the benefits of public health interventions (health promotion, screening and treatment) may not be realised for decades, so the principle of **discounting** becomes particularly pertinent (see Section 4D.1).

ECONOMIC EVALUATION

UK In the UK, NICE provides priority-setting advice to governments based on economic evaluations of the categories shown in Box 4D.8.1.

Box 4D.8.1

Health technologies	Drugs, procedures and other treatments
Clinical practice	Management of clinical conditions and diseases
Public health interventions	Health promotion and preventive medicine

For each of these categories, the process of economic evaluation consists of **identifying,** then **critically appraising** and **synthesising** the evidence, so that recommendations of cost-effectiveness (expressed as cost per QALY) can be made: see Box 4D.8.2.

Box 4D.8.2

Identification of evidence	A search of the published and unpublished evidence is performed (using economic search filters). Both qualitative and quantitative studies are sought from the literature and the grey literature
Critical appraisal	The quality of individual studies is appraised, following the hierarchy of evidence (see Section 1C.5), where RCTs (or cluster RCTs for community interventions) are particularly valued
Synthesis of evidence	Assessment of applicability, construction of evidence tables, meta-analysis and summaries of the evidence (in the form of evidence statements)

PUBLIC HEALTH INTERVENTIONS

Public health activities may be **direct** or **indirect**: see Box 4D.8.3.

Box 4D.8.3

Direct	Stop-smoking services Obesity clinics Health promotion posters
Indirect	Creation of parks and open spaces Restrictions on advertising fast foods

When evaluating a public health intervention, NICE follows the process shown in Box 4D.8.4. NICE uses a discount rate of 3.5% for costs and benefits in its analyses, and adopts the following economic perspectives: NHS, society and the individual.

Box 4D.8.4

Background work	Topics for evaluation are set by ministers Stakeholders are identified who will scrutinise and validate the evaluation process Quality assurance arrangements are put in place
Preliminary work	The scope of the evaluation is determined (i.e. definition, settings, population, exclusion criteria) A preliminary literature search is conducted
Detailed work	Key questions are set and an analytical framework is constructed A review of the evidence is conducted (selection of studies, quality assessment, evidence of implementation) The evidence is synthesised in the form of **evidence statements**

For each question addressed by the evaluation, an evidence statement is presented that documents both the **strength** of the evidence (i.e. its quality, quantity and appropriateness) and its **applicability** to the target population. NICE rates the strength of evidence according to the scale shown in Section 1C.5.

For example, a hypothetical evidence statement might be:

> *'A body of 2+ evidence of efficacy offers consistent findings about the impact of pogo sticks on weight loss. The evidence is directly applicable to the target population in terms of ethnicity, age and gender.'*

COST–CONSEQUENCE ANALYSIS

Because of the complexities of public health interventions, NICE sometimes uses a **cost–consequence** approach in addition to cost-effectiveness analysis. This technique enables non-quantifiable outcomes, equity and distributive justice to be considered in an explicit way (see Section 4C). It presents a table of the lifetime impact of the new treatment on individuals or groups of individuals in terms of:

- Resource use (both health-care and productivity losses)
- Health outcomes (symptoms, life-expectancy and quality of life).

Section 5

ORGANISATION AND MANAGEMENT OF HEALTH CARE AND HEALTH-CARE PROGRAMMES

Management is not limited to managers. Every member of an organisation has some managerial role. At its simplest level, this may require controlling one's own time and resources; at its most complex, it could demand directing projects, people and budgets.

Section 5 covers the major theories of management, from working with people to managing budgets. The challenges and benefits of working in teams are covered, together with an overview of how to motivate individuals. To enable team working to function optimally, an understanding of personalities, the ways that people think and how they function when they work together is needed. At a higher level, organisational structure and culture can account for the success or failure of a health system in meeting its goals. Understanding organisations, how they function (and more importantly, why they fail), is fundamental to implementing change in health systems.

Change is an essential part of any health service provision. As new technologies, diseases and different political priorities emerge, so health care must evolve and transform to meet the needs of its population. Robust strategies are needed to manage change carefully.

Health-care systems exist in a world with finite resources. Practitioners will be all too aware of resource constraints and their effects on the capacity of organisations to provide health care. Understanding the basics of finance and management accounting provides them with insights that are essential for establishing and sustaining health services.

5A
Individuals, Teams and Groups

Understanding individuals, teams/groups and their development

When things are going well, the benefits of working with others are clear: stimulation of ideas, enthusiasm, enjoyment, interest, increased productivity and a greater sense of reward. However, there are times when working with other people leads to conflict and stifled creativity, and impedes the progress of a project.

Public health practitioners often find themselves in the position of working with people from different professional groups, sometimes across organisational boundaries. A knowledge of management principles, team dynamics and interactions between professional groups is vital to understanding how groups function. It can also offer ways to analyse and address conflict, and enables us to modify our own attitudes and practices accordingly.

5A.1 INDIVIDUALS, GROUPS AND TEAM DYNAMICS

Motivation, creativity and innovation in individuals, and its relationship to group and team dynamics

People work differently when they are working alone or in groups. At its best, group work results in outcomes of a quality that the individuals working alone could never have produced. However, it is important that the **composition**, **dynamics** and **leadership** of the group be considered so as to motivate its members towards working together as productively as possible.

Groups may be convened **formally** (in order to perform a specific task) or **informally** (by individuals on the basis of a common interest). They are collections of people who:

- Interact with one another
- Are psychologically aware of one another
- Perceive themselves to be in a group (Schein and Bennis 1965).

GROUP MOTIVATION

Group motivation depends primarily upon individual motivation. Within a group, individuals will adopt one or more roles, as determined by a compromise between:

- How an individual wants to behave
- How other group members expect them to behave
- The group task.

Belbin (1996) described eight roles that should be represented by the members of an effective team, and can be remembered with the mnemonic ICE FIRST: see Table 5A.1.1. A ninth role of *specialist* is sometimes added, i.e. a person with a high level of skill in one given discipline. Note that team members may fulfil more than one role.

Table 5A.1.1 Belbin's eight roles for an effective team

Implementer/company worker	Make things happen High degree of self-discipline Deliver on time
Coordinator/chair	Default chairperson Step back to see the big picture
Evaluator/monitor	Fair and even-handed observers and judges of what is going on Can become almost machine like May have difficulty inspiring themselves or others to be passionate about their work
Finisher/completer	Perfectionists Strong sense of duty Complete painstaking and unpleasant tasks if they believe that this will improve quality May frustrate their team mates
Innovator/plant	Come up with unusual and innovative solutions to problems
Resource investigator	Vigorously pursue contacts and opportunities Excellent networkers Tend to lose momentum towards the end of a project
Shaper	Eager individuals that provoke their team into action May be insensitive to the feelings and perceptions of others
Team worker	Ensure that everyone in the working group is getting along Good listener and diplomat Talented at smoothing over conflicts

Groups can be motivated by means of:

- **Feedback** to remind the team members of the importance of their group's task, together with their individual roles within it
- **Leadership**: a good leader develops team spirit and elicits a high level of commitment from team members.

INNOVATION

Amabile (1998) has argued that three components are needed for creativity in companies:

1. **Motivation** (especially intrinsic motivation, i.e. satisfying own needs)
2. **Expertise** (technical, procedural and intellectual knowledge)
3. **Flexible thinking** (how flexibly and imaginatively people approach problems).

Creativity itself can be thought of as a three-step sequence consisting of an **input**, a **process** and an **output** (i.e. innovation): see Figure 5A.1.1.

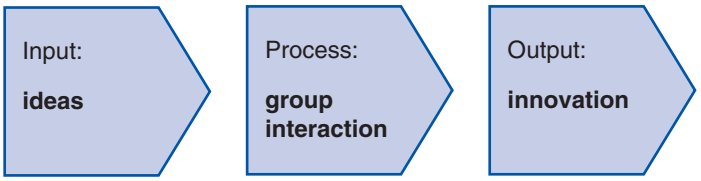


Figure 5A.1.1 Sequence for creativity

TEAM DYNAMICS

Tuckman (1965) described a four-stage process through which newly formed teams progress and develop: see Box 5A.1.1. A fifth stage, **adjourning**, was added in 1975 and refers to the disbanding of the team after its task has been completed.

Box 5A.1.1

Forming	Tasks and rules are established Resources are acquired Reliance is placed on the leader
Storming	Internal conflict Members resist the task emotionally
Norming	Conflict is settled Cooperation develops Views are exchanged Norms (i.e. new standards) are developed
Performing	Teamwork is achieved Flexible roles are developed Solutions are found and implemented

TEAMWORK

Charles Handy (1978) maintains that the ideal team not only achieves **formal** goals (i.e. those of its organisation), but also achieves **informal goals** [such as Pilowski and Spence (1997) satisfying its members' psychosocial needs]. The ideal team is described as having the traits listed in Box 5A.1.2.

Box 5A.1.2

Organisation	Common purpose Clearly defined task Clear team objective
Members	Members have specific expertise Members know their roles
Teamwork	Members support each other Members complement each other in skills and personalities Members are committed to accomplishing the task
Leadership	Leader who coordinates and takes responsibility

5A.2 CREATIVITY AND INNOVATION

Barriers to, and stimulation of, creativity and innovation (e.g. by brainstorming)

Creativity is one of the key strengths that public health professionals can bring to their working environment. By focusing on the 'bigger picture', the public health viewpoint is well placed to challenge established practices in a positive way.

STIMULATION OF CREATIVITY

The following techniques have all been used by successful organisations to stimulate creativity and innovation among staff:

- Brainstorming
- Suggestion boxes
- Team away-days
- Mind-mapping
- Reward.

Brainstorming is a method for promoting creativity in which team members '*think out loud*' in order to solve a problem. The process is conducted in an atmosphere that is conducive to independent thought and discussion. It is held to be a politically incorrect term by some, who prefer to use the term 'thought shower'. The Plain English Campaign considers that this suggestion 'reaches the point of real ridicule'. National organisations representing people with epilepsy or mental illness do not consider the term 'brainstorm' to be politically incorrect (see www.greencomms.com/index.jsp?i=100&s=1111).

DIFFUSION OF INNOVATIONS

New ideas are likely to spread rapidly if they are:

- Compatible with existing systems
- Simple
- Amenable to being trialled
- Demonstrably more efficacious.

An example of an innovation that was rapidly propagated is that of crack cocaine: crack met all four of the conditions above and its use spread rapidly among existing cocaine users.

BARRIERS TO CREATIVITY

The following factors hamper innovation:

- Uncertainty
- Over familiarity
- Fear of change
- Team conflict.

5A.3 INTERPROFESSIONAL LEARNING

Learning with individuals from different professional backgrounds

Several developments in health systems have led to an increased focus on interprofessional learning in recent years:

- Changing professional boundaries and requirements for clinical practice
- Changing health-care structures
- Emphasis within health care on multi-professional teams and clinical networks.

Interprofessional learning is more than simply sharing classes or training sessions with other professionals. According to Humphris and Hean (2004), this requires sessions that enable students to learn '*with, from and about one another*'. They also note that this requires more emphasis on small-group learning rather than large, didactic teaching sessions.

Eng In England, interprofessional education has gained increasing momentum since early commitments in the *NHS Plan* (DH 2000). The Bristol Royal Infirmary Inquiry (a high-profile enquiry into children's heart surgery at Bristol Royal Infirmary from 2001) made a number of recommendations to promote interprofessional learning from as early as possible in clinical training to reduce '***damaging intertribal rivalries***' in the NHS. The Department of Health's (2002) strategy to promote interprofessional learning, *Working Together – Learning Together* (see www.dh.gov.uk) includes a commitment to include interprofessional learning in:

- Education and training for all health professionals
- Undergraduate and pre-registration programmes, through to continuing professional education
- Practice placements as well as in the classroom.

A number of initiatives have been supported and funded, both nationally and locally, to support interprofessional and common learning in health professional education.

See Box 5A.3.1.

UK Box 5A.3.1

Example: Expansion of the public health workforce beyond medicine

Many disciplines are involved in delivering public health. In the past, a formal route for doctors to train in public health existed but the expertise of other professional groups was not always recognised or regulated.

Public health in the UK has changed in recent years to enable individuals from a range of non-medical backgrounds to enter formal specialist training programmes. Both medics and non-medics follow a **common recruitment process and training programme** to become public health consultants.

Once they have completed training, doctors remain registered with the General Medical Council. A new organisation, the UK Public Health Register, has been developed to maintain a register of non-medical public health professionals who have completed specialist training or have been assessed through a portfolio of work as competent to work as a public health consultant.

Reproduced from the UK Voluntary Register for Public Health Specialists: www.publichealthregister.org.uk/index.asp; Faculty of Public Health: www.fph.org.uk.

There are advantages and disadvantages involved in moving away from education in single professional groups towards interprofessional learning: see Box 5A.3.2.

Box 5A.3.2

Advantages	Disadvantages
Improve **communication** between different professions	Affect **professional support** by reducing traditional professional networks and support systems
Reduce the formation of **professional *'silos'***	**Resource intensive** to reconfigure training and education programmes to provide interprofessional learning
Promote clinical debate through opening issues to **different professional perspectives**	**Lack of evidence** that interprofessional learning produces the outcomes envisaged
Improve **teamwork** through enhanced appreciation of the roles of other staff	**Long lead time for evaluation** of pre-registration professional learning impact (10–12 years before students are practising consultants)
Improve **patient care** by developing care centred on the patient rather than on professional structures	
Enhance **capacity** by expanding clinical roles as required by the situation and needs of the team	

5A.4 PERSONAL MANAGEMENT SKILLS

Personal management skills (e.g. managing time, stress, difficult people, meetings)

Effective managers need skills in managing people and projects, but also in managing their own resources. The skills needed to manage other people are the same skills required to manage ourselves: the ability to plan, delegate, organise, direct and control.

TIME MANAGEMENT

Time management is a key component of self-management. Time can be managed by:

- Prioritisation into long-, middle- and short-term (today's) plans
- Organising offices, workstations and desks
- Managing paperwork and emails systematically
- Delegation.

Covey (1989) identifies a set of behaviours associated with efficient time management:

- Being proactive
- Beginning with the end
- Putting first things first
- Thinking win–win
- Seeking first to understand, then to be understood
- Synergising
- Sharpening the saw (taking time off).

STRESS MANAGEMENT

Several factors promote stress, including increased competition, de-regulation and rapid change. Methods of dealing with stress include:

- Recognising the **symptoms** of stress in yourself and others
- Understanding the **factors** that cause stress
- Applying a range of **strategies** to avoid, reduce and manage stress.

Well-run meetings reduce stress, save time and improve the effectiveness of the organisation. They tend to have the following features:

- Carefully **planned** with a realistic and succinct agenda
- Well **chaired**
- High **participation** before the meeting
- Participants are **aware of the contribution** that the meeting will make to the managerial process of the organisation.

External facilitators can improve the effectiveness of meetings by:

- **Taking the chair** (an independent facilitator focuses participants' thoughts on the matters at hand)
- Providing **training** for the chairperson and meeting participants
- Helping committees or boards to **establish protocols** for working effectively together
- Helping meetings decide the **level of issues** for which decisions can be made within the meeting, and therefore what topics should be discussed elsewhere.

MANAGING CONFLICT

The three principal methods for avoiding conflict are:

1. Negotiation
2. Mediation
3. Group arbitration.

Difficult situations can often be avoided by the careful use of **communication skills** to improve the delivery of bad or unpleasant news. By recognising **personality types** (especially your own personality characteristics), it is possible to identify the people with whom you are likely to clash and to develop strategies for dealing with them.

When conflict does occur, it may be prevented from escalating further by means of the following conflict–resolution principles:

- Mutual respect
- Identification of shared values
- Honesty
- Shared objectives
- Combating disinformation.

5A.5 THE EFFECTIVE MANAGER

An effective manager can help improve the quality and productivity of an organisation by means of delegation, feedback and listening.

DELEGATION

See Section 5A.6.

FEEDBACK

Feedback is crucial for the regulation of any system. It is the process whereby elements of the output of the system are returned to its input so as to regulate further output. Likewise, the feedback that a manager provides to an employee influences the performance of that employee.

The way in which feedback is delivered is pivotal, and managers need to practise giving feedback (an activity that many will actively avoid). A good manager will provide feedback only on specific, observable behaviours and will ensure that the feedback is provided non-confrontationally – perhaps as part of Iles's (2005) **criticism sandwich** (good news, bad news, more good news).

LISTENING

Iles (2005) argues that the success of organisations can be assessed by the speed with which bad news travels upwards through the management hierarchy. One way to ensure that the senior management team are kept informed of developments is to use the technique of **management by wandering around** (MBWA). Developed at the Hewlett-Packard Corporation, managers employing MBWA set aside time in their diaries each week to walk through their departments and engage in impromptu discussions. Although very simple (it has been described as the *'technology of obvious'*), MBWA is an extremely effective concept – particularly at times when the organisation is facing financial difficulties or reorganisation.

During their walk, managers should:

- **Listen** to what employees are saying
- **Explain** organisational policy face to face with employees
- Be prepared to offer on-the-spot **assistance** to employees.

5A.6 LEADERSHIP AND DELEGATION

There are many definitions of leadership. Most encompass the notion that a leader influences others to follow. Different models of leadership are described in detail in Section 5C.1.

LEADERSHIP

As well as communicating shared visions to their employees, successful leaders are able to foster an environment within their organisation that encourages:

- Appropriate risk taking
- Recognition and reward of success
- Empowerment that allows other leaders to emerge.

Leadership and management are separate concepts, with the former being a component of the latter. The principal aim of a manager is to maximise the output of the organisation through:

- Organisation
- Planning

- Staffing
- Leading
- Controlling.

In certain circumstances (e.g. within highly motivated groups) it is arguable whether leadership is required at all.

DELEGATION

Delegation is a key skill of an effective manager. It is the process of assigning authority and responsibility for a specific activity to another person. While a manager may delegate tasks, the ultimate responsibility for these tasks cannot be delegated.

As well as its most obvious benefit of sharing the burden of tasks, delegation can serve other purposes, such as:

- Reinforcing the role of leaders through promoting involvement
- Developing skills in team members.

However, many managers struggle to delegate successfully, fearing that a job will not be done properly by anyone else. Iles (2005) describes three rules that the effective manager should ensure when delegating any task:

1. The manager must be confident that the employee **understands** the task
2. The employee and manager must both be confident that the employee has the **skills** and **resources** necessary
3. The manager must provide **feedback** to the employee.

DELEGATION OF GOALS

Drucker (1950) described the technique of **management by objectives** (MBO). Managers following this method delegate **goals** rather than tasks. Employees are set a target to meet, but are free to choose the tactics and strategies that they will use to accomplish it. MBO has the following advantages:

- Managers avoid becoming so engrossed in day-to-day events that they lose sight of the organisation's objectives
- All employees participate in the strategic planning process
- The organisation's performance can be readily measured against defined objectives.

5A.7 NEGOTIATION AND INFLUENCING

Because of their relative lack of resource and position power (see Section 4C.9), public health practitioners must often rely on their negotiation skills to achieve their aims through influencing other people. The processes of negotiating and influencing both involve considering a situation from various points of view.

NEGOTIATION

Negotiation is the skill of resolving situations where two parties have conflicting desires. A negotiator investigates the situation with the aim of finding a solution that is acceptable to both sides. The most effective negotiating style will depend on the situation: it is influenced by the desire to meet your own needs and those of the other party.

DIFFERENT NEGOTIATING STYLES

See Figure 5A.7.1.

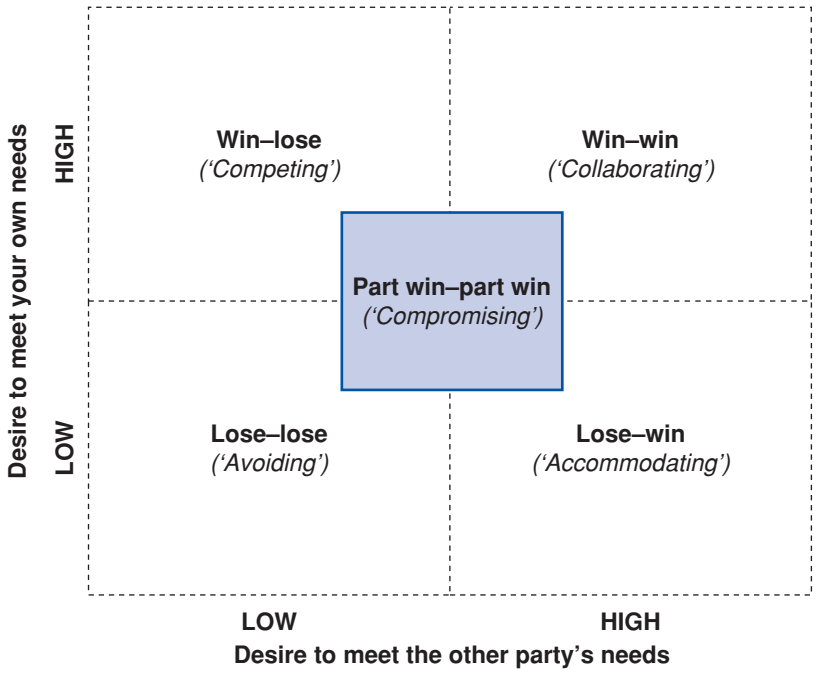


Figure 5A.7.1 Negotiating styles. *Reproduced from LSHTM Organisational Management notes*

INFLUENCE

To **influence** is to mobilise resources that modify the behaviour of others. There are three components to influence:

- **Conformity** (individuals change their behaviour in order to adhere to social norms)
- **Compliance** (request for a change in behaviour)
- **Obedience** (in order to change behaviour).

According to Covey (1989), every person has a **circle of concern** (i.e. a range of issues in which the person has emotional involvement) and a smaller **circle of influence** (i.e. a smaller range of issues that the person has the power to alter): see Figure 5A.7.2.

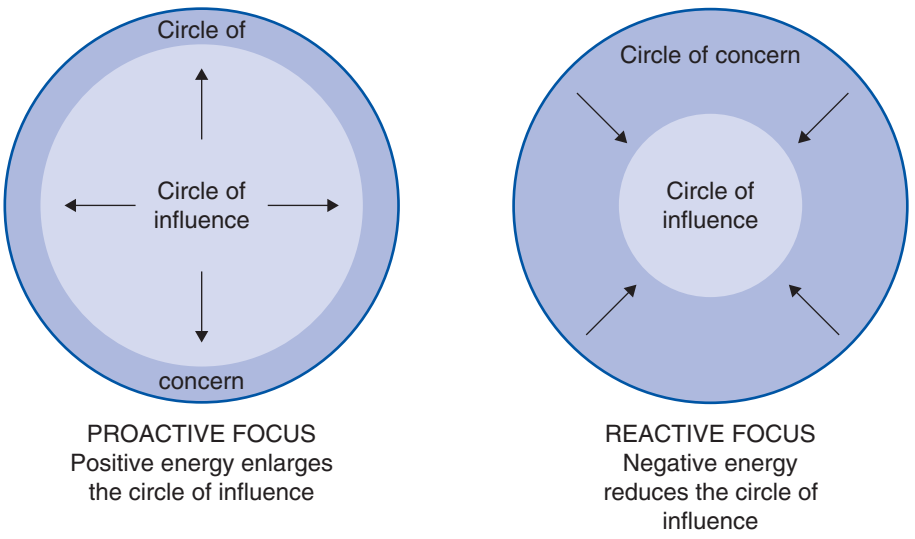


Figure 5A.7.2 Covey's circles of concern and influence. *Reproduced from Covey (1989)*

People who are **proactive** concentrate their efforts on their circle of influence – thereby enlarging it and thus becoming more powerful. In contrast, **reactive** people focus their attention on their circle of concern – to the demise of their circle of influence.

PUBLIC HEALTH ADVOCACY

Advocacy is the field within public health that aims to develop and shape public opinion in a strategic way, using whatever methods are the most effective in accomplishing this end. Even when the underlying goal being advocated is not contentious (e.g. increasing road safety), the means used to achieve it may be highly controversial (e.g. graphic television adverts that are designed to shock the audience).

Chapman (2004) identifies the following 10 steps that should be addressed when acting as a public health advocate:

1. Identify the **public health objectives**
2. Attempt to find a **win–win outcome**
3. Identify the key **decision-makers** and how they can be influenced
4. Identify the **strengths and weaknesses** of both sides of the argument
5. Set out the **media objectives**
6. Choose how to **frame** the key issue
7. Identify **symbols** and 'word pictures' to illustrate the argument
8. Compose '***sound bites***'
9. **Personalise** the topic by addressing the issue from the perspective of the ordinary citizen
10. **Mobilise** large numbers of sympathisers rapidly.

See Box 5A.7.1.

Box 5A.7.1

Example: Achieving cuts in global greenhouse gas emissions to halt climate change

While the science behind climate change is now widely accepted, the changes in policy required to reduce carbon dioxide emissions have not followed. One of the reasons for this is that the changes necessary are seen as being 'bad for business'. In October 2006, former head of the World Bank, Sir Nicholas Stern, produced a report on the potential effects of climate change on the global economy. The report is a good example of advocacy because it:

1. **Identifies decision-makers** for the changes required to address climate change: the target audiences for the report were the business and economic communities
2. **Understands how to influence** these groups: the report considered the effects of climate change on business. The arguments are presented to the decision-makers by someone who is credible to that community (Stern is a high-profile economist) rather than by an environmentalist or humanitarian
3. **Re-frames the issue** in these terms: the report highlights that it makes economic sense for the global economy to tackle global emissions now rather than leaving them unchecked until 2050

As a BBC commentary notes:

> *'The overall message of the report is not fundamentally new. In its 2001 report the Intergovernmental Panel on Climate Change (IPCC) calculated costs in the same ballpark .... The acid test of Stern is whether his economist's language can bring about the fundamental shift in political and economic direction which other financial analyses – as well as arguments posited on human rights, poverty alleviation, environmental services, health and simple concern for the natural world – have failed to do.'*

Reproduced from BBC (30 October 2006). Expert reaction to Stern review, available online at news.bbc.co.uk/1/hi/business/6098612.stm, BBC (30 October 2006); Climate costs: The next generation, available online at news.bbc.co.uk/1/hi/sci/tech/6098124.stm.

5A.8 EFFECTIVE COMMUNICATION

Principles, theories and methods of effective communication (written and oral) in general, and in a management context

See also Sections 2H.9 and 6C.

In a managerial context, the functions of communication within an organisation relate to:

- **Production** (direction, coordination and control of activities)
- **Innovation** (stimulation of change and development of new ideas)
- **Maintenance** (preservation of the values and relationships that bind the organisation).

COMMUNICATION METHODS

See Table 5A.8.1.

Table 5A.8.1 Communication methods

Formal	Communications that are routed through a so-called *'official channel'*, e.g. a written memorandum from a chief executive to the directors of an organisation
Informal	Information is passed among colleagues in an unstructured way
Diagonal	No obvious line of authority exists through which the information may be communicated
Vertical	This is the principal channel by which strategy, policy and tactics permeate from decision-makers down through the organisation to the frontline – and by which news from the frontline is fed back (see Section 5A.5)
Oral	Forms of verbal communication include speaking to another person over the telephone, or face to face in a discussion, debate, interview, presentation or meeting
Written	A form of verbal communication where the message is documented in writing, e.g. email, letters, fax
Non-verbal	There is no spoken language, e.g. eye contact, body language, sign language
Visual	A form of non-verbal communication where information is displayed in various ways, e.g. tables, advertisements. People express visual cues through their gestures and appearance (e.g. clothing, hair style, make-up). These may affect how other people receive their verbal and non-verbal communication
Internal	There are various methods of internal communication that can be used in an organisation. These include notices, bulletins, newsletters, tannoy, fax, letters, telephone, memos, email, instant messaging, intranet pages, face to face, reports, memoranda, etc.
External	Communication with outside bodies can be conducted using similar mechanisms to internal communication. However, the routes employed are likely to be more formal and less spontaneous, e.g. face-to-face communication is more likely to be in the setting of an arranged meeting than as a chat over coffee

5A.9 INTERACTIONS BETWEEN CLINICAL AND MANAGERIAL PROFESSIONALS

Interactions among managers, doctors and others

Strong links between managers and clinicians are essential in the delivery of health services. However, there is often perceived to be a conflict between the role of clinicians (particularly doctors) and that of managers. Public health

practitioners cross both managerial and clinical spheres. While this places practitioners in a strong position to align the two professional agendas, it may also place people in the middle of clinician/manager conflicts.

Tensions between doctors and managers are often described in relation to **finance** and **workload** but various authors have characterised fundamental distinctions between the groups, right from the time that they enter their professions: see Table 5A.9.1.

Table 5A.9.1 Distinctions between doctors and managers

	Doctors	Managers
Professional cultures (see Section 4B.3)	Strong established professional identity	Emerging identity
Training and learning	Specialised, prescribed body of knowledge Recognised qualifications essential to practice	No specific qualifications essential for role
Career path	Formal training route Low turnover (consultants often in posts for many years)	Different ways to enter profession (although formal NHS training exists) High turnover (22% of NHS chief executives change jobs in 3 years)
Responsibilities	Individual patient Clinic/practice	Systems Budgets

It is simplistic to describe differences just in terms of doctors and managers. Relationships between different clinical staff and management vary. Moreover, managers working in health services have often had clinical backgrounds and some clinicians adopt management roles in addition to clinical responsibilities. As Degeling et al (2003) highlight, nurses, doctors working as clinicians and those working in management positions show stereotypical differences in their views on the key elements of health service modernisation: see Table 5A.9.2.

Table 5A.9.2 Views on the key elements of health service modernisation

	Medical clinicians	Medical managers	General managers	Nurse managers	Nurse clinicians
Recognise connections between clinical decisions and resources	Oppose	Support	Equivocal	Support	Oppose
Transparent accountability	Oppose	Support	Support	Support	Oppose
Systematisation	Oppose	Oppose	Support	Support	Equivocal
Multidisciplinary teams	Oppose	Oppose	Equivocal	Support	Support

Reproduced from Degeling et al (2003).

Interactions between clinicians and managers can be improved by:

- Aligning corporate and clinical goals
- Involving clinicians in management decisions, e.g. resource allocation
- Involving managers in clinical decisions, e.g. clinical pathway development
- Shared education and training sessions
- Further developing clinical leadership.

For an example from England, see Box 5A.9.1.

Box 5A.9.1

Example: The Professional Executive Committee – clinical leadership in primary care trusts

When primary care trusts (PCTs) were set up in England, their management was divided between a trust board and a professional executive committee (PEC). The PEC is chaired by a practising clinician, often but not always a general practitioner, and the majority of its members are clinicians practising in primary or community care.

PECs were set up to ensure that clinicians were involved in making decisions about local service priorities. In turn, they enable clinicians to take responsibility for making decisions with financial implications and to become involved in some of the pressures faced by PCTs. While PECs provide a mechanism to involve clinicians in PCT management, the committees have not been without their problems:

- Some primary care clinicians remained **disengaged** with PCTs
- The PEC role has been seen as '**vague**' and sometimes overlapped with the trust board
- PECs can be **resource intensive**; as well as providing the administrative support for the meetings, PEC costs also include the costs of locum cover for the primary care contractors involved in those meetings

The function of PECs has recently been reviewed following other changes to NHS structures.

Reproduced from www.dh.gov.uk/assetRoot/04/13/89/37/04138937.pdf.

5A.10 POWER AND AUTHORITY

Theoretical and practical aspects of power and authority, role and conflict

Power is the ability to make choices or influence outcomes. See Section 4C.9 for details.

Authority can be seen as the right to make decisions and give orders. According to Weber (1958), authority manifests itself in three ways: see Box 5A.10.1.

Box 5A.10.1

Type of authority	Description	Example
Traditional authority	Authority is derived from preserved **customs**	The medical Royal Colleges rely on traditional authority
Charismatic authority	Authority comes from the personality and leadership qualities of the individual which inspire obedience and loyalty from others	Nelson Mandela. However, Adolf Hitler was also a charismatic leader; charisma is not always a force for good
Rational–legal authority	Authority is derived from powers that are bureaucratically and legally attached to certain positions	The Chief Medical Officer of a country is afforded this type of authority

5A.11 PROFESSIONAL ACCOUNTABILITY

Professional accountability – clinical governance, performance and appraisal

Professional accountability involves making individuals and organisations responsible for the quality of the service that they deliver. Systems of professional accountability have been strengthened in many countries in recent years.

UK When the NHS was first established, clinical professionals (particularly doctors) were held accountable for their actions solely through professional bodies such as the General Medical Council. In recent years, clinicians have partly ceded their high degree of professional autonomy and self-regulation in favour of:

- A regulatory system that involves members of the public
- Explicit standards of accountability such as **clinical governance**.

ROLE OF PROFESSIONAL BODIES

UK Several organisations are involved in setting and maintaining clinical standards for doctors working in the NHS. These include:

- Postgraduate Medical and Training Board (PMETB)
- National Clinical Assessment Service (NCAS)
- Medical Royal Colleges
- General Medical Council (GMC).

As well as the GMC, there are eight other major professional regulatory bodies for health care in the UK. These include the General Dental Council (GDC), the Nursing and Midwifery Council (NMC) and the Royal Pharmaceutical Society of Great Britain. Each of these bodies:

- Maintains registers of accredited professionals
- States how professional competence should be maintained
- Holds hearings when serious professional misconduct is alleged.

In the future they will require ongoing evidence of competence to practise.

In the UK, public health specialists may apply to join the **UK Public Health Register** (publichealthregister.org.uk). This is analogous to the public health specialist registers of the GMC and GDC, and its aim is to ensure high standards of public health practice so as to protect the public.

NZ The Medical Council of New Zealand requires that all registered medical practitioners participate in approved continuing professional development activities. For most practitioners this takes the form of an approved re-certification programme specific to their field of practice. Other doctors (particularly those in training positions) are required to have a formal collegial relationship with a nominated peer working in their field. Under the Health Practitioners Competency Assurance Act, this system of re-certification is being extended to other groups of health practitioners.

Accreditation of health-care agencies remains voluntary in New Zealand. A standardised system is provided by Quality Health New Zealand.

CHANGES IN PROFESSIONAL ACCOUNTABILITY

Changes in professional accountability have arisen for a number of reasons:

- Some serious and high-profile service failures have prompted suggestions that current systems for accountability were insufficient
- Recognition that the drive towards containing costs and reducing waiting times in health services in the early 1990s may have been at the expense of clinical quality
- Changes in the autonomy afforded to professionals in many fields
- Introduction of business practices into health, e.g. corporate governance.

Eng In England, recommendations to modernise professional regulation followed the **Fifth Report of the Shipman Inquiry**. This was a large-scale investigation into the reasons why Dr Harold Shipman, a GP who worked near Manchester, was able to murder an estimated 250 of his patients without detection over the course of several years. The recommendations from this enquiry are legion, but in summary they are designed to:

- **Align** systems of organisation-wide and professional regulation to ensure patient safety
- **Strengthen** systems of re-validation and fitness to practise
- **Support** the safe expansion in roles of professionals in regular contact with patients.

Following publication of this report, the Chief Medical Officer for England published a set of recommendations entitled *Good Doctors, Safer Patients*, aimed at improving systems of medical regulation (see www.dh.gov.uk/assetRoot/04/13/72/78/04137278.pdf).

APPRAISAL

Appraisal is a non-threatening, confidential dialogue that occurs between a manager and an employee, aimed at exploring and deciding:

- Progress in attaining objectives that were previously agreed
- Setting new objectives
- Identifying developmental needs.

Appraisal enables employees to achieve more from their work and to develop specific competencies. Its use is now widespread within the public and private sectors.

CLINICAL GOVERNANCE

Eng In the NHS, the profile of professional standards was raised in 1998 with the introduction of clinical governance. The government's consultation document *A First Class Service: Quality in the new NHS* (1998), defined clinical governance as:

> *'A framework through which NHS organisations are accountable for continually improving the quality of their services and safeguarding high standards of care by creating an environment in which excellence in clinical care will flourish.'*

The Health Act 1999 enshrines clinical governance as a statutory duty for all NHS organisations. Its introduction required a change in the culture of accountability, away from considering clinical standards the responsibility of single professional groups such as doctors, towards regarding it as the responsibility of the entire organisation, from the cleaners to the chief executive.

The Department of Health clinical governance support team outlines seven *'pillars'* of clinical governance, which are underpinned by five *'foundation stones'* of practice. Capping the framework is the partnership between patients and professionals. See Figure 5A.11.1.

Clinical governance is based on the principle that good **systems** of clinical care improve patient outcomes. Clinical governance involves **demonstrating** that these systems and practices are in place and are working well. For example, in terms of risk management, clinical governance would require a demonstration that an incident-reporting system (see Figure 5A.11.2) was working effectively.

Figure 5A.11.1 Seven pillars of clinical governance model. *Adapted and reproduced with permission from NHS Clinical Governance Support Team (1999), see www.cgsupport.nhs.uk/downloads/Seven_Pillars.pdf*

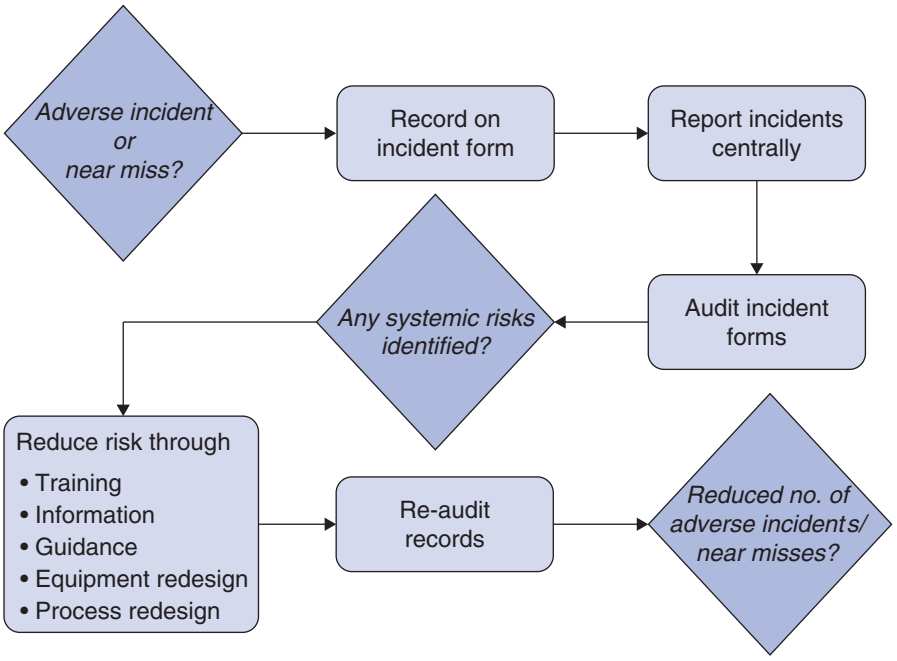


Figure 5A.11.2
Demonstration of an effective incident reporting system of risk management for clinical governance

5A.12 CHANGING BEHAVIOUR

Behaviour change in individuals and organisations

Binney and Williams (1995) describe 'types' of attitude that people may have towards change: see Box 5A.12.1.

Box 5A.12.1

Missionaries	Pleased to embrace change – they adopt it, adapt rapidly and actively encourage others to do so
Believers	Understand the merits of the changes, believe in them – but are a little more cautious, since they can see both benefits and the risks of the changes
People who pay lip service	Acknowledge that change is probably necessary but are typically not active in supporting or adopting it
Hiders and refugees	Ignore or try to hide from the change – often through fear or lack of interest
Members of the underground resistance	Actively try to block the changes
Honest opponents	Declare their resistance – they openly challenge the need for change
Emigrants	Simply *leave*, wanting nothing to do with the changes, preferring to seek their employment elsewhere

ADOPTING CHANGE

Rogers (1995) developed the '*diffusion*' model to explain how people generally move towards change: see Box 5.12.2 and Figure 5A.12.1.

Box 5A.12.2

Innovators	First to embrace change
Early adopters	Part of the first sizeable 'wave' of people who take up change, innovation – many of them becoming committed disciples of the change and innovation in question in due course
Early majority	Typically they have 'watched and waited' before either seeing the benefits and/or getting the confidence to take up the change themselves
Late majority	Follow in due course – they are less change oriented, slower to respond, need more convincing
Laggards	Those who really do not show an interest, do not want to 'get involved'

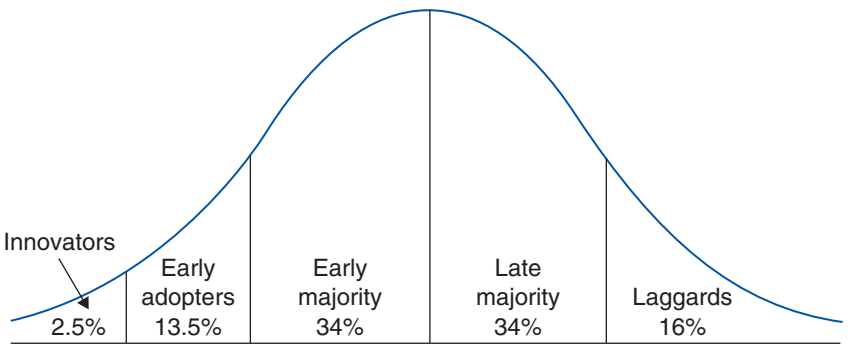


Figure 5A.12.1 Adoption curve. *Reproduced from Rogers (1995)*

ALTERING BEHAVIOURS

Theories from the behavioural and social sciences provide a platform for understanding why people engage in health-endangering behaviour or health-protective behaviour. The application of these theories can explain and predict behaviour, and can guide the design and implementation of health promotional activity. A summary is given in Table 5A.12.1 but see Section 2H.4 for more detail.

Table 5A.12.1 Summary of theories for altering behaviour related to health

	Theory	Focus	Key concepts
Individual level	Stages of change model	Individual's readiness to change or attempt to change towards healthy behaviours	Pre-contemplation Contemplation Decision/determination Action Maintenance
	Health belief model	Person's perception of the threat of a health problem and the appraisal of recommended behaviour(s) for preventing or managing the problem	Perceived susceptibility Perceived severity Perceived benefits of action Cues to action Self-efficacy
	Social learning theory	Behaviour is explained via a three-way, dynamic reciprocal theory in which personal factors, environmental influences and behaviour continually interact	Behaviour capability Reciprocal determinism Expectations Self-efficacy Observational learning Reinforcement
Organisational level	Organisational change theory	Concerns processes and strategies for increasing the chances that healthy policies and programmes will be adopted and maintained in formal organisations	Problem definition (awareness stage) Initiation of action (adoption stage) Implementation of change Institutionalisation of change
	Diffusion of innovations theory	Addresses how new ideas, products and social practices spread within a society or from one society to another	See above (Adopting change, p 467) Characteristics that determine rate of change include: • Relative advantage • Compatibility • Complexity • Trialability • Observability

5B
Understanding Organisations, Their Function and Structure

In public health practice there are many opportunities to learn how different organisations work. This could be through positive experiences such as effective liaison with stakeholders. Equally, it can come from clashes of cultures and structures, such as those that occur following organisational mergers: these too can provide valuable insights into organisations.

This chapter provides the tools for practitioners to develop systematic approaches to understanding organisations. These will be of use when forming partnerships, managing organisational change and adapting to a changing environment.

5B.1 ORGANISATIONAL ENVIRONMENTS

Understanding the internal and external organisational environments – evaluating internal resources and organisational capabilities

See Section 5C.2.

5B.2 STAKEHOLDER INTERESTS

Identifying and managing internal and external stakeholder interests

A stakeholder is a person or group that has an interest (a '*stake*') in the outcomes of a project or organisation. This interest may stem from:

- Professional interest
- Personal reasons
- Democratic representation
- Commitment to achieving a particular outcome.

Consulting stakeholders at the preliminary stages of project planning tends to garner their support, and the input from stakeholders invariably improves the quality of projects. The support of stakeholders can be helpful in securing additional resources, which in turn increases the likelihood that the project will succeed.

STAKEHOLDER ANALYSIS

The process of **stakeholder analysis** involves identifying a comprehensive list of stakeholders (both internal and external), and plotting their degree of interest and relative power: see Figure 5B.2.1.

The reaction of the stakeholders to the project may then be anticipated according to their perceived motivations from any ideological, strategic or financial interests in the project. If the reaction of a particular stakeholder is not likely to be positive, consideration can then be given to what alterations might win their support. Where winning such support is unrealistic, consideration can be given to how their opposition may best be managed, together with a contemplation of the powers of **leverage** that they may potentially exert.

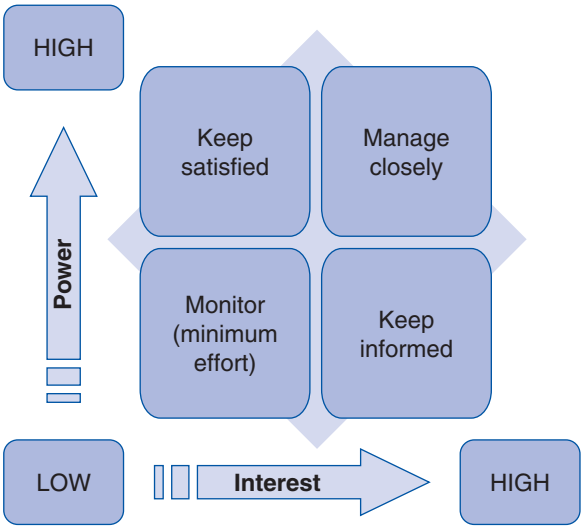


Figure 5B.2.1 Stakeholder analysis. *Reproduced from: Thompson R, MindTools.com. Available online at: www.mindtools.com/pages/article/newPPM_07.htm*

LEVERS OPERATED BY STAKEHOLDERS

The powers held by stakeholders differ between internal and external actors.

INTERNAL STAKEHOLDERS

Internal stakeholders, such as managers and individual employees, have their own interests that they will tend to pursue. For example, middle managers might seek promotion and employees may seek more favourable working conditions. All internal stakeholders possess a degree of negative power that they may be driven to use in order to impede the implementation of a particular strategy. These powers include threats of:

- Resignation
- Industrial action
- Refusal to relocate to another team or site.

EXTERNAL STAKEHOLDERS

Central government can exert influence on organisations through taxation, government spending, legal action, regulation and threatened changes in the law.

Community and pressure groups can exert influence by:

- Publicising activities that they regard as unacceptable
- Campaigning for changes in the law
- Refusing to cooperate with the organisation
- Conducting illegal actions such as sabotage.

5B.3 INTERORGANISATIONAL RELATIONSHIPS

Structuring and managing interorganisational (network) relationships, including intersectoral work, collaborative working practices and partnerships

There is growing academic and managerial interest in the field of interorganisational relationships (such as **strategic alliances**, **joint ventures** and **social networks**). This is because of their potential to benefit organisations at both the micro and the macro levels**:** see Box 5B.3.1.

Box 5B.3.1

Micro	Individual employees benefit from discussing professional practices with people from related fields. This additional insight may enable them to perform better
Macro	Possibility for innovation and efficiency gains

TYPES OF INTERORGANISATIONAL RELATIONSHIPS

Barringer and Harrison (2000) classify the principal types of interorganisational relationships as shown in Box 5B.3.2.

Box 5B.3.2

Joint venture	Two or more organisations pool a portion of their resources to form a separately owned venture. Advantages include economies of scale and of scope, as well as opportunities to innovate and to launch projects more rapidly than would otherwise be possible
Networks	These are collections of organisations that organise joint projects by means of informal arrangements rather than legally binding contracts. Such networks often operate in a hub-and-spoke configuration where a central organisation coordinates the others, with each organisation focusing on its particular specialty
Consortia	Here, organisations with a common need come together to form a new entity that satisfies that need on their behalf. For example, a group of neighbouring health authorities might form a human resources consortium because it would be more costly for the individual health authorities to operate their own
Alliances	This is an arrangement, often informal and short term, between two or more organisations that establishes an exchange relationship but no new entity is formed
Interlocking directorates	Here an executive director of one organisation sits on the board of another organisation. Such arrangements can help spread innovation and cooperation among organisations

Eng An example of interlocking directorates may be found in current government policy that encourages primary care trusts and local authorities to make joint appointments of Directors of Public Health. These appointments facilitate interorganisational work aimed at improving health and wellbeing, and are supported by the following initiatives:

- Statutory requirements placed upon local authorities to promote wellbeing
- Shared national targets
- Shared *Choosing Health* agenda.

An example of the need for interagency collaboration is shown in Box 5B.3.3.

Box 5B.3.3

Example: WHO Global Strategy on Diet, Physical Activity and Health (2004)

In an effort to reduce the global burden of non-communicable disease, the WHO released a strategy in May 2004 aimed at improving diet and promoting physical activity.

There was widespread recognition that the effort to promote physical activity at a population level required interagency collaboration and partnership working. Health promotion would succeed only by working closely with non-health sector agencies such as departments of transport, urban planning, education and sport.

Reproduced from World Health Organization (2006) Global Strategy on Diet, Physical Activity and Health, *see www.who.int/dietphysicalactivity/en/.*

5B.4 SOCIAL NETWORKS AND COMMUNITIES OF INTEREST

See also Section 4A.10.

SOCIAL NETWORKS

A *social network* is a group of people or organisations that are connected through social bonds. Relationships are viewed in terms of:

- **Nodes** (individual actors such as family members, neighbours, friends and colleagues)
- **Ties** (ranging from casual acquaintances to close familial ties).

In its simplest form, a social network is a map of all of the relevant ties between the nodes. These concepts may be displayed as a social network diagram, in which nodes are marked as the points and the ties as lines. The network can also be used to determine the social capital of individual actors (see Section 4A.10).

DUNBAR'S NUMBER

The maximum size of social networks is consistently found to be around 150 people (Dunbar 1992). This is known as **Dunbar's number** and it represents the maximum number of individuals with whom any one person can maintain stable relationships. The number was calculated in 1992 using a regression equation based on data from 38 primate genera, and it is thought to be determined by characteristics of the neocortex. Dunbar's number has subsequently been researched in the contexts of anthropology, sociology, statistics and organisational management.

ONLINE SOCIAL NETWORKS

Social networking also refers to a category of internet applications to help connect friends, business partners or other individuals together using a variety of tools. These applications, known as **online social networks**, are growing rapidly.

COMMUNITIES OF INTEREST

Communities of interest are groups of people with a common concern who may not necessarily be linked in terms of their location, profession, socioeconomic status or other characteristics. Members of a community are often scattered across a country or across the globe, and they come from many walks of life. Communities of interest have flourished with improved access to the internet. They offer members the opportunity to engage in discourse and critical thinking about topics from the perspective of their common interest.

EXAMPLES OF COMMUNITIES OF INTEREST

- Older people experiencing isolation and poverty: **older people's partnership**
- People with learning disabilities: **learning disability partnership**
- Disabled people
- Asylum seekers and refugees
- Ex-offenders
- Homeless people.

5B.5 INFLUENCES ON ORGANISATIONS

Assessing the impact of political, economic, sociocultural, environmental and other external influences

See Section 5C.2.

5C
Management and Change

Technological advancements, organisational restructuring and evolving populations mean that change is a constant feature of health-care systems. Public health practitioners have an important role in managing change in health-care systems to ensure that they continue to meet organisational goals and objectives. This chapter provides some key models of leadership and frameworks for managing change. These models can act as useful tools to evaluate how change has been managed in the past but can also be applied to addressing current situations.

5C.1 MANAGEMENT MODELS AND THEORIES

Understand the basic management models and theories associated with motivation and leadership and be able to apply them to practical situations and problems

Motivation can be regarded as a measure of a person's drive to initiate and persist in a given behaviour. Employees of an organisation may be highly able to perform a task, but if they are unmotivated then they will not fully dedicate their abilities to the job in hand.

In a managerial context, leadership is seen as the shaping of goals and development of ideas. It involves a connection with employees at an emotional level.

MOTIVATION

All organisations aim to promote motivation. This is fundamental to the role of managers in the workplace, namely to get things done through employees. There are several different theories of motivation in the workplace. Two major management theorists – **Maslow** and **Hertzberg** – have produced seminal theories to describe and predict what motivates individuals in their place of work.

MASLOW'S HIERARCHY THEORY

The **hierarchy of human needs** is among the most widely employed managerial theories. It states that:

- Only unsatisfied needs influence human behaviour
- Needs are ordered according to importance and complexity
- In the long run, people will be motivated by higher-level needs only once their lower-level needs have been satisfied (this is called **pre-potency**)
- The higher the person is up the hierarchy, the more individuality, humanity and psychological health they express.

See Box 5C.1.1.

Box 5C.1.1

	Hierarchy	Description	Work context
	**Self-actualisation**	Instinctive human need to make the most of one's unique abilities*	Promotion, opportunities for creativity/ innovation
	Self-esteem	Subjective appraisal of a person as intrinsically positive or negative	Job title, reviews, appraisals
	Love and belonging	Affection and happiness in the workplace	Professional associations, social events, supportive manager
	Safety	Absence of danger	Company pension, substantive contract
	Physiology	Basic needs such as warmth and shelter	Pay

**Definition of* 'self-actualisation' *is controversial.*

At the higher levels of the hierarchy, respect and recognition become much more powerful motivators than financial reward.

Advantages and disadvantages of the hierarchy theory are listed in Box 5C.1.2.

Box 5C.1.2

Advantages	Disadvantages
Identifies individuals who fail to progress beyond the lower levels of the hierarchy Highlights how basic problems (e.g. workplace temperature) can inhibit motivation Makes intuitive sense	Overly individualistic No allowance for altruism

Although Maslow stresses that not everyone experiences needs in this order, the very concept of an order of needs is disputed. For example, the need for warmth and shelter in a homeless person does not preclude them from simultaneously having strong needs for love and belonging.

HERZBERG'S TWO-FACTOR THEORY

This theory (also known as the **motivator–hygiene theory**) contends that certain workplace factors lead to **job satisfaction**, while others cause dissatisfaction. Factors are divided into motivators (which give positive satisfaction) and hygiene factors (which do not give positive satisfaction but the absence of which causes dissatisfaction): see Box 5C.1.3. Managers should aim to maximise both groups.

Box 5C.1.3

Motivators	Hygiene
Varied work	Good pay
Responsibility	Good working conditions
Recognition	Job security

LEADERSHIP

Leadership theory has been a topic of study throughout human history, and there are over 100 definitions of leadership. Management and leadership are sometimes used interchangeably, but Mullins (2005, p 283) distinguishes leadership from management in the following way:

> *'Management may arguably be viewed more in terms of planning, organising, directing and controlling the activities of subordinate staff. Leadership, however, is concerned more with attention to communicating with, motivating, encouraging and involving people'.*

In line with this definition, contemporary studies of leadership focus on change management and on empowering others. See Box 5C.1.4.

Box 5C.1.4

Participative theories	These models (by **Likert** and others) argue that participative styles of leadership lead to increased job satisfaction and improved performance. An example is *management by walking around* (**MBWA**) (see Section 5A.5)
Contingency theories	These theories all argue that the most effective leadership style depends on the context. For example, the managerial grid described by Blake can be used to determine whether a *boss-centred* or a *subordinate-centred* approach will work better
Instrumental theories	These contend that the leader's behaviour patterns (e.g. participation, delegation) can promote effective performance from others
Charismatic theories	These include inspirational and transformational leadership styles. By enthusing others using values and vision, the leader raises confidence in others. Charismatic leaders include those who are not appointed to authority but assume leadership in other ways
VMC model	In this model, leaders are seen as possessing the following qualities: **V**ision, **M**anagement skills and **C**ommitment. The three qualities are required in different proportions depending on the task at hand (e.g. vision is relatively unimportant in accountancy)

5C.2 FRAMEWORKS FOR MANAGING CHANGE

Critical evaluation of a range of principal frameworks for managing change

The main frameworks that are applied to change management are summarised in Table 5C.2.1.

Table 5C.2.1 Frameworks for managing change

Stage of change management	Analysis framework
Current situation	7S analysis and PESTELI analysis
Reason for change	SWOT analysis
Implementing change	Organisational development and action research

Figure 5C.2.1 demonstrates how the PESTELI and the McKinsey 7S tools can be used to support the central SWOT analysis, i.e. PESTELI and 7S can be used to ensure that each of the four components of SWOT (**S**trengths, **W**eaknesses, **O**pportunities and **T**hreats) is considered systematically and in sufficient depth.

McKINSEY 7S

This analysis identifies the **s**trengths and **w**eaknesses of the organisation, by considering the links that exist between the seven factors listed in Box 5C.2.1, each of which begins with a letter S: *strategy, structure, systems, staff, style, shared values* and *skills*. It is useful for assessing internal factors that influence performance.

Box 5C.2.1

Strategy	Action leading to allocation of resources
Structure	Hierarchy and interconnections
Systems	Procedures, processes including information flows
Staff	Personnel
Style	Style of key managers and how goals are achieved
Shared values	Guiding concepts shared by organisation's members
Skills	Capabilities of key personnel and organisation as whole

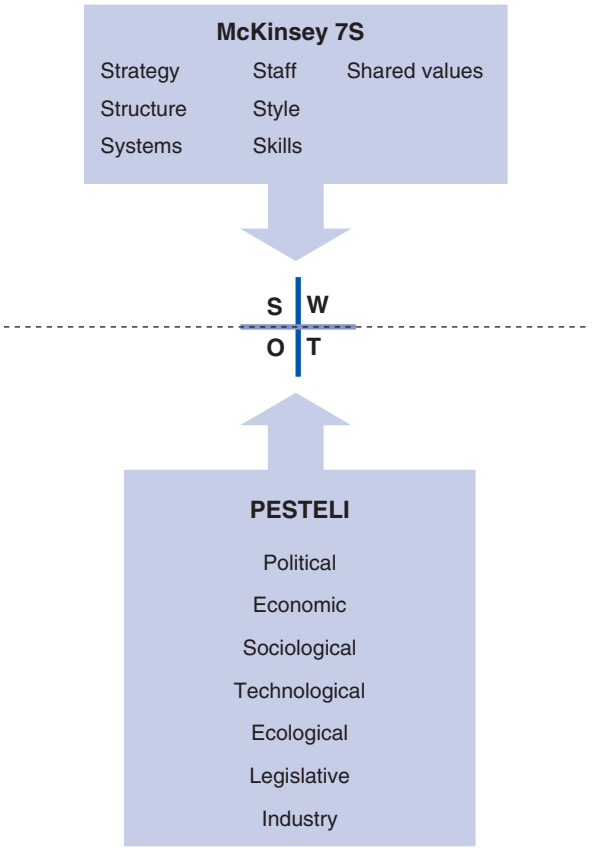


Figure 5C.2.1 Combining McKinsey 7S and PESTELI for SWOT analysis

There is conflicting evidence regarding the usefulness of the 7S model: its strengths and weaknesses are listed in Box 5C.2.2.

Box 5C.2.2

Strengths	Weaknesses
Survey of US companies in the 1970s using the 7S framework showed that the top 62 companies shared common characteristics	A repeat survey 5 years later showed that two-thirds of the previously top companies were no longer at the peak
Expert opinion generally regards the 7S approach as being useful	Ignores discord and rebellion within an organisation, which also shape the organisational culture
Dual emphasis on 'soft' and 'hard' factors	Tendency to focus on similarities between Ss and not conflicts

In health care, the 7S schema is often used to assess how changing one 'S' can impact on another

SWOT ANALYSIS

A SWOT analysis considers both internal and external factors, listing the *strengths, weaknesses, opportunities and threats* facing the organisation. These should be made with reference to the medium-term requirements of the environment in which the organisation operates. Strengths and weaknesses of SWOT analysis are listed in Box 5C.2.3.

Box 5C.2.3

Strengths	Weaknesses
Most widely used tool by UK businesses Considers both internal and external factors	Review of its use found that it generated lists of factors that were: • Too long • Too general • Often meaningless Rare that the output of the analysis is actually used Tendency to focus on the process and not the outcome

PESTELI ANALYSIS

This tool (see Box 5C.2.4) is used to identify potential **o**pportunities and **t**hreats. The acronym was initially PEST (*political, economic, sociological and technological*). *Ecological, legislative and industry* factors were added later to give a more complete description of the environment in which the organisation operates. It is useful for identifying factors that may help or impede progress but is often performed as a stand-alone activity.

Box 5C.2.4

Political factors	Affecting performance and options
Economic influences	Financial resources available and market factors
Sociological trends	Demographic changes, attitudes and beliefs
Technological innovation	Equipment, methods and approaches
Ecological factors	How the organisation interacts with the wider environment
Legislative requirements	Relevant laws that affect the organisation
Industry analysis	Attractiveness of the industry

ORGANISATIONAL DEVELOPMENT

In this approach it is the alteration of employees' on-the-job behaviour that is seen as key to implementing wider organisational change. Interventions are used to prompt desired behaviours; these can be targeted at the individual, group or organisational level. They include changes in:

- Technology
- Physical setting
- Goals and targets.

Strengths and weaknesses of the organisational development approach are listed in Box 5C.2.5.

Box 5C.2.5

Strengths	Weaknesses
Several reviews and meta-analyses showing positive outcomes	One review showed that a positive outcome was seen in 38% of cases, a negative outcome in 10% of cases, but no change in over 50% of cases.

5C.3 PERFORMANCE MANAGEMENT

An understanding of the issues underpinning the design and implementation of performance management against goals and objectives

Performance management is a form of holistic people management, which should be:

- **Integrated** (linking people management, and individuals and teams)
- **Strategic** (addressing broad issues and setting long-term goals).

For performance management to be successful, organisations need to develop a culture in which employees and teams take **responsibility** for their own contributions. The key steps involved are summarised in Box 5C.3.1.

Box 5C.3.1

Goal setting	Managers define their expectations and set strategic departmental and organisational goals
Agreement of a developmental plan	Plans agreed between managers and employees or teams
Continuous monitoring	Contemporaneous feedback and formal reviews

DESIGN

Like clinical audit, performance management is a cyclical process rather than an isolated event.

OBJECTIVES AND PERFORMANCE STANDARDS

Objectives or goals are expressed as:

- Targets to be met (e.g. achieving financial balance by the end of the financial year)
- Tasks to be completed by specified dates (e.g. by January 2004, 98% of patients are to spend under 4 hours in the emergency department)
- Ongoing targets, called **performance standards** (e.g. maintaining the intraoperative wound rate below a certain value).

Objectives should be SMART, i.e.

- **S**pecific
- **M**easurable
- **A**chievable
- **R**ealistic
- **T**imescale set for completion.

TEAMS

Team performance can be managed in connection with activity levels, outputs, patient satisfaction and financial results. Teams should agree their objectives and receive feedback in the same way as if they were individuals. Team members can contribute towards team evaluation through peer review.

360-DEGREE FEEDBACK

The concept of 360-degree feedback became popular in the 1990s. It involves collecting anonymous opinions and data from a number of sources about a person. Sources include:

- Individual's line manager
- Employees whom the individual line manages
- Peers (internal and external to the organisation)
- Self-assessment.

A 360-degree comment has the potential to provide a more rounded commentary that is less prone to bias than an assessment conducted by one individual.

IMPLEMENTATION

Performance management is difficult to implement and requires engagement by all members of an organisation – particularly line managers. The evidence suggests that performance management is well regarded by employees and managers alike, especially insofar as it emphasises personal development. However, performance-rating schemes often involve considerable amounts of red tape, which can be very time-consuming. Implementation of performance management requires:

- **Training** (especially initially) of managers and employees
- Clarification of the **definition of 'performance'**
- Understanding the organisation's **performance culture**
- Stressing the **personal benefits** to individual employees of participation in the process
- Remembering that performance management is a **tool** for line managers whose success depends on their ability to use it effectively.

5D
Understanding the Theory and Process of Strategy Development

There is a continued drive, both within health systems and externally, to improve, standardise and quality assure health care. A number of subjects covered in this chapter seek to achieve this aim, from **strategy implementation** to **risk management** and **guideline development**.

The **organisation and funding of health systems** are instrumental to their development. A range of diverse approaches to health system funding and organisation across different countries is summarised and compared.

5D.1 DEVELOPING HEALTH-CARE STRATEGY

Strategy communication and strategy implementation in relation to health care

Health-care strategies are medium- or long-term action plans that are designed to improve health or to focus on corporate priorities (such as containing health-care costs, reducing waiting lists or improving recruitment and retention of staff). In order to have an impact, a strategy needs to be communicated effectively to local stakeholders and decision-makers.

COMMUNICATION

In strategic communication, the following should be considered:

- **Audience**
- **Strategic objectives**
- **Key message**
- **Media for conveying the message**: often a range of communication channels will be employed, including press releases, annual reports and direct communication to stakeholders. See also Section 5A.8 for effective communication methods.

Strategy communication is more than presenting a 'finished product'. It is not enough for an individual or team to write a strategy document and to send it to stakeholders and decision-makers. Communication mechanisms should ensure that people responsible for delivery or affected by implementation of the strategy should be appropriately involved in:

- Agreeing strategic objectives
- Producing a baseline assessment of the current state of health-care provision
- Committing resources
- Monitoring implementation and delivery.

Communication will depend on the stage of strategic development and the audience being targeted. It typically includes the elements described in Table 5D.1.1.

Table 5D.1.1 Communication involved in strategy development

Stage of development	Audience and agencies	Communication methods and media
Strategy formation	Commissioners Partner organisations Experts and specialists Practitioners Service users	Working groups Surveys Focus groups Meetings
Dissemination	All stakeholders	Report publication (hard copy and on websites) Newsletters Media releases Public events
Monitoring implementation	Commissioners Scrutiny and review boards Public	Executive boards Local strategic partnerships Public reports

IMPLEMENTATION

Approaches to strategy implementation include all-out attack, and inside-venture and strategic alliances: see Box 5D.1.1.

Box 5D.1.1

All-out attack	In this approach (Kono 1984) all current strategic plans are abandoned and replaced with a new strategy. This typically occurs when a large organisation acquires, or is merged with, another
Inside-venture approach	This tactic (Zahra 1991) relies on the power of internal rewards and innovation. An organisation encourages employees to be innovative and, where their ideas are feasible, employees are rewarded with their own project team to develop the innovation
Strategic alliances	These include efforts such as **joint ventures** with other bodies: partnerships where two or more organisations pursue a set of agreed goals while remaining independent

Box 5D.1.2 describes the steps employed in implementing a national health strategy in England, the National Service Framework for Coronary Heart Disease.

Eng Box 5D.1.2

Example: National Service Framework (NSF) for Coronary Heart Disease

The Department of Health published a 10-year national strategy for heart disease in 2000.

The stated goals were linked to the wider strategies, *Our Healthier Nation* and the *NHS Plan*, and concerned reducing inequalities and improving outcomes for people with heart disease or at risk of heart disease through 12 standards. These were based on a baseline assessment of heart disease provision in England and reviews of available evidence.

An expert external reference group developed the strategy involving clinicians, managers, epidemiologists and patient representatives.

Dissemination and awareness-raising used a range of means including:

- **At publication:** press launch and publication of the report on the internet and in the health service circular to all NHS organisations and local authorities
- **After publication:** reports on progress and guides for implementation

The strategy's implementation was promoted through:

- **Milestones** with dates that were included in the strategy to indicate how NHS organisations were expected to progress towards meeting the standards
- Additional **funding** that was announced to support delivery of the strategy
- The standards were linked to **national performance** indicators, e.g. stop-smoking targets were a key part of the NHS star rating system for primary care trusts
- Parts of the NSF have been subject to **regular review** through Healthcare Commission/Audit Commission assessments

Reproduced from the Department of Health, National Service Framework for Coronary Heart Disease (2000), available online at: www.dh.gov.uk/PolicyAndGuidance/HealthAndSocialCareTopics/CoronaryHeartDisease/fs/en.

5D.2 THEORIES OF STRATEGIC PLANNING

Strategic planning is an activity through which an organisation confronts the major decisions that it faces. A decision is not deemed strategic merely by being important. Rather, strategic decisions or issues must fulfil the following criteria:

- **Define** the institution's relationship to its environment
- Generally take the **whole organisation** as the unit of analysis
- Depend on inputs from a number of functional areas
- Provide direction for, and constraints on, administrative and operational activities throughout the institution.

Theories of strategic planning include the **three-step models** and the **ranked order approach**.

THREE-STEP MODELS

Two models of strategy development in common usage are *STP* and *DST*: see Boxes 5D.2.1 and 5D.2.2. As can be seen, these models use similar steps in a different order.

Box 5D.2.1

STP process	
Situation	Describe the **current** situation and the factors that caused it to be thus
Target	Outline the **ideal** state of affairs and define specific goals and objectives that describe this state
Path	Map a possible **route** to the goals and objectives of the ideal state

Box 5D.2.2

DST process	
Draw	Describe the **ideal** state of affairs
See	Describe the **current** situation and the gap between the ideal state and the current state
Think	Outline the specific actions needed to **close the gap** between the current state and the ideal state

RANKED ORDER APPROACH

This approach involves defining the strategy in terms of a priority list. The following components of the strategy are first defined:

- Policies
- Plans
- Actions
- Goals
- Objectives
- Ideal state
- Strategies
- Tactics.

These items are then ranked into a hierarchy, such that:

- The item in a lower rank explains **how** the item immediately above it will be achieved
- The item in a higher rank addresses **why** the item immediately below it needs to be achieved.

In this way the **top rank objective** (TRO) does not address 'why', and therefore by definition this is the crux of the strategy upon which all else depends.

ACTION PLANNING

Once a strategy has been developed, the next stage in its implementation is the writing of an action plan. This must cover the following items:

- **Name** of the strategy
- **Actions** (tangible components of what will be done)
- **Location**
- **Personnel** (including the line-management structure of the programme)
- **Timing** (start and completion dates)
- **Resources** (staff, supplies, information, other resources)
- **Audit** (progress measurement and reporting)
- **Benefits** to the public of implementing this strategy
- **Performance rewards** (if any)
- **Contingency plans**.

5D.3 HEALTH SERVICE DEVELOPMENT AND PLANNING

The aims of health service improvement are to:

- Enhance the **quality** of patient care
- Improve strategic **outcomes**
- Contribute to improved public **health**.

HEALTH SERVICE DEVELOPMENT

Continual health care reform is necessary to ensure that services are effective, evidence based and appropriate for local needs. Drivers of change are outlined in Table 5D.3.1.

Table 5D.3.1 Drivers of change in health services

Driver	Example
Advances in technology	New drugs, equipment, procedures, screening programmes and health care settings
New information flows	Production of cost per item: can be used to inform clinicians of the financial impact of their decisions
National policy pronouncements	In the UK NICE makes assessments of both effectiveness and value-for-money
Demand management	Referral centres that act as a second gatekeeper to secondary care
Public perception	Rationing debate, particularly with regard to novel high-cost drugs

HEALTH SERVICE PLANNING

The degree to which health-care provision is planned varies between countries. However, even in the USA (which is generally regarded as an unregulated health market) there is an element of planning for large costly facilities, so that wasteful over-provision is avoided.

Eng In England, regional strategic health authorities are responsible for detailed planning. Their areas of responsibility are summarised in Table 5D.3.2.

Table 5D.3.2 Regional health authority responsibilities

Planning	Developing plans for improving health services in their area
Regulating	Assuring the quality and performance of local health services
Capacity building	Developing health services to provide adequate services for the population's needs
Priority setting	Ensuring that national priorities are included in local health service

5D.4 HEALTH SERVICE FUNDING

Methods of organising and funding health services and their relative merits, focusing particularly on international comparisons and their history

The funding and organisation of health care should be arranged so as to maximise efficiency and equity: see Box 5D.4.1.

Box 5D.4.1

Efficiency	Equity
Allocative efficiency (benefits exceed the costs) Operational efficiency (scarce resources used to their best advantage)	Financial equity (financial burden faced is proportional to ability to pay) Equity of opportunity (if not equality of access, equality of utilisation) Reducing health inequalities

ORGANISATION OF HEALTH SERVICES

A health organisation can be described in terms of its **commissioner**, its **setting**, the way in which it is **accessed**, as well as its **financial** flows.

COMMISSIONER

Eng The purchaser of health care determines which services are made available to patients. In England, the commissioner has traditionally been the local health authority or board, acting on behalf of the government. The level at which commissioning occurs is now being diversified, ranging from GPs who can choose to hold budgets for the provision of certain services, through groups of GPs, local health trusts, and up to supra-regional commissioners (who commission super-specialised services on behalf of a group of health trusts).

Wal In Wales, the commissioners of hospital, community and primary care services are local health boards.

Internationally, commissioners include mutual societies (as in France) or managed care organisations (as in the USA).

HEALTH-CARE SETTING

UK In the UK, primary health care is typically delivered in GP surgeries. Services based in the community are provided by primary care trusts in England, or by combined hospital and community trusts in Wales. Hospitals provide secondary and tertiary care in both inpatient and outpatient settings.

EU In many continental European countries, specialists work in town-centre offices rather than in outpatient departments of hospitals. **Polyclinics** also exist where several specialists share premises and diagnostic services.

Novel health-care settings include telephone-based services, hospital at home, and walk-in centres in shopping streets or transport hubs.

ACCESS TO HEALTH CARE

UK The UK, in common with some other countries, has a system of single-registration in which members of the public are entitled to register with a single GP. Although individuals are free to change GP at any time, they may not be registered with more than one GP concurrently. Advantages and disadvantages of single registration are listed in Box 5D.4.2.

Box 5D.4.2

Advantages of single registration	Disadvantages of single registration
Gate-keeping role avoids over-investigation and reduces pressure on secondary care Continuity of care Single person who receives all correspondence	Less consumer choice for patients Conflict of interest: GP is caring for the individual patient and bearing in mind implications for all other patients

GPs in the UK act as gatekeepers to elective secondary care. In contrast, members of the public in countries such as France are free to access secondary care direct. This is less cost-effective for the system but is popular with the public.

FUNDING OF HEALTH SERVICES

The five principal forms of funding health care are:

- Taxation
- Social insurance*
- Private insurance
- Out-of-pocket expenses
- Charities.

**A (more-or-less) compulsory insurance system in which employers and employees contribute to a fund. The government pays contributions for people out of work.*

Funding of health services (particularly through taxation) can be described as **progressive** or **regressive**. A tax is said to be progressive where a higher percentage of income is paid as income rises; a regressive tax is one that charges a lower percentage of income as income rises. Health-care systems that are funded through direct taxation (e.g. social insurance) tend to be more progressive than those based on direct payments or private insurance.

SHORTCOMINGS OF HEALTH-CARE SYSTEMS

A number of drawbacks inherent to health-care systems are described below.

CONSUMER MORAL HAZARD

Where the consumer does not face the full cost of a health service, there is a tendency to over-use services in the knowledge that the insurer (private or public) will foot the bill. This can lead to both over-consumption of health care when ill and a lack of engagement with preventive health care.

PROVIDER MORAL HAZARD

This occurs where a service becomes ignorant of costs to consumers, or where the remuneration system offers incentives to over-provide care (e.g. fee for service where the insurer always pays).

ADVERSE SELECTION

Where purchasers of health insurance are aware of their own personal risk, then those with low risk will tend not to purchase insurance – which is designed (and priced) to cover people of average risk. The average risk level of those remaining will rise, as will the premium to be paid, thereby exacerbating the problem. As a result, those at low risk will be uninsured, and those at high risk will be priced out of the market.

USER CHARGES

Introduction of cost sharing reduces utilization of health services by patients, but does so disproportionately among lower-income groups. Furthermore, demand is reduced for both effective treatments and minor ailments. Where supplier-induced demand exists, user charges may not reduce overall health expenditure. In lower-income countries there may be no alternative to user charges for raising funds for health care.

HISTORY

UK UNITED KINGDOM

Prior to the establishment of the NHS in 1948, patients generally had to pay for their own health care. Various charitable hospitals used to operate (e.g. the Royal Free Hospital in North London) and some local councils ran hospitals for their population: but provision was by no means universal. This all changed on 5 July 1948 when the Health and Housing Minister, Aneurin Bevan, founded the NHS. It was based on a cooperative that the coalminers ran in his hometown of Tredegar, South Wales, and followed the following principles:

- Services were provided free at the point of use
- Services were financed from central taxation
- Everyone was eligible for care (including foreign visitors and temporary residents).

The original structure of the NHS was **tripartite**: hospital service, primary care and community services. In the 1950s, rising costs led to out-of-pocket charges for prescriptions and dental treatment. To this day, these remain the major exceptions to the NHS being free at the point of use (although, as of 2006, the Welsh Assembly was planning to abandon prescription charges in Wales).

Other important NHS developments are summarised in Table 5D.4.1.

Table 5D.4.1 Key NHS developments

1950s	More equitable distribution of hospitals More medical staff Outpatient departments established
1960s	More equitable distribution of GPs Primary care teams established Shift away from large mental hospitals
1970s	English NHS re-organised: regional health authorities established Financial pressures mount
1980s	General managers appointed Internal market established
1990s	Fundholding GPs purchase care for their own patients Hospitals become semi-autonomous trusts

INTERNATIONAL COMPARISONS

A short comparison of health systems in eight countries is presented in Table 5D.4.2.

Table 5D.4.2 International comparisons

Country	Organisation		Funding provision
Australia	General	Universal free health care to all. Complex range of providers and regulators	Public health insurance system (Medicare) tax levy, which reimburses approximately 80% of health-care fees
	Primary care	Self-employed GPs	
	Secondary care	Public hospitals (70% acute beds) and private hospitals (which Medicare part subsidises)	
Canada	General	Public funding but private providers. Provincial government and Royal Colleges regulate providers	National health insurance plan (Medicare) that covers health services
	Primary care	GPs paid on fee-for-service basis	
	Secondary care	Mostly not-for-profit private hospitals	
France	General	Health care provided by private and public hospitals, and by private practitioners	National health insurance system funded by tax and compulsory social insurance from employers and employees Most citizens have supplementary mutual insurance funds that cover cost-sharing out-of-pocket expenses
	Primary care	Open access to generalists and to specialists with no gate-keeping GPs	
	Secondary care	Inpatient care provided by public and private hospitals (profit and not for profit) Outpatient care mainly provided by private specialists in their own offices	
Germany	General	Over 90% covered by statutory health insurance and remainder by private insurance. Self-regulating health-care system	General taxation and social insurance fund
	Primary care	Free access to office-based doctors (generalists and specialists) with full range of diagnostics	
	Secondary care	Outpatient care provided by office-based doctors, who act as gatekeepers to elective inpatient care	
England	General	Primary care provided by GPs. Hospitals mainly publicly owned. Government is (currently) main purchaser and provider of health care	General taxation
	Primary care	GPs act as gatekeepers. Through *Practice-based Commissioning* they can commission specialist care for their patients. Walk-in clinics and NHS Direct (helpline service) act as alternative forms of primary care	
	Secondary care	Access is through emergency ambulance, referral from a GP or self-referral to an A&E department. Hospitals are semi-autonomous trusts	

Table *contd overleaf*

Table 5D.4.2 *contd*

Country	Organisation		Funding Provision
USA	General	Health-care provision is mostly by private providers. Government funds health care through four routes: Medicare (people aged 65+) Medicaid (very-low-income people) Veterans State Children's Health Insurance Program	Voluntary private care funded by individuals and their employers. Medicare covers elderly, disabled and end-stage renal patients. Medicaid covers selected categories of the poor. Significant numbers with no health insurance
	Primary care	Except in Certain managed care plans, family doctors have no gatekeeper role	
	Secondary care	Variety of private, non-profit and public hospitals	
NZ	General	Costs of medical care largely funded by the state	General taxation revenue
	Primary care	Primary health organisations (PHOs) employ GPs and other staff Laboratory services and pharmaceuticals are largely provided at no cost to consumers	Capitated funding with co-payments for working age adults Dental care on a fee-for-service basis
	Secondary care	Public and private hospital systems. Private sector mainly provides elective surgical services	Taxation National public insurer covers costs of injury treatment Patient charges in private system
Hong Kong	General	A mixture of public and private Public sector is highly subsidised, providing primary through to tertiary care services Private sector is run as a business model, with services provided by generalists and specialists	General taxation
	Primary care	72% by private practitioners, 28% by public general outpatient clinics	
	Secondary care	82% by public (hospital authority) hospitals, 18% by private hospitals	

5D.5 RISK MANAGEMENT

Providing health care is a risky undertaking. There are risks of different kinds to the:

- Patient
- Practitioner
- Provider
- Commissioner.

Risk management involves identifying, monitoring and minimising these risks through a range of means. In England, systems of clinical governance (see Section 5A.11) provide the framework for organisational risk management.

RISKS TO PATIENTS

Patients trust health-care organisations to improve their health. However, patients are harmed in around 10% of hospital admissions. Patient safety is a particular concern because:

- There are risks associated with all types of health care
- Patients can be more vulnerable to existing hazards (e.g. many people carry MRSA in their nasopharynx, but immunocompromised patients are at greatest risk from infection).

Risks to patients cannot be eliminated but they can be minimised by ensuring that systems are reviewed and questioned regularly, e.g. by critical event audits and by learning from complaints. In England the Chief Medical Officer chaired a working group to devise recommendations to reduce adverse events in the NHS, which led to the report, *An organisation with a memory*: see Box 5D.5.1.

UK Box 5D.5.1

An organisation with a memory (2000)

This provided the platform for patient safety systems in the UK. It recommended systems to enable the NHS to learn lessons from previous incidents by:

- Focusing less on human error and more on systemic factors
- Learning from risk management in industry, particularly aviation
- Reporting incidents
- Analysing trends from reported incidents
- Ensuring that lessons learned from incidents are implemented

UK The National Patient Safety Agency was established in 2001 to promote systems of learning from incidents. The agency collates information on incidents in the NHS and shares lessons learned through:

- Agency's National Reporting and Learning System (NRLS)
- Patient Safety Observatory in the Agency, which synthesises information from incident reports sent by the NRLS, clinical negligence claims, data from death registrations, hospital activity and national surveys.

RISKS TO PRACTITIONERS

Important elements of quality insurance include:

- Ensuring that clinicians are immunised against infectious diseases
- Working in a safe environment (e.g. one that follows COSHH regulations)
- Keeping up to date.

RISKS TO THE ORGANISATION

Poor quality is a threat to any organisation. In addition to reducing risks to patients and practitioners, organisations can reduce their own risks by:

- Ensuring high-quality employment practice (including locum procedures and reviews of individual and team performance)
- Providing a safe environment (including estates and privacy)
- Ensuring adherence to safety standards and established policies.

Associated organisations (such as GP cooperatives, community pharmacists and residential care homes) should be covered by clinical governance frameworks through agreeing to comply with the standards of the organisations with which they are associated.

UK In the UK the clinical governance arrangements are complemented by and integrated with a strong risk management framework (including maintenance of risk registers introduced after the Turnbull committee made a report on corporate governance in the wake of the financial collapse of Bearings Bank). This allows clinical governance risks to feature highly in the ways that NHS organisations manage risk.

NEGLIGENCE

Negligence claims are now a feature of all health-care services. They are expensive, lengthy and undesirable for all concerned.

Eng In England, the NHS Litigation Authority is responsible for handling all claims made against the NHS in the country. According to the NHS Litigation Authority in 2006–2007:

- There were just under 9000 claims of negligence against NHS bodies
- Clinical negligence claims cost £579.3 million including damages to patients and the legal costs to the NHS
- On average it took just under 1½ years to deal with a clinical claim.

5D.6 GUIDELINE DEVELOPMENT

The concept of evidence-based medicine stipulates that guidelines based on scientific evidence should take precedence over individual judgement. In order to be acceptable to clinicians, however, clinical guidelines need to be **credible** and their development must be seen to have been **accountable.** See Box 5D.6.1.

Box 5D.6.1

Attributes of good guidelines

- Validity
- Reliability
- Clinical applicability
- Clinical flexibility
- Clarity
- Multidisciplinary process
- Scheduled review
- Good documentation

Reproduced from Field and Lohr (1990).

Instruments such as those used by SIGN (Scottish Intercollegiate Guidelines Network) and AGREE (Appraisal of Guidelines, Research and Evaluation in Europe) are based on these founding principles.

APPRAISAL CRITERIA

The *Appraisal of Guidelines for Research and Evaluation in Europe* instrument uses the criteria outlined in Box 5D.6.2. These show the steps that should be taken when developing guidelines.

Box 5D.6.2

Scope and purpose	Clear definitions of the guideline objective, clinical question and group of patients to whom the guideline applies
Stakeholder involvement	Range of professionals, together with patient involvement
Rigour of development	Systematic appraisal of evidence, explicit consideration of benefits and risks, evidence of external review prior to publication
Clarity and presentation	Specific, unambiguous recommendations
Applicability	Target users clearly defined; costs and other barriers discussed. Auditing criteria outlined. Guideline piloted
Editorial independence	Editors independent of funding body. Conflicts of interest recorded

5D.7 INTEGRATED CARE PATHWAYS

An integrated care pathway (ICP) is a **multidisciplinary outline of anticipated care**. It sets out explicit standards relating to how a patient with a specific condition is expected to progress through a clinical encounter. Typical timeframes for each step of the pathway are set out. It is important that ICPs should not be too rigid: there must be room for some degree of clinical freedom to meet the particular needs of individual patients. Hospitals favour ICPs because they have the potential to deliver health care at improved quality *and* lower cost. Advantages and disadvantages of ICPs are listed in Box 5D.7.1.

Box 5D.7.1

Advantages	Disadvantages
Facilitate clinical governance	Danger of being too rigid
Reduce unwarranted variations in patient care	Potential to disempower patients and carers
Tool for introducing clinical guidelines into everyday usage	Fail to account for the unique biology of the patient, with their special circumstances
Improve risk management	
Incorporate organisational strategy into patient care	
Avoid disputes over professional boundaries	

The ICP document itself becomes the record of all care. Clinicians are required to write in the document rather than in free-hand clinical notes. The ICP document sets out:

- **Timeframes** (e.g. how mobile the patient should be on day 3 post-surgery)
- **Decision guidance** (e.g. what should happen if a patient develops a postoperative infection)
- **Observations** (type, frequency and interpretation)
- **Investigations** (which tests should be performed on what day)
- **Referral criteria** (e.g. what should trigger a referral to a dietician)
- **Outcome measures** (e.g. patient's bowels should have opened by day 4 post-surgery).

DEVELOPMENT OF INTEGRATED CARE PATHWAYS

Writing a good ICP is a laborious process that requires a committed facilitator to coordinate the developmental steps. The process can be markedly accelerated if senior nursing and medical figures offer their public support.

STEPS IN THE DEVELOPMENT OF AN ICP

The steps involved in writing a new ICP include drafting the following:

- **Process map** (the current patient 'journey' is documented, followed by the ideal patient journey)
- **Sequence of steps** (ideal patient journey is divided into a series of stages)
- **Specific responsibilities** (roles of the different members of the multidisciplinary team are divided)
- **Problem areas** (careful attention is paid to potential variations in the patient journey).

Health services often begin by designing ICPs for conditions that are relatively common and potentially straightforward. Common examples include:

- Hip and knee replacement
- Stroke
- Coronary artery bypass graft (CABG)
- Heart failure
- Pneumonia
- Caesarean section.

EVALUATION OF INTEGRATED CARE PATHWAYS

The evaluation of ICPs is important because, although they have the potential to reduce average length of stay and to encourage evidence-based practice, they do require additional resources for their development and implementation. Assessing the impact of an ICP can be difficult since they involve many separate actions within a complex package of care. In a review of the literature (Bandolier Forum: Care Pathways, www.ebandolier.com) successful ICPs tended to:

- Examine the external evidence for individual technologies
- Combine this with local knowledge and experience and conditions
- Involve a number of different disciplines
- Measure the results of the actions
- Have information systems feeding back to the team on a timely basis
- Amend the pathway in the light of results.

5D.8 CONSULTATION ABOUT HEALTH SERVICES

Public and carer consultation and involvement in health service planning

The public can be involved in health services in a range of ways, from simply feeding back experiences of their own care to taking part in the delivery of new models of health services. Individuals can be involved as:

- Consumers of health care or carers
- Leaders or members of community groups (e.g. minority ethnic or religious groups)
- Representatives of groups with specific health interests (e.g. breast cancer support groups).

Consultation and involvement in health care are described in more detail in Section 2I.4.

5D.9 HISTORICAL DEVELOPMENT OF PERSONAL HEALTH SERVICES AND OF PUBLIC HEALTH

HISTORY OF PERSONAL HEALTH SERVICES

See Section 5D.4.

HISTORY OF PUBLIC HEALTH

See Table 5D.9.1.

Table 5D.9.1 Major events in the history of public health

Approximate date	Event
450 BC	Hippocrates studies medicine as a discipline in itself: records clinical experiences (e.g. obese people more prone to disease; differentiates epidemic and endemic diseases) and hypothesises about disease causation (e.g. the four humours: blood, phlegm, black bile and yellow bile)
1300	City of Venice introduces quarantine whereby incoming ships from ports infected with plague were required to sit at anchor for 40 days before docking
1500	Fracastorius writes about contagion in the context of syphilis
1650	Sydenham carefully describes a series of diseases and hypothesises about miasma (i.e. 'bad air' as the cause of disease)
1650	Gaunt analyses Bills of Mortality to describe disease patterns
1670	Leeuwenhoek first visualises bacteria using a microscope
1800	Jenner deliberately vaccinates a boy with the pus from cowpox sores, thereby immunising the boy against smallpox
1840	Chadwick notes differences in life-expectancy between the social classes. His campaign leads to the Public Health Act 1848 which established a Central Board of Health with powers to supervise street cleaning, refuse collection, water supply and sewerage disposal
1845	Snow identifies the cause of an outbreak of cholera in London as infected water coming from a water pump in Soho. Snow arranges for the handle of the Broad Street pump to be removed, and thereby terminates the outbreak
1850	Farr working as Registrar General (UK) uses statistical analysis as the basis for sanitary reforms
1875	Pasteur confirmed germ theory, produced artificial vaccines and described the process of pasteurisation (heat treatment of food to reduce its load of microorganisms)
1875	Koch defines four postulates that must be met to signify disease causation by a microorganism
1948	Bevan establishes UK's National Health Service
1974	Lalonde report sets out the health field concept: health is determined by the environment, lifestyle, biology and health care
1977	Alma Ata WHO Assembly sets out the vision of *'Health for All by 2000'*
1980	Black Report (made infamous by the failed attempt of the incoming Conservative government to suppress publication) outlines the correlation between social class and infant mortality rates, life-expectancy and inequalities in the use of medical services

5E
Finance, Management Accounting and Relevant Theoretical Approaches

Good health is often described as being priceless. Health care is an expensive commodity and, as such, adequate funding is essential for the adoption of new technologies and the continuation of existing programmes. Public health practitioners have an important role in promoting the use of effective, efficient technologies and successful programmes to benefit the population. Practitioners therefore need a strong grasp of the methods of **financial allocation** and **service commissioning,** and require a good insight into the **audit of health-care spending**.

5E.1 COST OF HEALTH SERVICES

Linkages between demographic information and health service information – its public health interpretation and relationship to financial costs

Demographic information can be thought of as a measure of need, and health service information as a measure of performance, and therefore they are proxies for certain outcomes.

Data linkage occurs when data about one set of features (e.g. demography) are linked to another (e.g. use of the health service or use of the A&E department). Such linkage (of demographic and health service data) can be used to assess the responsiveness of a health system to need.

DEMOGRAPHIC INFORMATION

This includes the items listed in Box 5E.1.1.

Box 5E.1.1

Population size, movement and projections	
Age, sex	
Ethnicity	
Deprivation	Employment status
	Home ownership
	Income
	Educational attainment

HEALTH SERVICE INFORMATION

This includes the items listed in Box 5E.1.2.

Box 5E.1.2

Tertiary care	Specialised units
Secondary care	Inpatient, A&E or outpatient activity Length of stay Re-admission rates Acute bed occupancy
Primary care	Referrals to secondary care Consultation rates Prescribing information

ADVANTAGES OF DATA LINKAGE

Linking demographic to health service information has both public health and financial benefits.

PUBLIC HEALTH

- **Highlights particular health or health-care issues** and prompts further research or investigation (e.g. identify inequalities in access, use of services)
- **Evaluates progress** by local agencies in improving health and reducing inequality (e.g. evaluating mental health service provision according to ethnicity, unemployment)
- Looks ahead to give **early warning** of public health problems (e.g. rising prevalence of childhood obesity)
- Uses health equity audits, needs assessment and health impact assessment.

FINANCIAL

- Identify **long-term savings** (average length of stay and re-admission rates – recognising areas of inefficiency benchmark against neighbouring PCTs, regions, national rates)
- An indication obtained of the intensity of resource utilisation can be used to **regulate the level of activity** contained within a service level agreement or contract.

WEAKNESSES OF DATA LINKAGE

While the benefits of data linkage vastly outweigh the disbenefits, some weaknesses of linking demographic and health service information include:

- Problems with data quality, reliability, access, completeness and lack of data capture
- Poor data quality can widen inequality/inequity and worsen public health
- Poor estimation of costs due to flawed data can lead to over- or under-spending on services
- Data protection requirements can make linkage difficult to negotiate, particularly where several parties act as guardians for particular data sources.

5E.2 PAYING FOR SERVICES

Budgetary preparation, financial allocation and service commissioning

Budget preparation is an important mechanism for ensuring organisational and financial management, and is crucial for achieving strategic goals. Accounting may be **financial** or **managerial**: see Box 5E.2.1.

Box 5E.2.1

Financial accounting	Financial accounting is a specialised field and relates to the use of accounting information for reporting to external bodies for auditing purposes
Managerial accounting	Managerial accounting is less technically complex and is both more amenable and more useful to the public health practitioner. By using a combination of historical data and estimated data, managerial accounting can be used to guide: • Day-to-day operations • Future operations • Organisational strategies

BUDGETARY PREPARATION

Budgets follow organisational priorities and can guide spending and decision-making. Budgets follow the fiscal year, which varies between countries: see Box 5E.2.2

Box 5E.2.2

Country	Fiscal year
UK*, Hong Kong, Canada	1 April–31 March
Australia, New Zealand	1 July–30 June
Ireland	1 January–31 December
USA	1 October–30 September

**In the UK, the fiscal year for personal tax affairs runs from 6 April to 5 April.*

In order to ensure that decisions reflect both economic realities and remain sensitive to the strategic mission of the organisation, budgetary preparation must involve both financial and managerial staff. The process begins several months before the end of the financial year, and involves the steps shown in Box 5E.2.3.

Box 5E.2.3

Service data	Collate service performance data (and compare against the objectives that were set for the year)
Fiscal data	Collate fiscal performance data (and compare against the budget that had been set for the year)
Cost per unit	Calculate the cost per unit of service by dividing the cost of the service by the number of patients seen
Service objectives	Determine service objectives for the forthcoming year based on the organisation's strategic plan
Costs projections	Estimate the costs required to achieve these objectives
Revenue/expense comparison	Compare revenue and expense projections
Budget setting	Prepare monthly budget breakdowns that reflect anticipated cash flows (not simply the full budget divided into 12 equal parts)

FINANCIAL BALANCE

Private sector organisations will, at different times, choose to:

- Incur a **deficit** (e.g. when investing prior profits into new developments)
- Realise a **surplus** (e.g. when establishing an operating reserve to guard against future cash flow shortfalls)
- Simply **break even**.

UK In contrast, government financial orders require NHS organisations to remain in continuous financial balance on every day of every year. This rule is designed to guard against bankruptcy and at the same time ensure that the organisation does not build up large surpluses rather than investing revenue in services for patients. It does, however, restrict flexibility of public bodies such as NHS organisations compared with the private sector. By using 3-year budget cycles, private companies can borrow for the future but this option is not available to NHS trusts. Trusts that are granted foundation status are, however, afforded additional financial flexibility.

COSTINGS

Budgeting must include the costs of staff, supplies and other resources (see also p 435). Managerial and financial staff should consult each other to ensure that all resources required for a service are considered. Historical or published costings can be used – with adjustments made for any impending cost changes.

In health care, staff costs typically amount to two-thirds or more of the expense budget, so this budget line item must be considered particularly carefully. Recruitment of new staff is especially costly.

See Box 5E.2.4.

Box 5E.2.4

Type of cost	Costs at a given level of activity	Example
Direct	Costs incurred exclusively for that output	Disposable surgical equipment
Indirect	Costs shared across several outputs	Autoclave that sterilises equipment from several operations
Overheads	Costs shared across the entire organisation	Cost of the organisation's press officer

If expenses need to be reduced to maintain financial balance, it is helpful to determine what each programme would cost at different service levels. A fixed percentage cut across all services is seldom the most effective way to reduce overall expenses.

See Box 5E.2.5.

Box 5E.2.5

Type of cost	Costs at changing levels of activity	Example
Fixed	Indispensable predetermined expense	Salary of a hospital's chief nurse
Semi-fixed	Predetermined expenses that are fixed in the short term	Total salaries bill
Semi-variable	Cost that is related to output but not directly proportional	Cost of equipment maintenance
Variable	Cost that is directly proportional to an output	Cost of materials

FINANCIAL ALLOCATION

UK When the NHS was founded, it assumed responsibility for voluntary hospitals, distribution of which across the country was patchy. Initially, funds had to be allocated in order to provide for services in the hospitals newly taken over. Until the 1970s, these geographical inequalities in NHS funding persisted because hospitals broadly received **historical** funding, i.e. the funding that they received in the previous year plus an allowance for growth.

In 1971 the **Crossman** formula was introduced in an attempt to rationalise the geographical allocation of resources. The formula was based on three variables:

- Population (adjusted for age and sex)
- Hospital beds (adjusted according to medical specialty)
- Cases (inpatient, outpatient and day-case).

Although transparent, this formula did not explicitly allocate resources on the basis of need and made no adjustment for deprivation.

Crossman was followed by the **RAWP** formula (*Resource Allocation Working Party*), which was based on the principle of *'equal opportunity of access to health care for people at equal risk'*. RAWP was the first **weighted capitation formula** to be used. Its intention was to distribute financial resources based on:

- Population size
- Need for health care based on standardised mortality ratios
- Unavoidable costs of providing health-care services.

The formula has subsequently been gradually adjusted (using small area census data, adjusted according to the square root of SMR, etc). Sudden swings from historical funding to weighted capitation funding would be destabilising. However, the NHS is committed to eradicating this historical legacy so current financial allocations are made according to a combination of weighted capitation, recurrent baselines, difference from target and the pace-of-change policy. For details see Section 4D.3.

CURRENT FINANCIAL ALLOCATION

UK Planned NHS spending in the UK for 2007–08 was £105.6bn – a rise from £65.4bn p.a. from 5 years earlier.

Eng The bulk of funds flows from the Department of Health as *'Unified Revenue Allocations'* to PCTs, which then fund primary and secondary care as shown in Box 5E.2.6.

MARKET FORCES FACTOR

Since a given amount of expenditure purchases different amounts of health care in different parts of the country, an adjustment called the **market forces factor** (MFF) is applied to the funding formula. The purpose of the MFF is to equalise the purchasing power of health authorities with regard to unavoidable variations in costs that are directly related to location.

COMMISSIONING OF SERVICES

Eng Since the time when the NHS was first established, the process for defining the services to be provided has evolved from one of lobbying by interested parties, through planning of district hospital services to be provided across the board, onto the separation of purchasing from providing, and into the current model of commissioning.

Eng **Box 5E.2.6**

Allocation	Fraction of total (%)	Basis of distribution	Weighted by
Hospital and community health services	82.76	Population	Age Needs Market forces Rurality
Prescribing	14.07	Population	Age Needs
Cash-limited general medical services	2.50	Population	Age Needs Market forces
HIV/AIDS	0.67	Population	Age Needs Market forces

In the health context, the term *'commissioning'* is used to describe the process by which a purchaser of health care (e.g. a health authority or a GP) identifies a local need for a service and then procures a service that meets this need. Commissioning is important because it:

- Identifies local health-care **needs**
- Exposes any **deficiencies** in current provision
- Determines the level of **investment** required to meet unmet needs
- Favours health-care services that are **cost-efficient**
- Ensures unambiguous **agreements** between purchasers and providers
- Guides **workforce development**.

PRINCIPLES FOR COMMISSIONING

The Department of Health identifies six principles that should guide commissioners of service. Each of the six is associated with a tool from the public health armamentarium: see Box 5E.2.7.

Box 5E.2.7

Commissioning principles	Public health tool
Population needs	Health needs assessment
Local service gaps	Health care evaluation
Equity	Health equity audit and equality impact assessment
Evidence based	Literature review and critical appraisal
Partnerships	Change management analysis
Value for money	Economic evaluation

FUNDHOLDING AND PRACTICE-BASED COMMISSIONING

UK In the early 1990s, organisations that previously both purchased and provided health services were split. This involved the establishment of separate:

- **Provider trusts** (e.g. acute hospital trusts and mental health trusts); and
- **Health authorities** (Wales and England) or **health boards** (Scotland and Northern Ireland) with a role confined to the purchasing of services. Later, in 2002–03 this responsibility moved to PCTs (England) and local health boards (Wales), which also provided services such as district nursing and community physiotherapy.

In the mid-1990s, the UK government experimented with the concept of **fundholding general practices** in which GPs were given a budget from which to purchase services for their patients including drugs and hospital care. In 1997 the incoming Labour government scrapped the fundholding initiative because of inequities: fundholding practices had disproportionately large budgets compared with non-fundholding practices, and were found more frequently in affluent areas. However, fundholding did achieve a relative reduction in prescribing costs and in hospital referrals between fundholders and non-fundholders.

Eng In 2004 the UK government announced a new system of local commissioning. Under this system, called **practice-based commissioning**, GP practices have the right to identify new providers of health care (including in-house arrangements or other primary care providers) to offer their patients. Practice-based commissioning allows practices to choose to commission a variety of services, and any cost savings relative to traditional (hospital) care are made available for reinvestment by the practice.

SPECIALISED COMMISSIONING

Random fluctuations in health-care activity present a relatively larger risk to small commissioners (such as a GP practice) than they do to large commissioners (such as a regional health authority). For example, in a GP list, a single patient with haemophilia requiring surgery could consume as much in health-care resources in a year as all other patients combined. In order to cope with inevitable over-spending and under-spending from one year to the next, commissioners can create a **risk pool**. This is a form of insurance for commissioners, which over-spending commissioners may access subject to explicit criteria.

UK The National Commissioning Group (NCG) covers risk-sharing arrangements for rare or expensive conditions. Until 2007, this was known as the National Specialist Commissioning Advisory Group (NSCAG). It exists to assist the groups listed in Box 5E.2.8.

Box 5E.2.8

Interested party	Benefit
Patients	Improving access to rare services
Health-care planners	Restricting the number of specialist centres so as to maintain high levels of expertise
Commissioners	Smoothing out risk
Providers	Cash flow to support rare and expensive treatments
Specialists	Focus point for discussion about service development

The 42 current services cover conditions such as:

- Craniofacial surgery service for congenital craniofacial disorders
- Epidermolysis bullosa service for children
- Extracorporeal membrane oxygenation (ECMO) service for neonates, infants and children
- Heart and lung transplantation service for adults and children

- Liver transplantation service for adults and children
- Secure forensic mental health service for young people.

5E.3 METHODS FOR AUDIT OF HEALTH-CARE SPENDING

The purpose of financial audit is to provide an independent and objective opinion on the performance of organisation with regard to:

- Financial control
- Risk management
- Governance.

INTERNAL AUDIT

In order to maintain credibility, auditors must strive for integrity, objectivity and confidentiality. A health-care organisation may be audited by:

- In-house audit teams
- In-house audit team with support from external contractors
- External 'whole-internal-audit' service.

EXTERNAL AUDIT

UK The National Audit Office (NAO) is an independent body that reports directly to Parliament. One of its roles is to assess the efficiency and effectiveness with which public sector bodies use their resources. The NAO audits the summary accounts of the NHS. In addition it appoints external auditors who audit the underlying health organisations.

FRAUD DETECTION

Fraud detection is, in general, a line-management issue rather than a direct responsibility of auditors. In the UK the NHS Counter Fraud Service is charged with tackling all fraud and corruption in the NHS.

Section 6
SKILLS TESTED AT PART A

Sections 1–5 cover the knowledge basis of public health. But practitioners also need a broad range of skills to function effectively. These skills are those of research (design and interpretation of studies), manipulation of information (data processing, presentation and interpretation) and communication.

Skills cannot be learnt through reading, but require application and practice. However, the use of the tips, formulae and principles in Section 6 will help practitioners as they develop these abilities.

6A
Research Design and Critical Appraisal

Being able to design and appraise research is a fundamental public health skill. Clearly, in offering study design advice, there is a balance to be struck between a scientific ideal and the pragmatics of conducting the study. Your effectiveness will depend on being able to sell the benefits of a good method coupled with the diplomacy of being able to achieve this within the realities of running a health service.

There is a flow in conducting research from the original source idea to a research hypothesis (aims, outcomes measurable?) and finally to the measurement of outcomes.

You need to consider the make-up of the research team (skills and personality types), as well as the ethical dimensions and how the research will be funded.

6A.1 SKILLS IN THE DESIGN OF RESEARCH STUDIES

See Box 6A.1.1.

An imperative to consider when designing a study is the target audience for the results:

- To whom will they be addressed?
- How do you intend to communicate with this audience?
- How much will this cost?
- What do you expect to be the results of your communication?

6A.2 CRITICAL APPRAISAL

Ability critically to evaluate papers, including the validity of the use of statistical techniques and the inferences drawn from them

See also Appendix A.

Box 6A.1.1

State the aim of study	Estimation of **parameters** (risk) **Association** between factor and outcome **Evaluation** (intervention such as new treatment)
Hypothesis	Evidence-based assumption to be tested by conducting the study
Choose an appropriate study design	Depends on the **aim**: Prevalence studies such as surveys, ecological or descriptive studies: *'How common is this?'* Case–control studies: *'What caused this?'* (particularly rare disease) Cohort studies: *'What effect does this have?'* (How does the level of the risk factor affect the outcome?) Interventional/experimental studies such as an RCT: *'What happens if …?'* (apply the intervention and observe the outcome, i.e. test the hypothesis) Consider what resources are available Ensure that the study is ethical
Choose study populations, sampling strategies and allocation methods	Use of **comparison** groups, i.e. controls (see Section 1A.7) **Sampling** methods (see Section 1A.25) **Allocation** (see Section 1A.26)
Measures to reduce errors (see Section 1A.9)	**Chance** (sample size, power) **Bias** (randomisation, blindness, standardisation, etc) **Confounding** (restriction, matching, stratification)
Exposure measurement in the context of observational research	Attention to aspects of **data collection** and processing (including coding, data entry, cleaning, quality control, etc)

Critical appraisal is the **systematic process** of assessing and interpreting evidence. It involves judging a paper or study with regard to three factors:

- **Validity** (could the findings be explained by chance, bias or confounding?)
- **Results** (are the statistical methods sound?)
- **Relevance** (are the findings useful to your organisation?).

An underlying principle of appraisal is expressed as '*just because it is published does not make it true*'.

The process should be objectively balanced. It must be neither overly deferential nor harshly critical of the authors.

CRITICAL APPRAISAL SKILLS PROGRAMME

The Critical Appraisal Skills Programme (CASP) developed by Milton Keynes PCT can be viewed at www.phru.nhs.uk/casp. This contains the appraisal tools listed in Box 6A.2.1, which can be downloaded free, then used for critically appraising papers.

Box 6A.2.1

- Systematic reviews
- Randomised controlled trials
- Diagnostic test studies
- Cohort studies
- Case–control studies
- Economic evaluation studies
- Qualitative research studies

In addition, the CASP workbook and CD-ROM package (which can be purchased through the website) offer practical experience of critically appraising papers similar to those found in the UK MFPH Part A examination.

CRITICAL APPRAISAL OF STATISTICAL TECHNIQUES

The number of statistical tests available is infinite, so you may well not be familiar with the exact analysis used in the paper that you are appraising. This need not be intimidating: simply interpret the statistical test in terms of a *p* value and confidence interval.

Other issues to address when critically appraising statistical techniques include:

- Do the figures in the **tables add up**?
- Was a **power** calculation used?
- Were **multiple comparisons** made without using the Bonferroni correction?
- Are **confidence intervals** given?
- Is a ***p* value** quoted for null results?
- Is the **absolute risk reduction** (as well as the relative risk reduction) quoted?
- Can you suggest an **alternative statistical test** that might have been used?

6A.3 DRAWING CONCLUSIONS FROM RESEARCH

Ability to draw appropriate conclusions from quantitative and qualitative research

In a public health examination, ask yourself why has this paper been put into today's examination? Are there any traps? Any conclusions or inferences that should *not* be drawn? Any biases or fundamental flaws? Think of all the possible explanations other than the orthodox.

Conclusions drawn from research should clearly state whether the findings support or refute the hypothesis posed. Their public health relevance can then be ascertained by considering whether the findings:

- Justify or prove the effectiveness of a programme
- Serve to refine an existing theory
- Can be used to develop a new theory.

The *BMJ* (bmj.bmjjournals.com/advice/sections.html) recommends that conclusions should be structured as follows:

- Statement of **principal findings**
- **Strengths** and **weaknesses** of the study
- Strengths and weaknesses in **relation to other studies**, discussing important differences in results
- **Meaning** of the study: possible explanations and implications for clinicians and policy-makers
- **Unanswered questions** and **future research**.

6B
Drawing Conclusions from Data

6B.1 DRAWING CONCLUSIONS FROM DATA

Ability to sort and manipulate data and to draw appropriate conclusions from quantitative and qualitative data

The ability to handle data and draw appropriate conclusions is a fundamental skill for all public health practitioners. In some instances, this may involve presenting data that practitioners have collected and analysed themselves. At other times, it requires the conclusions and data presented by others to be critically reviewed.

Basic summary statistics and graphical techniques can be used to highlight trends and make comparisons within quantitative data. For example, frequencies may be illustrated using a bar or pie chart (for categorical data) or using a histogram (for continuous data).

TRANSFORMING RAW DATA INTO MEANINGFUL INFORMATION

See also Section 1B.8.

Well-presented tables, figures and graphs enable the reader to identify patterns and contrasts in the data that would otherwise not be immediately apparent.

TABLES

When displaying data in tables, aim to follow the following rules:

- Sort in a meaningful order (e.g. largest to smallest) rather than random or alphabetical order
- Label the table correctly: rows and columns (with units), together with the title for the table itself
- Two significant figures usually provide sufficient information for the reader
- Rates are often more useful than numbers for comparing data (but the denominators must be comparable).

GRAPHS

When generating graphs, follow the principles shown in Box 6B.1.1.

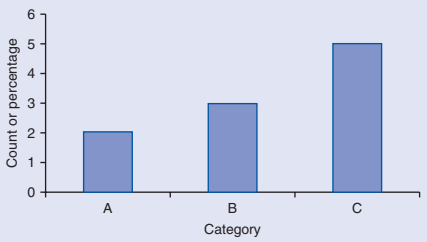
Box 6B.1.1

Clarity	Label the graph (title, axes, units) completely but succinctly Use scales appropriately – fill as much of the graph's space with data Maintain convention by starting the axes of the graph from zero Ensure that the colour or pattern of lines or bars clearly differentiates each category
Simplicity	Consider Edward Tufte's **ink-to-data** ratio: use the least ornate format of data to convey the message (i.e. table rather than bar chart, bar chart rather than pie chart) Use few gridlines Use three-dimensional designs only where a two-dimensional design would not be appropriate Be selective with content: only include relevant data
Transparency	Provide a source (and date) for the data wherever possible If data have been transformed (e.g. divided into categories or calculated as percentages), then also provide the raw data

WHICH GRAPH TO USE?

See Table 6B.1.1 and Section 1B.8.

Table 6B.1.1 Summary of display graphs

	Type of display	Strengths	Weaknesses
	Table	Display data details that would be lost in charts or text	Difficult to visualise relationships and trends
Categorical data	Pie chart	Displays percentages within a whole Easily understood, popular in lay materials, e.g. newspapers	Difficult to gauge size of pie slices by eye Difficult to compare proportions across two pie charts
	Bar graph 	Summarises large amounts of data in a visual form Trends and relationships easy to visualise Relatively simple, so accessible to a range of audiences	Scaling effects (i.e. when one variable is much greater than others, this reduces the scale of the graph and therefore makes it difficult to visualise small changes in other variables)

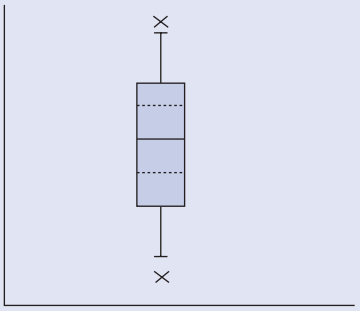
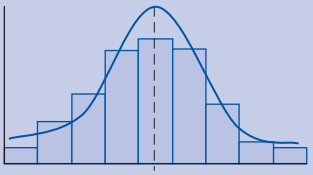
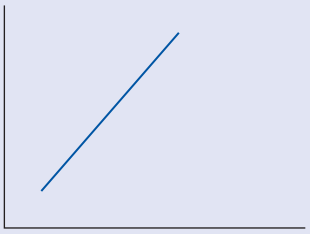
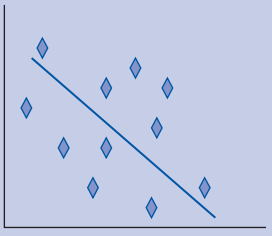
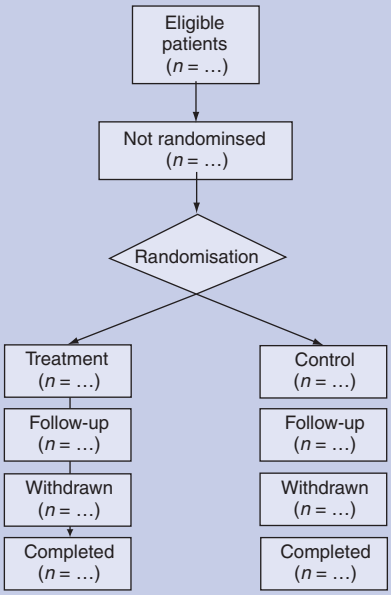
	Type of display	Strengths	Weaknesses
Univariate numerical data	Box and whisker 	Median, range of data points easily identifiable Useful to compare two or more sets of data	Relatively complex – not readily accessible for wide range of readers Exact values not retained
	Histogram 	Indicate mean, median and mode Illustrate shape of frequency distribution (e.g. symmetrical or skewed)	Does not display exact numbers, only category bands
Bivariate numerical data	Line graph 	Clear display of relationships between several dependent variables and an independent variable Displays trends (e.g. changes over time)	Temptation to extrapolate beyond data points Difficult to distinguish between lines (particularly if printed on black and white or poorly photocopied) Scaling effects – can bias results (e.g. on a 0–100 scale, a change of two points looks small; whereas on a scale of 30–40, a change of two points looks large)
	Scatterplot 	Shows minimum, maximum and outliers Illustrates relationship between variables	Hard to visualise results/ trends and relationships for large data-sets

Table *contd overleaf*

Table 6B.1.1 *contd*

	Type of display	Strengths	Weaknesses
Other	Flow chart  	Illustrates a series of steps in a procedure, decision or other 'stepwise' process Essential in RCTs to show the proportion of original units retained by the end of the study	Potential to over-simplify processes or procedures
	Maps: • Qualitative: mind maps/organisational maps • Coded mapping (e.g. MapInfo)	Shows relationships between areas Allows rates to be compared between and among regions and countries	Potential to distort (e.g. large rural areas with small populations appear bigger on a map than small urban areas with a high population)

COMMON PITFALLS IN GRAPHICAL DISPLAY

There are several ways in which a poor graphical display can cloud or distort the message that it intended to convey. These include:

- *'Chart junk'* – unnecessary graphics or text on charts
- Overuse of three-dimensional or complex graphical designs that obfuscate the main message and can bias the presentation (e.g. pie slices at the back of a three-dimentional pie chart look smaller than those at the front)
- Incorrect or insufficient labelling
- Scaling problems: too small to see trends; no zero; distorted scales – the 'gee whiz' graph (see below)
- Spurious comparisons (unequal denominators; numbers given rather than rates)
- Too many data points
- Too few data points.

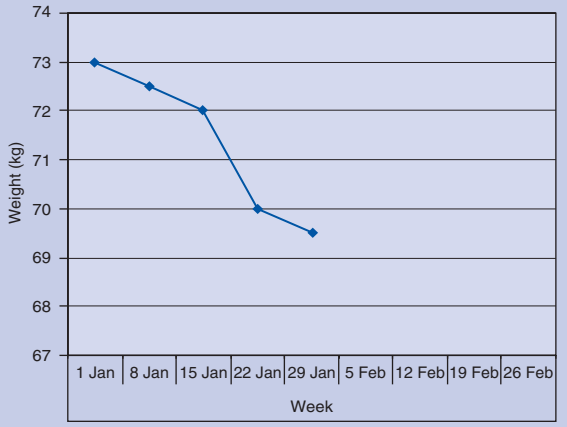
EXAMPLE 1: THE EFFECTS OF SCALING ON DATA PRESENTATION ('GEE-WHIZ' GRAPHS)

Depending on how the data in Box 6B.1.2 are presented, the effectiveness of the New Year diet plan can appear very different: see Figure 6B.1.1.

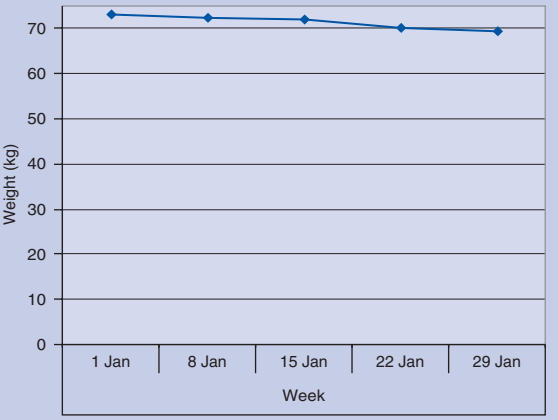
Box 6B.1.2

Individual's weight by week following New Year diet plan

	1 Jan	8 Jan	15 Jan	22 Jan	29 Jan
Weight (kg)	73	72.5	72	70	69.5



'Gee whiz' – amplified change:
Most of the *y* axis scale is missing (67–74 kg)
2 months shown; *x* axis: compressed, relative to vertical



Normal scaling – reduced change:
All of the scale is shown (*y* axis 0–75 kg)
x axis only 1 month

Figure 6B.1.1 Illustration of how the gee-whiz graph amplifies change

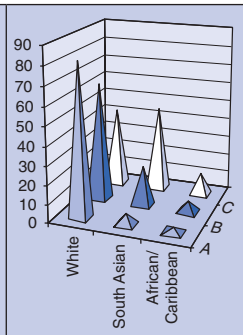
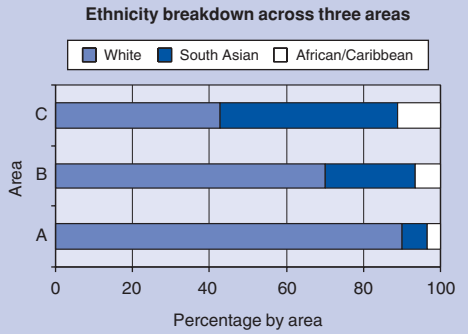
EXAMPLE 2: CHOOSING APPROPRIATE GRAPHICAL DESIGN FORMATS

See Box 6B.1.3 and Table 6B.1.2.

Box 6B.1.3

Area	Ethnicity (%)		
	White	South Asian	African/Caribbean
A	81	6.1	3.2
B	63	21	6.0
C	42	45	11

Table 6B.1.2 Choosing a graphical design format

		
	Weaknesses	**Strengths**
Design	Three-dimensional, complex design Takes up more space, confuses message	Two-dimensional, simple design Enables easy comparison of relative population constituents
Labelling	No title No label or units on *y* axis, obtrusive labelling of *x* axis	Title describing what the graph shows Axes labelled with units Legend at the bottom of graph, taking up least room possible
Scaling	Size disparity between categories: difficult to see relative differences in each ethnic category, cannot read off the values	Scale appropriate for category sizes – enables comparison between three areas

DRAWING CONCLUSIONS FROM DATA DISPLAYS

See also Appendix A.

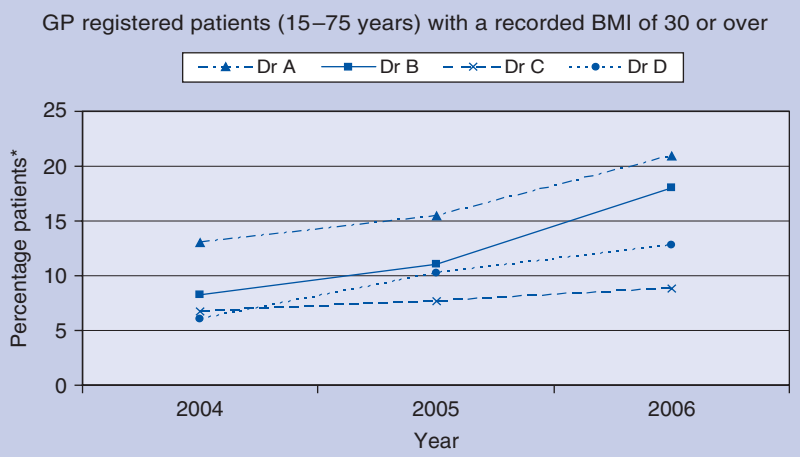
When presented with data, it is tempting to focus immediately on the detail. Instead, it is more effective to describe observations in a logical order, such as:

1. What type of display (chart/plot) is it?
2. What information is it attempting to show? Title, *x* and *y* axes and scale
3. What shapes do you observe? (e.g. bell?, skewed?, linear increase?, sharp increase or spike?, J, S shapes)
4. What similarities/differences do you notice?
   - Range of data – highest/lowest
   - Trends
5. How do you interpret what you see? Are there alternative explanations?
6. What caveats would you note?
   - Quality of the presentation (simplicity, clarity, transparency)
   - Quality of data handling, manipulation
   - Comprehensiveness of the data (other information that you would need to strengthen your interpretation).

See Boxes 6B.1.4 and 6B.1.5.

Box 6B.1.4

Example: interpreting a line graph



*Registered patients with recorded BMI 30+ as a proportion of total GP registered population

Box 6B.1.5

Suggested commentary

This is a line graph, showing the proportion of patients registered with four different general practices with a recorded body mass index (BMI) of over 30 (clinically obese) in a 3-year period (2004–2006).

In all practices the proportion of patients considered clinically obese rose from 2004 to 2006. In 2004, the proportion ranged from about 6% in Dr D's practice to 14% in Dr A's practice. In 2006 the proportion ranged from just under 9% in Dr C's practice to 21% in Dr A's practice.

The increase was approximately linear for Dr C's practice, with an increase of roughly 1% for each year. The increase was steeper in Dr D's practice in 2004–2005 than in 2005–2006. In contrast, the increase for Dr A's and Dr B's practices was greater in 2005–2006 than in 2004–2005.

The graph indicates the following:

- There appears to be a clear trend of growing obesity prevalence in all four practices
- The proportion of obese patients varies by practice and this variation has increased since 2004, possibly indicating increasing health inequalities within the area

However, the data should be interpreted with caution because:

- The graph indicates the proportion of patients registered with the GP with a recorded BMI of 30 and above; it may simply show that recording of BMI has increased. It would have been more useful to use a different denominator for the percentage calculations, e.g. people with a BMI 30+ as a percentage of all patients registered with the practice where BMI was recorded
- The data have been divided into just two categories – BMI of 30 and above and BMI below 30. It would be helpful to have information on the raw values to find the range and dispersal of BMI by practice over time
- Only 3 years of data are presented here; data from previous years would be helpful to forecast with greater confidence how the prevalence of obesity will develop over time

DATA ANALYSIS PACKAGES

For both quantitative and qualitative data (particularly large datasets), statistical packages are now available to accelerate the process of data manipulation and display: see Table 6B.1.3. However, their effective use still relies on fundamental understanding of what the data mean.

Table 6B.1.3 Statistical packages for quantitative and qualitative data

Qualitative data	Packages such as ***NUD*IST***, ***MaxQDA*** and ***NVivo*** allow textual data to be searched and sorted. Segments of interest can then be noted and marked with code words. The packages allow analyses to be performed upon these code words which can be saved, exported or analysed further
Quantitative data	Most data need to be processed before they can be analysed. Data columns (variables) need to be checked for accuracy, and may be re-arranged, re-coded or re-ordered. Some tables may need to be combined with others, especially if the data are from a relational database. Programs such as ***Microsoft Access***™ can do much of this, although programming skill is required A most versatile programme, in terms of ease of use and the fact that it is to be found on most PCs, is ***Microsoft Excel***™. This is very useful for re-shaping and re-presenting data. Its graphs and tables are easy to produce, a particular bonus for novice users In epidemiology, ***EpiInfo***™ can create questionnaires, store the data collected, manipulate, analyse and display the results In general public health, the following statistical packages are often used: **SAS**™, **SPSS**™ and **Stata**™. The relative merits of these packages depend on the type of data, the analysis that needs to be performed and the user's technical abilities
Geographical data	Geographical Information Systems (GIS) such as ***MapInfo***™ offer a range of ways of manipulating and displaying information with a geographical component

6C
Written Presentation Skills

6C.1 WRITTEN PRESENTATION SKILLS

Communicating clearly and effectively is a vital public health skill. It can make the difference between your efforts leading to significant changes and your efforts leading to a report gathering dust on a shelf.

Throughout the MFPH Part A examination (and its equivalents), candidates are tested on their written presentation skills. It is a temptation in the examination setting to skimp on the preparation and presentation of answers, but to do so is a false economy. Always ensure that you have addressed each of the following three principles before committing pen to paper: **preparation**, **organisation** and **customisation**.

PREPARATION

Before you start writing, consider the issues listed in Box 6C.1.1.

Box 6C.1.1

Brief	What are you writing for? What are the constraints, e.g. time, word count?
Audience	Who is your intended audience? Different language, structures and content will be appropriate for different audiences. Think about your audience's: • Familiarity with the subject matter • Education/understanding (is technical language appropriate? Will it alienate your audience or will it provide you with credibility?) • Culture • Point of view

ORGANISATION

Ways to **organise** your writing include those shown in Box 6C.1.2.

Box 6C1.2

Structure	Use a **clear structure** for your written presentations (in the rest of this chapter some standard formats and structures for particular audiences and purposes are given). In an examination, it is *absolutely vital* to sketch this out before you start writing an answer
Subheadings	Use **subheadings** to help your reader navigate through your writing, and to help you keep to the topic at hand
Lists	Use **lists** and **bullets** interspersed throughout your prose

CUSTOMISATION

Ensure that the **language** that you use is appropriate for your audience. Aim for the characteristics shown in Box 6C.1.3.

Box 6C.1.3

Clarity and **simplicity**	Use the simplest language appropriate for your audience: choose short words rather than long; make sentences as short as possible
Precision and **brevity**	Technical language and abbreviations (which should be defined at the first usage) may be appropriate. Only use them if they provide a more precise way of expressing yourself in fewer words, and you anticipate that your audience will be familiar with them
Neutral and **respectful language**	Avoid using words or phrases that could be construed as pejorative or insulting by some of your audience. Common pitfalls include the term 'innocent victims' (implying that there are some guilty ones); defining groups by disease or characteristics (e.g. schizophrenics or insomniacs, as opposed to people with schizophrenia or sleeping problems), and sexist terms (e.g. workmen)

CONCLUSIONS

If you have time, **check and summarise** what you have written. Note, however, that in an examination it is debatable whether an extensive summary should be included. This is because conclusions do not contain any new material and therefore do not attract new marks.

1. Refer back to original instructions periodically: does what you have written meet the original **brief?**
2. Although the executive summary will normally be at the start of your document, **write the summary last,** after reviewing what you have written. This will ensure that it closely reflects what follows in the rest of the document.
3. **Read what you have written**. In an ideal world, it is useful to leave written work overnight before checking and submitting it. Clearly, in an examination this will not be possible.
4. If you have written electronically then **use a spelling and grammar checker**. This is not a substitute for reading the document yourself but it will pick up many errors and may suggest simpler sentence structures.
5. Remove any embedded comments and tracked changes before submitting a document electronically.

6C.2 PREPARATION OF PAPERS FOR PUBLICATION

Most journals will have their own requirements on what they will accept for submission. If you are re-submitting an article to a different journal because of a rejection, then the article should be thoroughly re-drafted and re-formatted for the new journal. Goulding (2003) summarises the main points to consider for different types of scientific publication.

A paper should be structured along the lines shown in Box 6C.2.1.

Box 6C.2.1

Title

Keyword list

Used by the indexing and abstracting services, in addition to those already present in the title

Abstract

(*Most journals specify a length, typically not exceeding 250 words*)

Principal objectives and scope of the investigation

Summarise the results

Principal conclusions

Introduction

Context, background, literature review

Why was there a need to conduct the study? Introduce pertinent literature

Methods

Major study design elements:

- Sample
- Measurements
- Statistical models and testing

Include ethics approval

Results

Combine the use of text, tables and figures to condense data and highlight trends – refer to publisher's guidelines for preparing tables and figures

Discussion (see also Section 6A.3)

Generalisations that can be drawn

How findings compare with the findings of others or expectations based on previous work

Any theoretical/practical implications of your work

Consider the limitations of work

References

Check the publication's style

Reference list should contain all references cited in the text

Include with each reference details of the author(s), year of publication, title of article, name of journal or book, and place of publication of books, volume and page numbers.

Be consistent in the use of journal abbreviations

Authorship

Order: modern journals have strict rules regarding authorship and contributorship

Acknowledgements

Grant-awarding body

Clerical support, etc.

Conflicts of interest

SUBMISSION PROCESS

When dealing with reviews:

- Organise the final version of the paper and all ancillary data carefully before submission
- Incorporate any helpful comments and re-submit.

Choose the journal carefully, ensuring:

- Relevance to the subject
- Potential audience
- Impact factor.

6C.3 PRESENTING TO DIFFERENT AUDIENCES

Preparation of material for different audiences, including expert and non-expert audiences, the media and information handling and use of media in advising the public about health services, disease prevention and health promotion

SLIDE PRESENTATIONS

PowerPoint™ is the most common way of delivering a professional presentation. For slides to be effective, it is important that the presenter produces effective slides, and prepares for and performs the presentation in a format appropriate to the audience.

PRODUCING EFFECTIVE SLIDES

1. Minimise content per slide: 3–5 bullets, 5–6 words long (or use a simple graph, figure or picture)
2. Ensure that the presentation is readable:
   - Font: size 20+; sans serif; high contrast with the background
   - Template consistent on all slides
   - Inconspicuous background colour and design
3. Use slide animation sparingly if at all.

PREPARATION

1. Practise the presentation (e.g. audio or video recording)
2. Time its length
3. Arrive early on the day and check that your presentation runs on the projector and computer. Better still, bring your own tried and tested equipment.

PRESENTATION

1. Talk to the audience not the screen
2. Use slide text as key points; do not just read off the screen or from a script
3. Use a loud, low-pitched, slow voice
4. Combine with other formats (e.g. interact with the audience; include video; provide handouts).

PRESS RELEASES

General principles for press releases are described in Table 6C.3.1.

Table 6C.3.1 Principles for press releases

Information pyramid	Put the most important information at the beginning of the release, with less important material further down. (Imagine that the release could be 'cut' at any point from the bottom of the page upwards)
Keep it short	• Short **sentences** (~20 words) • Short **paragraphs** (~2 or 3 sentences each) • Short **release** (~2 pages in total, with notes to editors)
Active voice	Use the active, not passive, voice (i.e. say *'Researchers found that …'* not *'It was found that …'*)
Non-technical language	Wherever possible (e.g. use *'link'* not *'epidemiological association'; use 'breathing'* not *'pulmonary'*)
Messages	Use the release to include **established public health messages**, e.g. effects of smoking

The format and content of a release will vary to some extent depending on the organisation producing it and on the story itself. A standard layout is shown in Box 6C.3.1.

INTERVIEW BRIEFINGS

Be clear about what **information** and what **impression** you want to leave with the audience. You should plan your own responses in advance and prepare for related issues – or topical matters – that may also be asked about. Always consider the **perspective** of the interviewer and that of the audience.

Organisations are likely to have their own format for briefings, but the following kinds of information will typically be generated in consultation with the **communications officer.**

ARRANGEMENTS

- Interview for: (station/programme/presenter/newspaper)
- Time
- Date
- Telephone/studio
- Live/pre-record.

KEY POINTS

Identify **three** key points and ensure that they are:

- **In plain English**, suitable for a general audience
- **Short** and memorable.

Note that it is useful here to have *'bridging phrases'* handy. These acknowledge the question that was asked but also ensure that the key points are covered, e.g. *'and this leads me on to …'* and *'… but the real issue is …'*.

BACKGROUND

- **Target audience** of radio station/newspaper, e.g. ABC1 women, teenagers/young people, professional men …
- Agenda and **point of view** of the media (e.g. anti/pro public health issue?)
- Any other people due to be interviewed? Their viewpoint(s)?
- Topic(s) of the interview: some media will send through a list of questions/topics for the interview in advance if requested.

Box 6C.3.1.

Organisation

Date

Press release

TITLE, *e.g. NEW RESEARCH LINKS SMOKING AND COT DEATH*

EMBARGOED UNTIL … HOURS, DATE (or FOR IMMEDIATE RELEASE)

Subtitle (e.g. '*Reducing parents' smoking may cut baby deaths*')

Paragraph 1

Cover: **Who? What? Where? When?**

For example, *researchers (who) at X University (where) have linked cot deaths with smoking (what) in a paper published today (when)*

Paragraph 2

More details, answering '**How?**'

For example, *over 1000 families answered various questions about their lifestyle, health and living conditions*

Paragraph 3

More detail/**Why?**

For example, *researchers believe the effects of smoking could be …* or *It is too early to understand why smoking has these effects*

Paragraph 4

Quote from someone in authority or connected with the study

For example, *X Director of Public Health, Jane Smith, said, 'This could help reduce cot death …'* OR

John Smith, who led the study, said, 'This tells us something new about cot death'.

-ENDS-

Notes to editors:

- Contact details for more information
- Background information: e.g. *x babies die from cot death every year in the UK; smoking is the biggest preventable cause of death*

CORRESPONDENCE

LETTERS

See Box 6C.3.2.

EMAILS

Emails combine the immediacy of face to face communication with the permanence of traditional written correspondence and the audience of broadcast media. Email used well can be invaluable. Email used badly can, at best, be ignored and, at worst, alienate. The style and manner of email correspondence (see Table 6C.3.2) can ensure that emails are effective and well received.

Box 6C.3.2

Headed paper containing:

Organisation

Reply address

Date

Recipient address

Dear ... ***[title/name as they have written to you]***

Re: [subject of the letter]

Paragraph 1 **Thank you for your letter of ....**

Paragraph 2 **Acknowledge concern/query ...**

For example, 'I acknowledge that X is a particular issue ...'

Paragraph 3 **Background and evidence base for your point of view**

For example, 'X services are provided currently ...'

Finish **This is the current situation ...**

We shall, of course, keep the situation under review.

Further contact ***Please contact me if I can be of further help.***

Yours sincerely*

Name, Qualifications

Position

**Use 'Yours faithfully' for letters addressed to Dear Madam or Dear Sir*

STYLE

- Be **concise:** people receive hundreds or thousands of emails per month and can be impatient with unnecessarily long and uninterrupted text.
- Be **sensitive**: the tone of communication implicitly conveyed through speech and even handwriting is lost in emails – it can be easy to offend through overly terse text, or misplaced humour.

Retain the formality of traditional written correspondence when emailing people professionally:

- Address people whose first name you have not been given as Ms, Mr, Dr, Professor
- Follow rules of grammar, punctuation and sentence structure
- Use conventional spellings, few abbreviations, no emoticons or text language.

RESPONDING TO EMAILS

Reply within 24 h wherever possible. **Never send an email too hastily**:

- If responding to an offensive message, wait until you are calm enough to respond politely before sending a response
- Check the contents for errors
- Observe good email etiquette (see www.emailreplies.com).

Table 6C.3.2 Rules for email

Use always	Use sparingly
Email signature: set as a template including the following: • Disclaimer and confidentiality message (viruses, if received in error) • Your full name and position • Organisation name, phone number and postal address	**Attachments**: at best unnecessary attachments are just ignored, at worst they could alienate recipients by clogging email boxes or transmitting viruses
Subject header: this can determine whether your email is even opened. Good subject headers are concise, but indicate what the message will contain. Ideally they should also give some insight into what the recipient needs to do.	**Urgent priority** (marks emails as '!'): the recipient is unlikely to respond any quicker to these emails if there is no obvious reason for its use
Bcc for mass mailings: ensure that recipients do not need to scroll down a list of names before the message (write your own email address in the To: field and all other names in the Bcc field)	**Reply to all**: ensure that emails are sent only to those who need to see them
Spell checker	**Formatting and graphics:** html formats, e.g. font styles, bullets, tables, may not be retained when sent to different systems

STRATEGY DOCUMENT

1. Current position (identify issues) – **where are we now?**
2. Future priorities and objectives (including targets, evidence base) – **where do we want to be?**
3. Strategy (include time table, implementation) – **how are we going to get there?**
4. Monitoring and evaluation – **how will we know we are there?**

REPORT TO MANAGEMENT

GENERAL PRINCIPLES

- Keep to around four sides A4
- **Why is** this paper written, and **why now**?
- Main points only: you need to make it clear what exactly it is that you are asking management to decide
- Intersperse text with **bullet points**
- Use **subheadings** to break up the text.

A standard layout is shown in Box 6C.3.3.

Box 6C.3.3

Example: Standard report template

Report to: *name of management board*

Report from: *name, position, department*

Title: *covering the main subject of the paper in non-technical language*

Date:

Purpose: *what the board should do: i.e. for information, for approval of recommendations, for discussion*

Executive summary (and recommendations if appropriate)

5–7 sentences summarising the report

Enough information needs to be included for someone to read the executive summary only

Background

- Set out the context
- What is known from policy and/or research
- Issue under question

Heading(s) specific to subject

Consider the audience:

- Lay members (use non-medical language, outline relevant medical principles)
- Responsibility of the board (e.g. if commissioning, consider cost and contracting issues)

Recommendations/options

Make clear:

- Who is responsible for implementing what
- Timescales involved
- Resource implications of options

6D
Formulae Required to Pass Part A

This chapter lists all of the formulae that candidates must learn and be prepared to apply in the UK MFPH Part A examination. Formulae found elsewhere in the book that are not listed here are included only to facilitate understanding.

6D.1 SENSITIVITY AND SPECIFICITY

See Box 6D.1.1.

Box 6D.1.1

Sensitivity = proportion of those who *truly* have the disease and are picked up by the test

= the doubly positive cell / sum of both truly positive cells

$$= \frac{a}{a + c}$$

		Truth		
		POSITIVE	**NEGATIVE**	
Test result	**POSITIVE**	*a*	*b*	*a + b*
	NEGATIVE	*c*	*d*	*c + d*
		a + c	*b + d*	

Specificity = proportion of those who *truly* do not have the disease and are left alone by the test

= the doubly negative cell / sum of both truly negative cells

$$= \frac{d}{b + d}$$

		Truth		
		POSITIVE	**NEGATIVE**	
Test result	**POSITIVE**	*a*	*b*	*a + b*
	NEGATIVE	*c*	*d*	*c + d*
		a + c	*b + d*	

6D.2 POSITIVE AND NEGATIVE PREDICTIVE POWER

See Box 6D.2.1. See also Section 2C.2.

Box 6D.2.1

Positive predictive power = proportion of those *testing* positive who truly have the disease

= the doubly positive cell / sum of both test positive cells

$$= \frac{a}{a + b}$$

		Truth		
		POSITIVE	**NEGATIVE**	
Test result	**POSITIVE**	a	b	$a + b$
	NEGATIVE	c	d	$c + d$
		$a + c$	$b + d$	

Negative predictive power = proportion of those *testing* negative who truly do not have the disease

= the doubly negative cell / sum of both test negative cells

$$= \frac{d}{c + d}$$

		Truth		
		POSITIVE	**NEGATIVE**	
Test result	**POSITIVE**	a	b	$a + b$
	NEGATIVE	c	d	$c + d$
		$a + c$	$b + d$	

6D.3 NUMBERS NEEDED TO TREAT

See Box 6D.3.1. See also Section 1A.22.

Box 6D.3.1

$$\text{Number needed to treat (NNT)} = \frac{1}{\text{Absolute risk reduction}}$$

6D.4 RELATIVE RISK

Relative risk is calculated as the **risk ratio**, the **rate ratio**, the **odds ratio** (for case–control studies) and the **standardised mortality ratio** (SMR) (in an occupational setting). In cohort studies the time period must be stated (because the risk of dying will always be 100% in the long run for both groups). See Box 6D.4.1. See also Section 1A.10.

Box 6D.4.1

Risk ratio	= (Risk of disease in exposed) ÷ (Risk of disease in non-exposed)
Rate ratio	= (Incidence rate in exposed) ÷ (Incidence rate in non-exposed)
Odds ratio	= (Odds of exposure in cases) ÷ (Odds of exposure in controls)
SMR	= (Number of cases observed) ÷ (Number of cases expected) × 100%

6D.5 ODDS RATIO

The odds ratio is a measure of relative risk used in reporting case–control studies. See Box 6D.5.1. See also Section 1A.10.

Box 6D.5.1

Odds ratio	= (Odds of exposure in cases) ÷ (Odds of exposure in controls)

6D.6 ATTRIBUTABLE RISK FRACTION

See Box 6D.6.1. See also Section 1A.10.

Box 6D.6.1

$$\text{Population attributable risk fraction} = \frac{\text{Population attributable risk}}{\text{Rate of disease in population}}$$

$$\text{Attributable risk fraction} = \frac{\text{Attributable risk}}{\text{Risk in exposed group}}$$

6D.7 APPLICATIONS OF STANDARD ERROR

Standard error and confidence interval of a proportion and of a difference in proportions

See Boxes 6D.7.1 and 6D.7.2. See also Sections 1B.3 and 1B.7.

Box 6D.7.1

Proportion

Standard error (proportion) $= \sqrt{\frac{p(1-p)}{n}}$

95% confidence interval = sample value ± (1.96 × standard error)

Box 6D.7.2

Difference in proportions

Standard error (difference in proportions) $= \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

95% confidence interval = sample value ± (1.96 x standard error)

6D.8 CHI-SQUARED FOR A 2 × 2 TABLE

Box 6D.8.1

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

See Box 6D.8.1. Chi-squared is used only for **actual** numbers: not proportions, percentages, etc.

1. For each observed number calculate the expected number
2. Subtract the expected number from the observed number ($O - E$)
3. Square the result and divide this by the expected number $(O - E)^2 \div E$
4. X^2 = total of these results for all cells, i.e. sum of (3)
5. Look up X^2 (using degrees of freedom = [rows – 1] × [columns – 1]) to find the p value.

See also Section 1B.12.

6D.9 McNEMAR'S TEST

Box 6D.9.1

$$\chi^2 = \frac{(A - B)^2}{A + B}$$

Box 6D.9.2

Example of McNemar's test: calculating significance of difference between matched smokers trying to give up receiving nicotine replacement and receiving placebo

Pairs of smokers were matched according to age, sex and ethnicity. The first in the pair received nicotine replacement patches and the second in the pair received a placebo patch. All participants were assessed at 6 weeks after beginning the patches.

		First in the pair (nicotine replacement)		
		Still smoking	**Not smoking**	**Total**
Second in the pair (placebo)	**Still smoking**	**86**	**25**	111
	Not smoking	**56**	**17**	73
	Total	142	42	184

1. Ignore the concordant cells

		First in the pair		
		Still smoking	**Not smoking**	**Total**
Second in the pair	**Still smoking**	~~86~~	**25**	111
	Not smoking	**56**	~~17~~	73
	Total	**142**	**42**	**184**

2. Assign A to be 25, and B to be 56 for the equation $\chi^2 = \frac{(A - B)^2}{A + B}$

3. $\chi^2 = \frac{(25 - 56)^2}{25 + 56} = 11.9$

Since 11.9 is greater than 3.84 (1.96^2), it can be concluded that there is a significant difference at the p = 0.05 level between the matched pairs, i.e. that there is a difference between placebo and nicotine replacement

See Box 6D.9.1. McNemar's test is used for matched analyses. McNemar's test is used only for **actual** numbers: not proportions, percentages, etc.

1. Ignore the concordant cells
2. Treat one of the discordant cells as *A*, and the other discordant pair as *B* (it makes no difference which way round these are assigned)
3. Calculate χ^2 using the formula in Box 6D.9.1
4. Assess significance at the $p = 0.05$ level using the 3.84 cut-off.

An example is shown in Box 6D.9.2.

6D.10 STANDARDISATION – DIRECT AND INDIRECT

Standardisation is necessary to make fair comparisons between populations of differing demographic structures, where simply using crude mortality or morbidity rates would be misleading.

DIRECT STANDARDISATION

See Box 6D.10.1.

Box 6D.10.1

1. Begin with a reference population*
2. Break down the size of the reference population into individual age bands
3. Take the age-specific mortality rates for the comparator population, and multiply them by the size weighting of the reference population
4. Sum the values in (3) to obtain the age-standardised mortality rate.

**One of the populations being compared, their average, or an outside population.*

INDIRECT STANDARDISATION

See Box 6D.10.2.

Box 6D.10.2

1. Start with the stratum-specific death rates of a standard population (e.g. European Standard Population)
2. Use these to calculate expected number of deaths in the study population, according to its age and sex structure
3. Add up the expected number of deaths for each age band
4. $\text{SMR} = \dfrac{\text{Observed deaths}}{\text{Expected deaths}} \times 100\%$.

6D.11 WEIGHTED AVERAGES

See Box 6D.11.1 and, for an example, Box 6D.11.2.

Box 6D.11.1

Weighted mean $= \dfrac{(\bar{x}_1 n_1) + (\bar{x}_2 n_2)}{(n_1 + n_2)}$

where $\bar{x}_1$ = mean of sample 1, n_1 = number in sample 1
and $\bar{x}_2$ = mean of sample 2, n_2 = number in sample 2

Box 6D.11.2

Example: weighted averages

Two A&E departments are striving to meet a government target of seeing and treating their patients within 4 h of presentation. In the month of October, Department A sees 80% of its patients within 4 h, and Department B sees 90% of its patients within 4 h. During that month, Department A saw 3690 patients and Department B saw 2697 patients.

Arithmetic mean = 85% of patients seen within 4 h

Weighted mean = **Weighted mean** $= \dfrac{(\bar{x}_1 n_1) + (\bar{x}_2 n_2)}{(n_1 + n_2)}$

$= \dfrac{(80 \times 3690) + (90 \times 2697)}{(3690 + 2697)}$

$= 84.2\%$

6D.12 CONFIDENCE INTERVALS AND STANDARD ERRORS OF THE MEAN

See Box 6D.12.1 and Section 1B.3.

Box 6D.12.1

Standard error (mean) $= \dfrac{s}{\sqrt{n}}$

95% confidence interval = sample value $\pm$ (1.96 $\times$ standard error)

where s = standard deviation for the sample, n = number in the sample

Appendix A
Revision Tips

The revision tips here are generally focused on the UK MFPH examination, although the principles of exam technique will apply to any exam.

The UK MFPH Part A examination is notorious for its low pass rate: barely one candidate in five passes at some sittings. The elaborate marking algorithm ensures that candidates who do not perform consistently well in all questions will be heavily penalised. Accordingly, we advise you to ensure that:

- Your knowledge of the syllabus is **broad** rather than deep
- You divide your time in the examination across the questions (and sub-sections of questions) **proportionately** according to the marks available for each sub-question.

EXAMINATION STRUCTURE

UK Passing the exam (known formally as the *'Part A Examination for Membership of the Faculty of Public Health of the Royal Colleges of Physicians of the United Kingdom'*) entitles you to apply for:

- Diplomate membership of the Faculty of Public Health (DFPH)
- Entry to the Part B examination – success in which leads to full membership of the Faculty (MFPH).

Although the Part A examination actually consists of four papers (one each morning and afternoon over two consecutive days), the Faculty's marking scheme refers simply to papers I and II, one on each day.

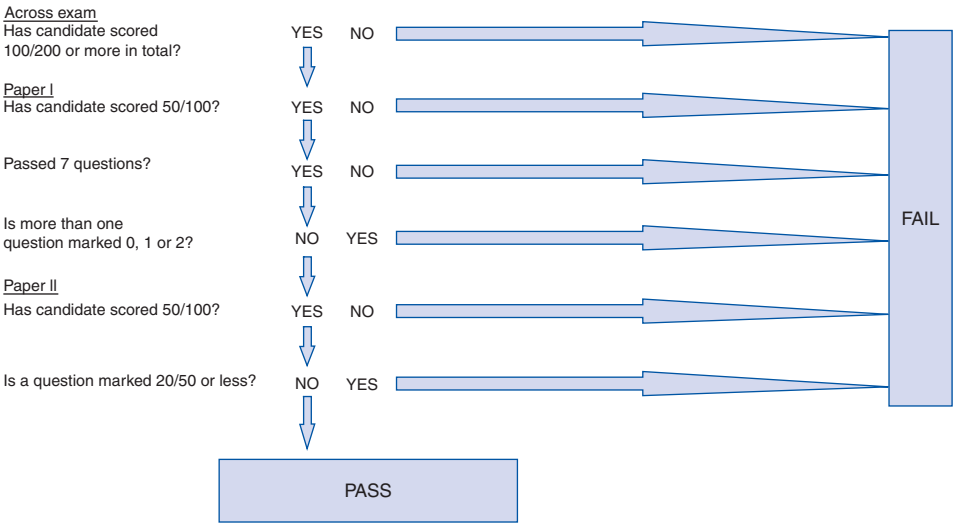
DAY 1

Session	Paper	Duration	Description	Details	Marks	Timing
am	**Paper Ia**	2½ hours	Short-answer questions (SAQs)	Research methods (epidemiology, statistics, qualitative research, health information sources), health promotion and health protection	60	6 questions in 150 min = 25 min per question
pm	**Paper Ib**	1½ hours	SAQs	Medical sociology, social policy, health economics and organisational management of health care	40	4 questions in 90 min = 22½ min per question

DAY 2

Session	Paper	Duration	Description	Details	Marks	Timing
am	**Paper IIa**	2½ hours	Critical appraisal and discussion	Candidates are provided with a journal paper to read (usually from the *BMJ*) Questions include composing a structured abstract of the paper, critically appraising the paper and then addressing more general questions on the topic	50	50 marks in 150 min = allow approximately 1 hour to read and digest the paper, then roughly 20 min per 10 marks
pm	**Paper IIb**	1½ hours	Data handling and strategy writing	Candidates are provided with evidence in the form of raw data, tables of data, maps or graphs Questions involve interpretation of these data (which may include statistical manipulation) and then the writing of a policy document based on the above	50	50 marks in 90 min = roughly 20 min per 10 marks

Each of the two papers carries 100 marks, so that the examination is marked out of 200 in total. There is a complex marking algorithm which contains six hurdles. A candidate who stumbles on any of these hurdles will fail the examination.



Reproduced from www.fphm.org.uk/exams/downloads/PtI_marking_algorithm.pdf.

In order to pass the examination overall, candidates must pass both paper I and paper II. However, if candidates pass only one of the papers (score 50/100 or more) but also obtain an overall score of at least 100/200, then that paper can be **banked**. Once a paper has been banked, candidates need only attempt the remaining paper when they re-sit the examination on a subsequent occasion.

In practice the intricacies of the algorithm are irrelevant, and the bottom line is that you must avoid doing badly in any question.

REVISION

The amount of time required for revision obviously depends on your prior knowledge and experience, and on your personal revision style. That said, most successful candidates seem to devote at least 4–6 months to revision. A good way to gauge how much work you will need to do is to familiarise yourself with the syllabus and with one of the past papers and examiners' comments. We would advise you not to look at the most recent paper until a week before the exam, when you should use it as a mock under self-imposed examination conditions.

On page 544 is a suggested revision schedule that worked for us – and which you might want to use or adapt to your personal revision style.

Revision phase	Subject matter	Details
Initial	Syllabus	• Read the **syllabus** and scan through the contents of this book
	Sample past paper	• Read a sample **past paper** (not the most recent), and the associated examiners' **comments** to gauge the required standard
Early	Epidemiology	• Ensure that you are *au fait* with all of the epidemiological components of the syllabus. **Epidemiology** is the backbone of the examination and it is tested not only in the epidemiology questions in paper Ia, but also throughout paper II • Study Chapter 1 of this book and read the recommended epidemiology text below
Mid	Cover the whole syllabus	• Study the rest of this book • Consult the further reading as recommended below • Work your way through the **CASP** workbook and CD-ROM (crucial for paper IIa)
	Past papers	Again setting aside the most recent past paper, work your way through the previous five or six past papers as follows: • For each question, write down the **introductory sentence** that you would use and the **essay structure** that you would employ • Compare what you have written with the examiners' key points and comments
Late	Context	You can add weight to your answers by adding material and examples from the following: • Selected editorials from the last few months' editions of the ***BMJ*** (see Dr Edmund Jessop's weekly reading list [www.edmundjessop.org.uk]) • Items on the **websites** of the following organisations: King's Fund (www.kingsfund.org.uk), Food Standards Agency (www.food.gov.uk), National Statistics (www.statistics.gov.uk), National Institute for Health and Clinical Excellence (www.nice.org.uk), Healthcare Commission (www.healthcarecommission.org.uk), Health Protection Agency (www.hpa.org.uk), Department of Health (www.dh.gov.uk) and the WHO (www.who.int)
Week before	Mock paper	Exactly 1 week before the exam you should attempt the most recent past paper in 'real time' and under **exam conditions**. Write your answers longhand and use only the materials that will be available to you in the examination itself (www.fph.org.uk/exams/part_1/exam_preparation.asp)
	Finalise	Consolidate and review your revision notes
	Short-term memory	Learn the following: • Question timings/marks-per-minute rates for each paper • Statistical formulae • Key definitions (infant mortality, etc.) • Essay structures

We know from personal experience that this book contains more than enough information to pass the MFPH Part A. The books below were useful for us for addressing uncertainties.

Subject area	Further reading
Epidemiology	Hennekens C, Buring J (1987) *Epidemiology in Medicine*. Philadelphia: Lippincott Williams & Wilkins
Statistics	Swinscow TDV, Campbell MJ (2002) *Statistics at Square One*. London: BMJ Books Campbell MJ (2001) *Statistics at Square Two*. London: BMJ Books
Health information	www.connectingforhealth.nhs.uk; www.ic.nhs.uk
Health promotion	Nutbeam D, Harris E (2004) *Theory in a Nutshell: A Guide to Health Promotion Theory*. Maidenhead: McGraw-Hill Education
Environmental public health	Chapter within Donaldson LJ, Donaldson RJ (2003) *Essential Public Health*. Petroc Press
Communicable disease	Hawker J et al (2005) *Communicable Disease Handbook*. Oxford: Blackwell Science
Medical sociology	Scambler G (2003) *Sociology as Applied to Medicine*. Philadephia: WB Saunders
Qualitative research	Green J, Browne J (2005) *Principles of Social Research*. Milton Keynes: Open University Press
Health economics	Wonderling D et al (2005) *Introduction to Health Economics*. Milton Keynes: Open University Press
Health care management	Iles V (2005) *Really Managing Health Care*. Oxford: Oxford University Press
Critical appraisal	CASP: Evidence-based Health Care Workbook and CD-ROM. Update Software Ltd (www.update-software.com)
Public health practice	Pencheon D et al (2006) *Oxford Handbook of Public Health Practice* (2nd ed). Oxford: Oxford University Press
Strategy writing	Model answer Paper 2a, June 2005
Clear communication	*The Economist Style Guide* (2003) London: Economist Books

EXAMINATION TECHNIQUE

Unfortunately the examination papers rarely start exactly at the published time: a delay of 5–10 min is usual. This delay can make the timing of questions rather tricky, so you will need to spend the first minute or two of the exam calculating the exact times at which you should start each question. Write down these times alongside each question on the question paper. Note that the short answer questions (SAQs) are of different durations in paper Ia (25 min each) and paper Ib (22½ min each).

You need to remain acutely aware of the time throughout the examination. Therefore, during the MFPH Part A examination you should regard your watch as the equivalent of the rear-view mirror in a driving test: force yourself to keep looking at it frequently. This is the only way to ensure that you allow a proportionate amount of time to each question and sub-question.

PAPER I

As the questions in all parts of paper I are compulsory, there is little to be gained from reading through the whole question booklet at the start. Instead we would suggest treating each SAQ as a mini-exam, and then starting each new question afresh at the calculated time.

The first 5 minutes of each SAQ should be spent on planning. You are advised to use this time to:

- Re-read the question and underline the keywords

- Brainstorm the question, i.e. write down the key elements that you think the examiners are expecting you to cover, together with any 'gems' that you can throw in from your own work experience or from your revision (particularly the contextual material that you may have read in the later stages of revision) to impress the examiners
- Choose a structure (either one of the structures in Appendix B or a bespoke structure for that question)
- Craft the first sentence of your answer very carefully: first impressions count
- There is generally no need for a conclusion.

Note that where a question asks you to answer in relation to a 'country of your choice', you are expected to state that country explicitly at the beginning of your answer.

PAPER II

In contrast to paper I, you should read through all of the questions as soon as you are allowed to turn over the paper. This is because the questions build on each other and therefore give an indication of the examiners' line of thought.

For **paper IIa** we would advise you to set aside the first **hour** to read and digest the study that you are being required to appraise critically. We suggest that you follow Dr Edmund Jessop's advice (see www.edmundjessop.org.uk/partI_main.doc) of reading the following parts of the paper first:

- Title
- Last paragraph of the introduction
- First paragraph of the discussion
- Last paragraph of the discussion.

This will provide you with the gist of the paper (and therefore the bulk of the abstract that you will be required to write) and you should try to make this completely clear in your mind before you read any further.

For **paper IIb** you should begin by studying and describing in a systematic way any data presented to you (see 6B1). Medically qualified candidates will be familiar with the standard way to report a chest radiograph, namely:

- Type of image
- Name and date of birth of the patient
- Date of examination
- Striking features
- Systematic approach to bones, soft tissues, zones of the lungs, etc.

For example, the report may read, *'This is a postero-anterior chest radiograph of Mr David Jones (DOB 24/7/46) taken on 14 June 2006. The most striking abnormality is a left-sided pneumothorax. The bones appear normal …'*.

You should adopt a similar methodical approach to whatever data source you are asked to describe in paper IIb – paying particular attention to any axes, units or denominators shown.

Later in the paper – when it comes to writing the policy document – you should build on your answers to earlier questions. You should refer to what you wrote earlier in the exam but you must not regurgitate the same material.

TECHNIQUE FOR SHORT ANSWER QUESTIONS

For each SAQ you must convince the examiners that you understand:

1. What exactly the question is asking
2. Why this question is being asked today (i.e. why it is important to contemporary public health)
3. How to set about tackling the question in a logical fashion.

We would encourage you to approach each SAQ by covering these three points. Always begin with a carefully constructed sentence that teases out the issues in the question and explains why they are important to the world of

public health today. Next, set out the structure that you are going to adopt to answer the question, i.e. the answer framework.

You should then proceed to answer the question, writing out and underlining the individual headings contained in your box as you go along.

Aim for a clear writing style that follows the advice of the *Economist's Style Guide,* i.e. you should:

- Use short, snappy sentences
- Avoid trite turns of phrase
- Steer clear of unnecessary jargon.

The examiners encourage you to employ the *'rationed use of bullet points, tables and diagrams'* to illustrate your answers. You should aim to include relevant examples, and to name any eponymous theories or structures that you use.

Example (January 2006, Paper Ia, Question 1): Describe how you would undertake a formal survey to determine the prevalence of angina in a local area of a developed country (population 100 000), e.g. the United Kingdom.

The public health importance of angina is that:

(a) *it is an indicator of unmet need in relation to the treatment of ischaemic heart disease (currently the commonest cause of death in developed countries)*

(b) *in most cases, it is treatable either medically or through revascularisation (surgically or percutaneously)*

(c) *differences in its prevalence can therefore be used to compare inequalities in access to health care within and across populations.*

1. Definitions (setting; angina; formal survey)

2. Steps:
- *Obtaining a sampling frame*
- *Obtaining a sample*
- *Assessing presence of angina*
- *Calculation of prevalence*

3. Strengths and weaknesses of method chosen

Definitions

I shall conduct this survey in the relatively deprived inner-London borough of Islington, UK. Although there are many types of angina (e.g. Ludwig's angina, Prinzmental's angina), for the purpose of this study, angina will be defined as stable angina of cardiac origin, specifically …

Appendix B
Answer Frameworks

A recurrent gripe mentioned in the examiners' comments is that candidates' answers lack structure. To avoid this, candidates must ensure that – without exception – every word that they write during the examination fits into a structure.

Answers can be organised using one of the frameworks below, or a framework can be custom-made for a particular question during the examination. Either way, it should be made explicit to the examiners what framework is being used (e.g. by drawing a box at the start of each answer and writing the framework for that answer inside the box). The contents of the box should be repeated as headings within the answer, and these should be underlined.

GENERIC FRAMEWORKS

The following two structures can be used either as the stand-alone framework for an entire question (especially if none of the more specific frameworks fits) or alternatively as the subheadings for a component of another framework.

Jessop framework

Definition
Example
Advantages
Disadvantages

Areas of public health

Health intelligence
Health services
Health promotion
Health protection

PAPER I FRAMEWORKS

NEEDS ASSESSMENT

Needs for a population group

Physical	Public health	Basic (food, water, shelter)
		Lifestyle (smoking, alcohol)
		Screening
		Immunisations
	Primary care	Medical
		Dental
		Pharmacy
	Secondary care	Physical
		Mental
	Mental	
	Social	

Needs for a disease

Epidemiological	Definition
	Numbers
	Current set-up
	Alternative set-ups
Comparative	Other places
	Gold standard
Corporate	Central government
	Health authority

HEALTH CARE EVALUATION

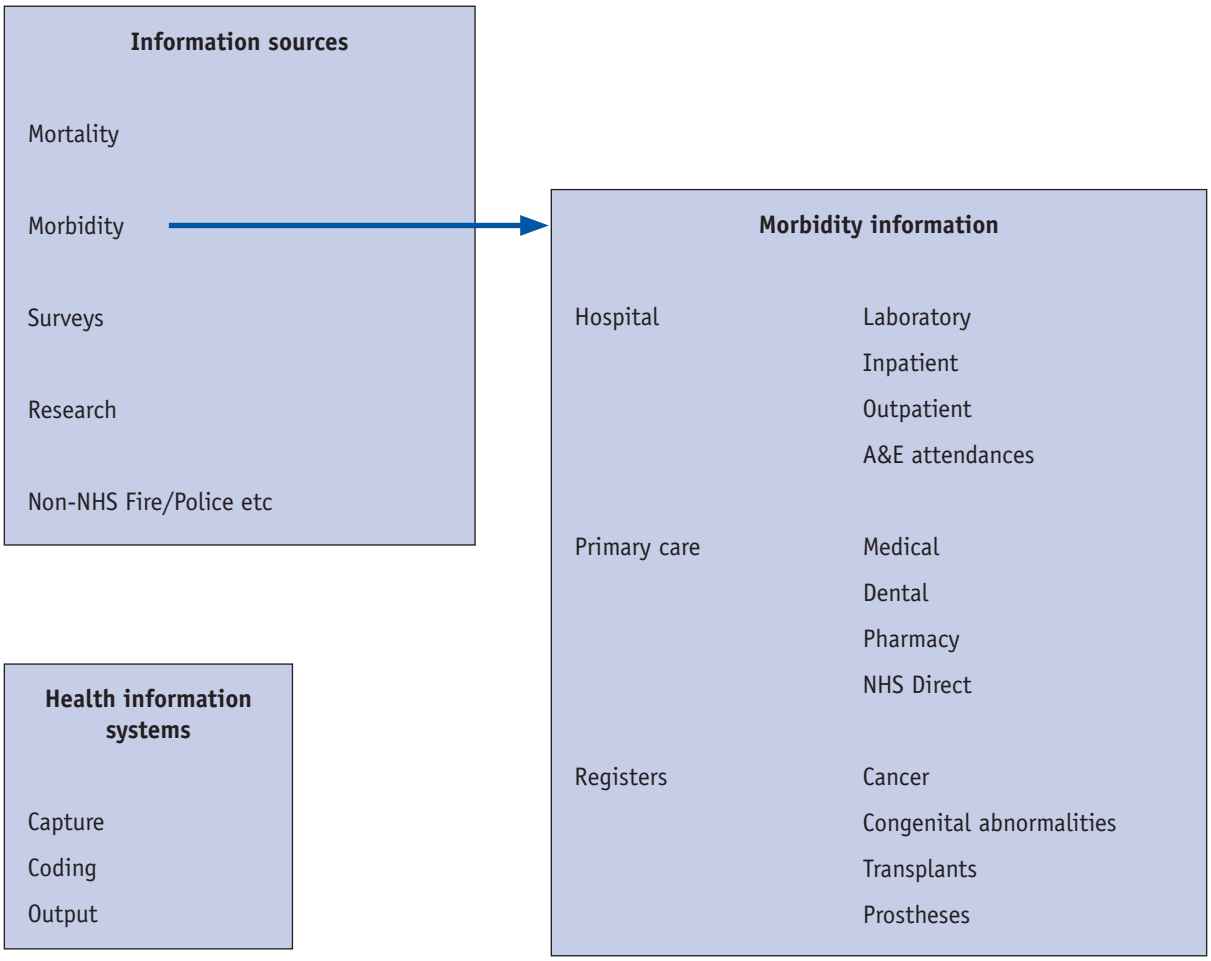
The names of Maxwell or Donabedian should be quoted if their frameworks are used.

Donabedian	
Structure	Staff numbers
	Staff qualifications
	Bed capacity
Process	Admissions
	Procedures
Outcomes	Survival
	Quality of life

Maxwell	
A	Access
	Acceptability
	Appropriateness
E	Equity
	Efficiency/economy
	Effectiveness
R	Relevance

Ongoing evaluation ('HADEPO')	
Health	Public health indices
Access	Equality/equity
Delivery	Evidence-based practice
Efficiency	Costs
Patient experience	Questionnaires
Outcome	Audit cycle

HEALTH INFORMATION



COMMUNICABLE DISEASE

Epidemiology of a communicable disease

'AgORMICS'

Agent	Virus/bacteria/protozoa
	Illness caused
	Method of diagnosis
Occurrence	In named country
	Seasonal pattern/sporadic/imported
Reservoir	
Mode of transmission	Parenteral/Faeco-oral/other
Incubation	(Omit if unsure)
Communicability	e.g. communicable while still excreting in stool
Susceptibility	e.g. infection confers resistance

Acknowledgement: Edmund Jessop's notes

Control of a communicable disease

'PIDQuICS'

- Prevention
- Isolation
- Disinfection
- Quarantine
- Immunisation
- Contacts
- Specific measures

Acknowledgement: Edmund Jessop's notes

Outbreak investigation

CCDCs HATE IT

- Count cases
- Control outbreak
- Diagnosis verified
- Communication
- Surveillance enhanced
- Hypothesis formulation
- Additional microbiology samples sent
- Test hypotheses
- Epidemic confirmation
- Identify cases
- Tabulate data

Acknowledgement: NCL revision course notes

MISCELLANEOUS

Needs	Harms	Lalonde health fields	Epidemiology
Felt	Source	Lifestyle	Time
Expressed	Path	Health service	Place
Normative	Receptor	Environment	Person: Age, Sex, Class, Ethnicity, Occupation
		Genes	Intervention

PAPER IIA FRAMEWORKS

ABSTRACT

Abstract (taken from *BMJ*)

Objective
Design
Setting
Participants
Intervention
Outcome measures
Results
Conclusions

CRITICAL APPRAISAL

RCT

Screening	Focused issue (population/ intervention/outcome)
	Randomisation
	Follow-up complete
Fairness	Triple blinding
	Patient characteristics equal both groups
	Identical management except intervention
Results	Treatment effect
	Confidence intervals and *p* value quoted
Locally applicable	Similar population
	All potential outcomes considered (e.g. QoL)
Cost:benefit	

Meta-analysis

Screening	Focused (population/ intervention/outcome)
	Randomised
Rigour	Relevant studies included
	Databases
	Reference lists
	Personal contacts
	Unpublished work
	Non-English work
	Assessed quality
	Reasonable to merge (similar results)
Results	Bottom line (odds ratio/ NNT)
	Confidence intervals
Locally applicable	Similar population and setting
	All outcomes
Cost:benefit	

Acknowledgement: Critical Appraisal Skills Programme (CASP) www.phru.nhs.uk.

PAPER IIB FRAMEWORKS

Health impact assessment

Screening	*(Define the problem)*
Steering group	*(Define the options)*
Negotiation	*(Options appraisal)*
Implementation	
Monitoring	
Evaluation	
Dissemination of evaluation	

Health needs assessment (HNA)

Objective	
Setting	
Methods	(*HNA for a group or for a disease – as in the frameworks for paper I)*
Results	
Conclusions	

Service specification

Pathway	Diagnosis	
		Patient empowerment
		Medication
		Interventions
Clinical guidelines		Best practice
		Type and dose of drug
		Skills
		Minimum activity
Clinical governance		
Skill mix		
Accreditation		

Strategy

Current	*Where are we now?*
Future	*Where are we going?*
Strategy	*How do we get there?*
Monitoring }	*Are we there?*
Evaluation }	

Briefing

Preamble
- Title
- Author
- Audience
- Date

Purpose
- Strategy
- Methods
- Results
- Conclusions

Executive summary

Introduction and background

Key issues

Options

Recommendations

Timetable

Letter

Business letter layout

Thank you

Acknowledge concerns

Background to the issues raised

Evidence

Statement of current situation

Situation will be kept under review

Please contact again if

Appendix C
Top Fives

In the examination, you will need to quote standard criteria, classic studies and key names in order to demonstrate your knowledge of the specialty of public health. We suggest that you use the lists below as a starting point for creating your own lists which you then memorise prior to the examination.

EPIDEMIOLOGY

	Concept	More information in Section
John Snow (1854)	Epidemiological method: cholera and Broad Street pump	5D
Richard Doll and Austin Bradford Hill (1954)	Doctors' cohort: Smoking causes lung cancer	
Archie Cochrane (1972)	Evidence-based medicine	1A
David Barker (1987)	Risk of coronary heart disease in adults is linked to *in utero* development	2A
Geoffrey Rose (1992)	Population-based prevention, and the prevention paradox	2H

UK PUBLIC HEALTH POLICY

	Concept	More information in Section
Chadwick (1842)	Sanitation and health	5D
Tudor Hart (1971)	Inverse care law	1C
Black (1980)	Health inequalities are growing in the UK and are linked to social class	2I
Acheson (1998)	Reducing health inequalities under the new Labour government	2I
Wanless (2004)	Economic case for investing in public health	2I

HEALTH PROMOTION: FRAMEWORKS

	Concept	More information in Section
Tannahill (1985)	Health promotion = overlapping spheres of protection, prevention and education	2H
Lalonde (1974)	Health field concept	2H
Dahlgren and Whitehead (1991)	Policy rainbow	2H
Evans and Stoddart (1990)	Health field model	2H
Diderichsen and Hallqvist (1998)	Social determinants	2H

HEALTH PROMOTION: MODELS

	Concept	More information in Section
Hochbaum et al (1958)	Health belief model	2H
Bandura (1970s)	Social learning	2H
Prochaska and DiClemente (1984)	Stages of change	2H
Beattie (1991)	Dimensions of health promotion: authoritative – negotiated; individual – collective	2H
Ewles and Simnet (1995)	Health field concept	2H

SOCIOLOGY

	Concept	More information in Section
Talcott Parsons (1951)	The sick role	4A
Edwin Lemert (1967)	Primary and secondary deviance	4A
Irving Goffman (1963)	Stigma; institutionalisation	4A
Ivan Illich (1975)	Iatrogenesis	4A
Emile Durkheim (1897)	Social integration and suicide	
John Rawls (1971)	Social justice	4C

MANAGEMENT

	Concept	More information in Section
Donabedian (1966)	Health service quality: structure–process–outcome	1C
Maxwell (1984)	Health service quality: access, equity, efficiency, effectiveness, economy, appropriateness, acceptability	1C
Belbin (1996)	Team roles	5A
Maslow (1943)	Hierarchy of needs	5A, 5B
Handy (1987)	Organisational types	5A

KEY STUDIES

STUDY DESIGN EXAMPLES

Study design	Example	Participants	Major finding(s)	More information
RCT	**Women's health initiative** (2003)	64 500 women over 15 years	Contrary to observational study evidence, HRT does *not* protect against CHD	Wassertheil-Smoller S et al 2003. *JAMA* 289:2673–2684
Cohort	**Whitehall** (1967 onwards)	Whitehall I: 18 000 male civil servants Whitehall II: 10 000 male and female civil servants	Risk of death from CHD linked to social status/ employment grade	www.ucl.ac.uk/ whitehallII
Case–control	**UK Childhood Cancer Study** (1999)	*Cases*: records of 30 000 children with cancer *Controls*: children matched for age, sex, area of birth	Proximity to powerlines at birth linked to childhood leukaemia	Draper G et al 2005. *BMJ* **330**:1290
Cross-sectional	**Health Survey for England** (annual)	6000 adults + 3000 children (in 2004 – random selection and boosted sample in high minority ethnic group areas)	Prevalence of risk factors and health behaviours, e.g. fruit and vegetable consumption	www.ic.nhs.uk/ pubs/healthsurvey 2004ethnicfull/ hse2004vol1/file
Case reports	***Pneumonia jirovec* pneumonia (PJP) in Los Angeles** (1981)	Five homosexual men aged 29–36	Abnormal epidemiology of PCP, early sign of emergence of AIDS	www.cdc.gov/ MMWR/preview/ mmwrhtml/june_5. htm

EFFECTS OF DIET ON HEALTH

	Study design	Major finding(s)	More information
North Karelia (1972–1982)	Intervention (before/after): residents of North Karelia	Altering lifestyle (including saturated fat intake) reduced CHD mortality	Section 2E
Framingham (1948 …)	Cohort: 5000 residents of Framingham, MA	Factors affecting risk of CHD include cholesterol	www.nhlbi.nih.gov/about/ framingham/index.html
7 Countries (1958–1970)	Ecological cohort: 11 000 men aged 40–59 in 7 countries in Europe and the USA	Link between unsaturated fat and lower risk of death from CHD	Section 2E
Intersalt (1988)	Cross-section: 10 000 men and women	Dietary salt levels linked to blood pressure	Section 2E
UK Women's cohort (1993 …)	Cohort: 35 000 women in UK	Exploring links between diet and cancer	www.leeds.ac.uk/medicine/ ceb/NutEp/ukwcs/ publications.htm

OTHER REVISION LISTS

UK SCREENING: UK NATIONAL SCREENING COMMITTEE SUMMARY

The **condition:**	• An **important health problem** • Epidemiology and natural history understood • Cost-effective primary prevention interventions implemented first
The **test**	• **Simple, safe, precise and validated** • Distribution of test values in the target population should be known and a cut-off level defined and agreed • **Acceptable** to the population • Policy on diagnostics for individuals with a positive test and choices available • Criteria to select mutations to be covered by screening
The **treatment**	• **Effective treatment** exists for patients identified through early detection • Early treatment has better outcomes than late treatment • Agreed criteria for which individuals to be offered treatment and what should be offered • Providers prepared to manage patients before programme starts
The **programme:**	• **Reduces mortality or morbidity** • Where screening aimed solely at providing 'informed choice', test accurately measures risk, provides valuable information readily understood by the individual being screened • Complete programme clinically, socially and ethically acceptable to health professionals and the public • **Benefit outweighs harm** (physical and psychological) • Value for money • **Managing and monitoring** arrangements in place • Adequate resources available – staffing and facilities for testing, diagnosis, treatment and programme management • Other options for managing the condition considered • Evidence-based **information for potential participants** explaining screening consequences • Anticipate public pressure for widening eligibility criteria. Decisions about screening parameters should be justifiable to the public.
More information	• Section 2C • www.nsc.nhs.uk/pdfs/criteria.pdf • Wilson and Jungner (1968)

CAUSATION AND ASSOCIATION

Bradford Hill 'viewpoints' for studying causation (Bradford Hill 1965) (see Section 1A)

Strength	High relative risk or odds
Consistency	Similar results from several studies in different populations
Specificity	Single cause produces a single effect
Temporality	Cause must precede effect
Biological gradient	Dose–response curve
Plausibility	Biologically acceptable or relevant reason for the cause to produce effect
Coherence	Does not conflict with current knowledge
Experimental	Introduction or removal of putative cause leads to change in effect
Analogy	Consistent with previous experience in similar situations

Further Reading

CHAPTER 1A

BOOKS

Donaldson LJ, Donaldson RJ (2003) *Essential Public Health*. Oxford: Petroc Press. Contains readable summaries on a wide range of subjects in the Part A syllabus.

Hennekens CH, Buring JE (1987) *Epidemiology in Medicine*. Philadelphia: Lippincott, Williams & Wilkins. Comprehensive description of epidemiology concepts, study designs and analysis.

Kirkwood BR, Sterne J (2003) *Essential Medical Statistics* (2nd ed). Oxford: Blackwell Science. In-depth description of statistics and their use in epidemiological studies including key elements required for Part A.

Petrie A, Sabin C (2005) *Medical Statistics at a Glance* (2nd ed). Oxford: Blackwell Publishing. Short summary covering all key elements of statistics required for Part A with examples from epidemiological studies.

Swinscow TDV, Campbell MJ (2002) *Statistics at Square One* (10th ed). London: BMJ Books. Guide to statistics, containing several exercises to practice calculations and embed understanding of statistical concepts. 9th edition available online at: www.bmj.com/statsbk.

ARTICLES

Northridge ME (1995) Public health methods – attributable risk as a link between causality and public health action. *American Journal of Public Health*. **85**:1202–4. Provides examples of calculations of attributable risk percent and population attributable risk percent, and illustrates how it can be used to inform public health priorities.

WEBSITES

Centre for Evidence-Based Medicine – www.cebm.net. Promotes evidence-based health-care and provides support and resources to anyone who wants to make use of them.

Cochrane Collaboration – www.cochrane.org. Produces and disseminates systematic reviews of health-care interventions and promotes the search for evidence in the form of clinical trials and other studies of interventions.

GreyNet – www.greynet.org. Grey literature network service which facilitates dialogue, research and communication between people and organisations in the field of grey literature.

National Institute for Health and Clinical Excellence – www.nice.org.uk. Organisation responsible for providing national guidance on the promotion of good health and the prevention and treatment of ill health (merged with the Health Development Agency in 2005).

National Research Ethics Service (NRES) – www.nres.npsa.nhs.uk. NRES is part of the National Patient Safety Agency. It coordinates certain parts of the research ethics approval process and provides research ethics approval process, and guidance, training and support to local committees and applicants.

Office of National Statistics – www.statistics.gov.uk. Compiles and publishes statistics on Britain's population and society at national and local level.

Wikipedia – en.wikipedia.org/wiki/Main_Page. Reference resource that anyone can contribute to, containing within its 1000 000 articles, information on many of the subjects in the Part A syllabus.

CHAPTER 1B

BOOKS

Campbell MJ (2001) *Statistics at Square Two: Understanding modern statistical applications in medicine*. London: BMJ Books.

Kirkwood BR, Sterne J (2003) *Essential Medical Statistics* (2nd ed). Oxford: Blackwell Science. In-depth description of statistics and their use in epidemiological studies including key elements required for Part A.

Petrie A, Sabin C (2005) *Medical Statistics at a Glance* (2nd ed). Oxford: Blackwell Publishing. Short summary covering all key elements of statistics required for Part A with examples from epidemiological studies.

Pereira-Maxwell F (1998) *A-Z of Medical Statistics: A companion for critical appraisal*. London: Hodder Arnold.

Rowntree D (1991) *Statistics without Tears: An introduction for non-mathematicians*. London: Penguin Books Ltd. Simple introduction to statistics, which assumes no previous knowledge of the subject.

ARTICLES

Altman DG, Bland JM (1996) Education and debate. Statistics notes: Comparing several groups using analysis of variance. *BMJ* **312**:1472–3.

Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**:629–34.

Harrison WN, Mohammed MA, Wall MK, Marshall TP (2004) Analysis of inadequate cervical smears using Shewhart control charts. *BMC Public Health* **4**:25.

Spiegelhalter DJ (2005) Handling over-dispersion of performance indicators. *Quality and Safety in Health Care* **14**:347–51. Discusses funnel plots in the context of different statistical methods for dealing with variations in available data.

Sterne JAC, Egger M, Davey Smith G (2001) Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* **323**:101–5. Describes statistical and graphical methods for detecting and correcting for bias.

Tekkis PP, McCulloch P, Steger AC et al (2003) Mortality control charts for comparing performance of surgical units: validation study using hospital mortality data**.** *BMJ* **326**:786–8.

WEBSITES

Bandolier – www.jr2.ox.ac.uk/bandolier. Monthly journal and resources on evidence-based health care. The *Learning Zone* section is particularly useful for Part A.

Children's Mercy Hospitals and Clinics – www.childrens-mercy.org/stats/definitions/conditional.htm. Explanation of conditional probability.

Public Health electronic Library – www.healthknowledge.org.uk. The health knowledge section contains succinct summaries on key subjects tested at Part A.

CHAPTER 1C

BOOKS

Fulop N, Allen P, Clarke A, Black N (eds) (2001) *Studying the Organisation and Delivery of Health Services: Research methods*. London: Routledge. Health services research from the perspective of several different research disciplines, including action research, organisational psychology and operational research.

McPake B, Kumaranayake L, Normand C (2002) *Health Economics: An international perspective*. London: Routledge. An introduction to health economics in three sections – covering supply and demand concepts and markets, economic evaluation and a comparison of different health-care systems.

Pencheon D, Guest C, Melzer D, Muir Gray JA (eds) (2006) *The Oxford Handbook of Public Health Practice* (2nd ed) (Oxford Handbooks Series). Oxford: Oxford University Press. Covers many essential public health topics including needs assessment, advocacy and management issues with step-by-step descriptions and practical examples.

Wright J (ed) (1998) *Health Needs Assessment in Practice*. London: BMJ Books.

ARTICLES

Bradshaw J (1972) A taxonomy of social need. *New Society* **March**:640–3.

Carr-Hill RA (1992) The measurement of patient satisfaction. *Journal of Public Health Medicine* **14**:236–49.

Lock K (2000) Health impact assessment. *BMJ* **320**:1395–8.

McColl A, Roderick P, Gabbay J, Ferris G (1998) What do health authorities think of population based outcome indicators? *Quality in Health Care* **7**:90–7.

National Institute for Clinical Excellence (2002) *Principles for Best Practice in Clinical Audit*. Oxford: Radcliffe Medical Press. Also available online at: www.nice.org.uk/page.aspx?o=29058. A comprehensive guide to practising clinical audit in the health service.

WEBSITES

Department of Health – www.dh.gov.uk. Government department in England responsible for health-care policy and the strategic direction of health care.

Healthcare Commission – www.healthcarecommission.org.uk. Responsible for reviewing the performance and quality of NHS and private health care.

National Institute for Health and Clinical Excellence – www.nice.org.uk. Responsible for providing national guidance on the promotion of good health and the prevention and treatment of ill health.

CHAPTER 1D

BOOKS

Green J, Thorogood N (2004) *Qualitative Methods for Health Research*. London: Sage Publications Ltd. A clearly written, practical guide that contains case studies illustrating how to carry out qualitative research and some of the practical and theoretical issues that need to be addressed.

CHAPTER 2A

ARTICLES

Ben-Shlomo Y, Kuh D (2002) A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology* **31**:285–93. This editorial accompanies a series of articles with a 'life course' theme. It outlines some of the major key conceptual issues around life course epidemiology.

CHAPTER 2B

WEBSITES

Association of Public Health Observatories – www.apho.org.uk/apho. For statistics and reports relating to health in regional areas.

Clinical and Health Outcomes Knowledge Base – www.nchod.nhs.uk. For the compendium of 500 indicators enabling health comparisons at regional and organisational levels.

Department of Health – www.dh.gov.uk. For national strategies, e.g. National Service Frameworks, statistics or guidance relevant to England.

National Institute for Health and Clinical Excellence – www.nice.org.uk. For nationally recommended cost-effective interventions (technology assessments) and guidelines on the management of specific conditions.

World Health Organization – www.who.int. For international guidelines and statistics. In particular, the global burden of disease estimates are useful for key diseases nationally and internationally, available at: www.who.int/healthinfo/statistics/bodgbddeathdalyestimates.xls.

Major charity websites, e.g.:

- Mental health: MIND – www.mind.org.uk; Rethink – www.rethink.org; Alzheimer's Society – www.alzheimers.org.uk
- Cardiovascular disease: British Heart Foundation – www.bhf.org.uk; Stroke Association – www.stroke.org.uk
- Long-term conditions: Diabetes UK – www.diabetes.org.uk; Asthma UK – www.asthma.org.uk; Multiple Sclerosis Society – www.mssociety.org.uk.

CHAPTER 2C

ARTICLES

Deeks JJ, Altman DG (2004) Diagnostic tests 4: likelihood ratios. *BMJ* **329**:168–9.

Department of Health (2002) *Screening/Case Finding: National Service Frameworks: A practical aid to implementation in primary care*. London: Department of Health. Available online at: www.dh.gov.uk/assetRoot/04/05/08/68/04050868.pdf.

Gaeta T (2005) *Screening and Diagnostic Tests*. eMedicine, WebMD. Available online at: www.emedicine.com/emerg/topic779.htm.

Holtzman NA, Shapiro D (1998) Genetic testing and public policy. *BMJ* **316**:852–6.

Karimi A, Kadivar MR, Fararoee M, Alborzi A *(*2000) Active case-finding of communicable diseases in the south of the Islamic Republic of Iran. *Eastern Mediterranean Health Journal* **6**:487–90.

Marks D, Wonderling D, Thorogood M et al (2000) Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis. *Health Technology Assessment* **4**:1–123.

Marteau TM, Dormandy E, Michie S (2001) A measure of informed choice. *Health Expectations* **4**:99–108.

WEBSITES

Jarrett J (2004) Health economics of screening – www.cmgp.org.uk/resources/ppt/jj_guys_talk.pps. Cambridge Genetics Knowledge Park and University of East Anglia Presentation at Guy's Hospital, London.

National Library for Health (NLH) – www.library.nhs.uk. NLH organises and provides access to the best available evidence on screening.

Public Health electronic Library – www.library.nhs.uk/publichealth.

UK National Screening Committee (NSC) – www.nsc.nhs.uk. The NSC advises Ministers, the devolved National Assemblies and the Scottish Parliament on all aspects of screening policy. The website has information on the evidence base for recommended programmes and those in discussion. In particular, UK National Screening Committee's Policy Positions – www.nsc.nhs.uk/pdfs/policy_position_chart_july06%5B1%5D.pdf.

Genome programmes of the US Department of Energy Office of Science - www.doegenomes.org. In particular, Human Genome Project Information: Ethical, Legal, and Social Issues – www.ornl.gov/sci/techresources/Human_Genome/elsi/elsi.shtml.

CHAPTER 2D

BOOKS

Zimmern R (2001) Genetics in disease prevention. In: Pencheon D, Guest C, Melzer D, Muir Gray JA (eds) *Oxford Handbook of Public Health Practice*. Oxford: Oxford University Press, pp 544–9. Concise account of public health genetics, linking concepts to issues that public health practitioners may encounter in their work.

ARTICLES

Kaslow RA, Moser SA (2000) Role of microbiology in epidemiology: Before and beyond 2000. *Epidemiologic Reviews* **22:**131–5. Commentary on the ways in which molecular biological techniques have been used in epidemiology.

Kirk M (2005) The role of genetic factors in maintaining health. *Nursing Standard* **20**:50–4. Introduction to genetics in public health with an emphasis on use in clinical practice.

WEBSITES

Centers for Disease Control and Prevention, Office for Genomics and Disease Prevention – www.cdc.gov/genomics. US-based website providing information about human genetic developments and potential uses in improving health and preventing disease at the population level.

Foundation for Genomics and Population Health – www.phgfoundation.org/. Contains primers on basic genetics, including genetic epidemiology, principles, techniques and applications to health and disease suitable for health-care professionals.

CHAPTER 2E

ARTICLES

Department of Health (2004) *Choosing Health: Making healthy choices easier*. London: Department of Health. Available online at: www.dh.gov.uk. Public Health White Paper for England, including a range of dietary interventions to reduce the prevalence of obesity.

WEBSITES

British Nutrition Foundation – www.nutrition.org.uk. Charity providing evidence-based nutritional knowledge and advice.

Food Standards Agency – www.foodstandards.gov.uk. Agency for the English government and the devolved administrations in the UK providing advice, information, monitoring and enforcement about food safety. Its website contains nutritional advice, food safety information and results of nutritional surveys.

WHO Global Strategy on Diet, Physical Activity and Health – www.who.int/dietphysicalactivity/en. These webpages provide information on risk factors, WHO strategy and effective interventions related to diet.

WHO Child Growth Standards – www.who.int/childgrowth/standards/en/. These webpages provide information on recommended child height, weight, body mass index, for children up to age 5 and the background to their development.

CHAPTER 2F

BOOKS

Donaldson L, Donaldson RJ (2003) *Essential Public Health* (2nd ed). Oxford: Petroc Press.

Yassi A, Kjellstrom T, de Kok T, Guidotti T (2001) *Basic Environmental Health*. Oxford: Oxford University Press.

WEBSITES

BBC Weather Centre: Climate Change – www.bbc.co.uk/climate/evidence. Basic guide to climate change, covering the evidence base, impact and policies to reduce global warming.

Communities and Local Government – www.communities.gov.uk. Government department in England created in May 2006 that replaces many of the functions of the revised Office of the Deputy Prime Minister.

Department for Environment, Food and Rural Affairs – www.defra.gov.uk. Government department developing policy around the interests of farmers and the countryside, the environment and the rural economy.

Department for Transport – www.dft.gov.uk. Government department developing policy around transport, which includes environmentally friendly and sustainable transport options.

Drinking Water Inspectorate – www.dwi.gov.uk. Agency responsible for assessing and enforcing the quality of drinking water in England and Wales.

Environment Agency – www.environment-agency.gov.uk. Government agency in England and Wales responsible for inspecting and regulating businesses, providing information and advice on environmental issues and taking action, after an environmental incident such as a flood or pollution.

Food Standards Agency – www.food.gov.uk. Government department that provides public information on food safety, circulates food alerts and food legislation, and undertakes local authority audits and other enforcement activities to assure food safety and quality.

Friends of the Earth – www.foe.co.uk. UK-based charity campaigning for initiatives to address climate change.

Health Protection Agency – www.hpa.org.uk. In particular, for checklists for chemical incident management, see: www.hpa.org.uk/chemicals/checklists/htm.

Sustainable Development – www.sustainable-development.gov.uk. Website for the government's sustainable development unit, based in the Department for Environment, Food and Rural Affairs (defra), provides information and updates on sustainable development issues.

Sustainable Development Commission – www.sd-commission.org.uk. The government's independent advisory body for England on sustainable development.

UK Air Quality Archive – www.airquality.co.uk/archive/index.php. Provides general information on the causes and nature of air pollution and current levels of local air pollution condition across Britain.

UN Economic Commission for Europe – www.unece.org/env/welcome.html. The UN Economic Commission for Europe's environment website covers the formation of international policies, countrywide inspections and standards, and international treaties on the environment.

CHAPTER 2G

BOOKS

Chin J (ed) (2000) *Control of Communicable Diseases Manual* (17th ed). American Public Health Association. A relatively concise handbook for the management of communicable disease and the definitive text in the USA.

Department of Health (1996) *Immunisation Against Infectious Disease 1996 – 'The Green Book'*. London: Department of Health. Revised versions are available online at: www.dh.gov.uk. Contains the most recent national advice on immunisation with descriptions of the epidemiology and clinical features of immunisable diseases. Several areas, e.g. MMR, have been updated since 1996.

Donaldson L, Donaldson RJ (2003) *Essential Public Health* (2nd ed). Oxford: Petroc Press.

Hawker J, Begg N, Blair I et al (2005) *Communicable Disease Control Handbook* (2nd ed). Oxford: Blackwell Publishing. UK-based publication describing the clinical features of communicable diseases and providing guidance for management.

WEBSITES

Department of Health – www.dh.gov.uk. The Department of Health sets health policy for England. Policies on issues relevant to infection control on the website include:

- *The Health Bill* (2006) – includes Hygiene Code of Practice for the Prevention and Control of Healthcare Associated Infections
- *Saving Lives: A delivery programme to reduce healthcare associated infections including MRSA* (2005)
- *Matrons' Charter: An action plan to cleaner hospitals* (2004)
- *Getting Ahead of the Curve: A strategy for combating infectious diseases (including other aspects of health protection)* (2002).

Health Protection Agency – www.hpa.org.uk. The statutory body in England with responsibility for protecting the public from communicable diseases and chemical, poisonous or radioactive hazards, and preparing for new and emerging threats to health, such as bioterrorism or new disease strains. The HPA's website contains succinct descriptions of many communicable diseases. Its online journal, *CDR Weekly*, provides updates on the incidence of major communicable diseases across England.

National Resource for Infection Control – www.nric.org.uk. Maintained by City University and funded by the Department of Health, the site provides resources and links to up-to-date, UK-based information on infection control for health care staff.

Public Health electronic Library – www.healthknowledge.org.uk. The health knowledge communicable disease section includes summaries on outbreaks, surveillance and immunisation.

World Health Organization – www.who.int. The WHO website has several factsheets on infectious diseases, provides descriptions of WHO activities, and downloadable reports – see Section 2H.

BOOKS

Naidoo J, Wills J (2000) *Health Promotion: Foundations for practice*. London: Baillière Tindall. Comprehensive textbook covering health promotion, from theory to practice with examples from practice.

Nutbeam D, Harris E (1999) *Theory in a Nutshell: A guide to health promotion theory*. Maidenhead: McGraw-Hill Education. Very short, digestible summaries of major health promotion theories.

Downie RS, Tannahill C, Tannahill A (1996) *Health Promotion: Models and values* (2nd ed). Oxford: Oxford University Press. Textbook describes the Tannahill health promotion framework, and covers the policy and practice of health promotion from various different clinical and social perspectives.

O'Sullivan GA, Yonkler JA, Morgan W, Merritt AP (2003) *A Field Guide to Designing A Health Communication Strategy*. Baltimore: Johns Hopkins Bloomberg School of Public Health/Center for Communication Programs. Available at: www.maqweb.org/iudtoolkit/marketing_comm/afieldguide.shtml. A series of steps and tools for effective health communication, including health behaviour theories to audience segmentation and communication strategy.

WEBSITES

World Health Organization – www.who.int. A subsidiary of the United Nations, the most important source of guidance and information regarding international health and health promotion initiatives.

Action on Smoking and Health (ASH) – www.newash.org.uk. A campaigning organisation working to eliminate the harm from tobacco; its website contains information about the effects of tobacco and useful summaries of current tobacco policy both in the UK and globally.

CHAPTER 2I

ARTICLES

Cancer Research UK *SunSmart Campaign Strategy*. Available online at: www.cancerresearchuk.org/healthyliving/sunsmart/forprofessionals/campaignstrategy. A description of the origins, research, range of activities and evaluation of a national social marketing campaign in England to reduce harm from the sun.

Wanless D (2004) *Securing Good Health for the Population*. London: HMSO. Available online at: www.hm-treasury.gov.uk/consultations_and_legislation/wanless/consult_wanless04_final.cfm. Derek Wanless's second report into health, commissioned by the Treasury, provides a summary of key public health policies in England in the twentieth and early twenty-first century, and summarises the evidence base around smoking, falls, obesity/physical activity and salt.

CHAPTER 3A

BOOKS

Pencheon D, Guest C, Melzer D, Muir Gray JA (eds) (2006) *The Oxford Handbook of Public Health Practice* (2nd ed) (Oxford Handbooks Series). Oxford: Oxford University Press. Includes chapters on information collection and use. In particular, see Chapters 1.1 Information, 1.2 Acute health trends: surveillance, 1.3 Longer term health trends: registers and 5.2 Evaluating health care using routine data.

ARTICLES

Butler RN (1997) Population aging and health. *BMJ* **315**:1082–4. This article covers trends in ageing and the impact on the population, and appears in a themed issue of the *BMJ* devoted to ageing.

McMichael AJ (2002) Population, environment, disease, and survival: Past patterns, uncertain futures. *Lancet* **359**:1145.

WEBSITES

Communities and Local Government (CLG) – www.communities.gov.uk. Formerly known as the Department of the Deputy Prime Minister, CLG provides key 'non-health' information and policies relevant to public health, including local government, housing/homelessness and transport.

Government Actuary's Department – www.gad.gov.uk. Provides demographic information, life-tables and population projections for the UK with descriptions of the methodologies used.

Office for National Statistics – www.statistics.gov.uk. The primary source for information on Britain's population, lifestyles, economy and society at national and local levels. Statistical tables and commentaries on the statistics are available.

CHAPTER 3B

The fast pace of change in England's NHS means that few resources stay current for long and this is particularly true for information sources. Most of the further reading suggested therefore comes from websites, rather than textbooks. However, the degree of organisational change may mean that some of these sites become obsolete in the near future.

BOOKS

Donaldson LJ, Donaldson RJ (2003) *Essential Public Health* (2nd ed). Oxford: Petroc Press. Key public health text with useful and relatively current section on health and related information.

WEBSITES

NHS Connecting for Health – www.connectingforhealth.nhs.uk. The organisation responsible for the NHS IT modernisation programme. The website contains information on many aspects of information use, communication and storage within the NHS.

The Information Centre for Health and Social Care – www.ic.nhs.uk. It commissions and produces a range of health-care information relevant to public health, including:

- Primary care: Quality and Outcomes Framework (QOF) data relating to general practices' performance on key clinical and organisational domains

- Surveys, e.g. Health Survey for England data
- Service statistics, e.g. maternity, community care.

Medicines and Healthcare products Regulatory Agency – www.mhra.gov.uk. The MHRA is the government agency for England responsible for ensuring that medicines and medical devices work and are acceptably safe.

National Prescribing Centre – www.npc.co.uk. Formed by the Department of Health to promote high-quality, cost-effective prescribing to relevant professionals and senior managers working in the NHS. Contains information on medicines management, non-medical prescribing, education and development.

Office for National Statistics – www.statistics.gov.uk. The primary source for information on Britain's population, lifestyles, economy and society at national and local levels.

Prescription Pricing Authority – www.ppa.nhs.uk/index.htm. This is part of the NHS Business Services Authority and processes all NHS prescriptions dispensed in England. The website contains information on prescribing volumes, trends and costs in the NHS.

United Kingdom Association of Cancer Registries – www.ukacr.org. Organisations involved in collecting and coding cancer information in the UK and Ireland.

World Health Organization, *International Classification of Diseases* (10th ed) – www.who.int/classifications/icd/en. This website provides a list of ICD-10 codes and subcodes.

CHAPTER 3C

BOOKS

Berg M (ed) (2004). *Health Information Management: Integrating information and communication technology in health care work*. London: Routledge. Recent and international description of health information management theories and implementation.

Deutsch T, Carson E, Ludwig E (1994) *Dealing with Medical Knowledge: Computers in clinical decision making*. New York: Plenum Press.

Donaldson L, Donaldson RJ (2003) *Essential Public Health* (2nd ed). Oxford: Petroc Press. Provides a summary of sources of health and health-care information available in England and Wales.

ARTICLES

Musgrave S (2004) Public health information systems – tools to widen their accessibility and impact. *Healthcare Computing* 237–44. Available online at: www.health-informatics.org/hc2004/P30_Musgrave.pdf.

Marshall T, Rouse A (2002) Resource implications and health benefits of primary prevention strategies for cardiovascular disease in people aged 30 to 74: mathematical modelling study. *BMJ* **325**:197–9. An important example of a cost-effectiveness exercise (within the National Service Framework for Coronary Heart Disease).

Stevens A, Gillam S (1998) Needs assessment: from theory to practice. *BMJ* **316**:1448–52.

WEBSITES

Association of Public Health Observatories – www.apho.org.uk/apho. The umbrella body of the regional public health observatories (PHOs) in England. The Association includes profiles on health themes such as determinants of health and links to the regional PHOs, which collate and analyse local data to provide local health and health service profiles for their areas.

Centre for Health Economics, Health Policy Team, University of York – www.york.ac.uk/inst/che/research/hpolicy.htm Information on applied and methodological economics research relevant to health care.

Department of Health – www.dh.gov.uk. The website contains some information on the English NHS performance.

National Institute of Economic and Social Research – www.niesr.ac.uk. In particular, O'Mahony M (2006) *Outputs, Inputs and Productivity in the NHS*.

NHS Health Informatics Community – www.connectingforhealth.nhs.uk. Website relating to NHS informatics, information and developments.

NHS Connecting for Health – www.connectingforhealth.nhs.uk. News and information on the National Programme for IT, which aims to introduce modern computer systems throughout the NHS.

Public Health electronic Library – www.healthknowledge.org.uk.

Virtual classroom of Health Informatics – www.robin-beaumont.co.uk/virtualclassroom/contents.htm.

CHAPTER 4A

BOOKS

Halpern D (2005) *Social Capital*. Cambridge: Polity Press. Textbook that covers virtually all of the Part A syllabus relating to medical sociology.

Scambler G (2003) *Sociology as Applied to Medicine* (5th ed). Philadelphia: WB Saunders.

ARTICLES

Wilkinson R, Marmot M (2003) *Social Determinants of Health: The solid facts* (2nd ed). Copenhagen: WHO. Available at: www.euro.who.int/document/e81384.pdf

Brimlow D, Cook JS, Seaton R (2003) *Stigma and HIV/AIDS: A review of the literature*. US Department of Health and Human Services Health Resources and Services Administration. Available online at: hab.hrsa.gov/publications/stigma/introduction.htm.

Gray AJ (2002) Stigma in psychiatry. *Journal of the Royal Society of Medicine* **95**:72–6.

Macintyre M, Telban B, Mitchell WE et al (2004) Book review forum on Gilbert Lewis's *A Failure of Treatment*. *Journal of Ritual Studies* **18**:121–51. Available online at: www.pitt.edu/~strather/Gilbert%20Lewis%20Review%20Forum.pdf.

WEBSITES

Equality and Human Rights Commission – www.equalityhumanrights.com/pages/eocdrccre.aspx. Contains educational materials that illustrate the social model of disability and information relating to the three merged organisations, the Disability Rights Commission, Commission for Racial Equality and the Equal Opportunities Commission.

Healthcare Commission – www.healthcarecommission.org.uk. Commission for Health Improvement. In particular, see *Unpacking the Patients' Perspective: Variations in NHS patient experience in England* (2004), available at: www.healthcarecommission.org.uk/_db/_documents/04003496.pdf.

National Grid for Learning Cymru, Glossary of Sociological Terms – www.ngfl-cymru.org.uk/vtc/ngfl/sociology/detailed_glossary.htm. Authored by David Bown, assisted by Janis Griffiths.

Public Health electronic Library – www.healthknowledge.org.uk. Contains short summaries on medical sociology applied to Part 1.

CHAPTER 4B

BOOKS

Gabe J, Bury M, Elston MA (2004) *Key Concepts in Medical Sociology*. London: Sage Publications. Lay language overview of medical sociology.

Scambler G (2004) *Sociology as Applied to Medicine* (5th ed). London: WB Saunders. Standard medical sociology text, covering the Part A syllabus.

CHAPTER 4C

BOOKS

Donaldson C, Gerard K, Milton C et al (2004) *Economics of Healthcare Financing: The visible hand*. Basingstoke: Palgrave Macmillan. Readable explanation of modern health economic policy concepts and dilemmas.

Hogwood B, Gunn LA (1985). *Policy Analysis for the Real World*. Oxford: Oxford University Press, 52–3.

McPake B, Kumaranayake L, Normand C (2002) *Health Economics: An international perspective*. Routledge. An introduction to health economics in three sections – covering supply and demand concepts and markets, economic evaluation, and a comparison of different health-care systems.

Pencheon D, Guest C, Melzer D, Muir Gray JA (eds) (2006) *Oxford Handbook of Public Health Practice* (Oxford Handbooks Series). Second Edition Oxford: Oxford University Press. Covers many essential public health topics, including policy formation and priority setting descriptions, and gives practical examples.

ARTICLES

Department of Health (2003) *Code of Practice on Openness in the NHS*. Leeds: Department of Health. Available at: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4050490.

Dollar D (2001) Is globalization good for your health? *Bulletin of the World Health Organization* **79**:827–33.

Leon DA, Walt G, Gilson L (2001) International perspectives on health inequalities and policy. *BMJ* **322**:591–4.

Maynard A (2001) Economics-based medicine: an evolving paradigm. *Journal of the Royal College of Physicians of Edinburgh* **31**(Suppl 9):16–7.

NICE (2006) Rapid review of the economic evidence of physical activity interventions. Available at: www.nice.org.uk/download.aspx?o=528518.

Stevens A, Gabbay J (1991) Needs assessment needs assessment. *Health Trends* **23**:20–3.

WEBSITE

Wikipedia, Procedural Justice – en.wikipedia.org/wiki/Procedural_justice.

CHAPTER 4D

BOOKS

Farmer R, Miller D, Lawrenson R (2004) *Lecture Notes on Epidemiology and Public Health Medicine* (5th ed). Oxford: Blackwell Publishing.

Jefferson T, Demicheli V, Mugford M (2000) *Elementary Economic Evaluation in Health Care* (2nd ed). London: BMJ Books.

ARTICLES

Cohen D (1994) Marginal analysis in practice: an alternative to needs assessment for contracting health care. *BMJ* **309**:781–4.

National Institute for Health and Clinical Excellence (2006) *Methods for Development of NICE Public Health Guidance*. Available at: www.nice.org.uk/download.aspx?o=299970.

Office of Health Economics (1999) *The Economics of Health Care Scarcity* (3rd ed). Available at: www.oheschools.org/ohech1pg3.html.

Palmer S, Byford S, Raftery J (1999) Economics notes: Types of economic evaluation. *BMJ* **318**:1349.

Palmer S, Raftery J (1999) Economics notes: Opportunity cost**.** *BMJ* **318**:1551–2.

Palmer S, Torgerson DJ (1999) Economic notes: Definitions of efficiency *BMJ* **318**:1136.

Sassi F, Le Grand J, Archard L (2001) Equity versus efficiency: a dilemma for the NHS. *BMJ* **323:**762–3.

WHO (2001) *About Health Systems Performance*. Available at: www.who.int/health-systems-performance/about.htm#Why%20assess%20health%20system%20performance.

CHAPTER 5A

BOOKS

Iles V, Cranfield S (2005) *Really Managing Healthcare* (2nd ed) Maidenhead: Open University Press/McGraw-Hill.

Pencheon D, Guest C, Melzer D, Muir Gray JA (eds) (2006) *Oxford Handbook of Public Health Practice* (2nd ed) (Oxford Handbooks Series). Oxford: Oxford University Press. Covers many essential public health topics including management.

ARTICLES

Edwards N (2003) Doctors and managers: poor relationships may be damaging patients – what can be done? *Quality and Safety in Health Care* **12**:21–4.

Finch J (2000) Inter-professional education and team-working: a view from the education providers. *BMJ* **321**:1138–40.

Garelick A, Fagin F (2005) The doctor–manager relationship. *Advances in Psychiatric Treatment* **11**:241–50.

Barr H, Goosey D (2002) *Interprofessional Education: Selected case studies*. London: Department of Health. Available at: www.dh.gov.uk/assetRoot/04/03/45/41/04034541.pdf.

WEBSITES

Public Health electronic Library health knowledge resources – www.healthknowledge.org.uk. The website contains summaries of a range of subjects relevant to Part A, including organisational management relating to health.

Brown B (2000) *Organisational Communication Theories* – uva.ulb.ac.be/cit_courseware/research/theories3.htm. This is a summary of the information in Littlejohn S (1992) *Theories of Human Communication* (5th ed). California: Wadsworth Publishing.

Evidence-based medicine – www.evidence-based-medicine.co.uk. In particular, Storey N (2001) *What is Clinical Governance?* Volume 1, number 12. Available at: www.evidence-based-medicine.co.uk/ebmfiles/WhatisClinGov.pdf.

The Shipman Inquiry: Independent public enquiry into the issues arising from the case of Harold Fredrick Shipman - www.the-shipman-inquiry.org.uk.

Managing stress – www.cambsmentalhealthinfo.nhs.uk/positive/managing_stress.html. Website provided by Cambs Mental Health Info.

CHAPTER 5B

ARTICLES

Miller TR (1996) Structuring networks for maximum performance under managed care. *Healthcare Financial Management* **50**:37–9.

Pluye P, Potvin L, Pelletier J (2004) Community coalitions and health promotion: Is it that important to develop an inter-organisational network? *Promotion and Education* **11**:17–24.

WEBSITES

Public Health electronic Library health knowledge resources – www.healthknowledge.org.uk. The website contains summaries on a range of subjects relevant to Part A, including organisational management relating to health.

Classics in the history of psychology. *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*. Sherif M, Harvey O, White BJ et al (1954/1961) – psychclassics.yorku.ca/Sherif. An internet resource developed by Christopher D Green, York University, Toronto, Ontario.

CHAPTER 5C

BOOKS

Iles V, Sutherland K (2001) *Organisational Change: A review for health care managers, professionals and researchers*. London: NCCDSO, London School of Hygiene and Tropical Medicine. The guide provides readable, concise summaries of change management models and the experience of their application to health care.

Mullins LJ (2004) *Management and Organisational Behaviour* (7th ed). Harlow: Financial Times/Prentice-Hall. A comprehensive management textbook covering the Part A syllabus and much more in considerable depth. Examples come from a range of organisations, not just health care.

Silbiger S (1999) *The 10-Day MBA: A step-by-step guide to mastering the skills taught in top business schools*. London: Piatkus Books.

CHAPTER 5D

BOOKS

Donaldson L, Donaldson RJ (2003) *Essential Public Health* (2nd ed). Oxford: Petroc Press.

Iles V, Sutherland K (2001) *Organisational Change: A review for health care managers, professionals and researchers*. London: NCCSDO. A review of management literature as it relates to health care in the NHS.

McDougall A, Duckett P, Manku M, Robertson J (2003) *International Health Comparisons: A compendium of published information on healthcare systems, the provision of healthcare and health achievement in 10 countries*. London: National Audit Office. Available online at: www.nao.org.uk/publications/Int_Health_Comp.pdf. A description of health systems in ten industrialised nations.

WEBSITES

Appraisal of Guidelines Research and Evaluation (AGREE) – www.agreecollaboration.org/intro.

Bandolier – www.jr2.ox.ac.uk/bandolier. Monthly journal and resources on evidence-based health-care.

Scottish Intercollegiate Guidelines Network (SIGN) – www.sign.ac.uk.

CHAPTER 6A

BOOKS

Petrie A, Sabin C (2005) *Medical Statistics at a Glance* (2nd ed). Oxford: Blackwell Publishing. Short summary covering all key elements of statistics required for Part A with examples from epidemiological studies.

Silman AJ (2002) *Epidemiological Studies: A practical guide* (2nd ed). Cambridge: Cambridge University Press. Contains a useful section on the design and application of epidemiological studies, and a diagram that illustrates differences between observational studies.

CHAPTER 6B

ARTICLES

Stanford X (2002) *Organizational Mapping: Knowing the pitfalls*. Alberta: Stanford Solutions Inc. No. 59. Available at: www.skyrme.com/updates/u59_f1.htm.

WEBSITES

Betty C Jung's website – www.bettycjung.net/Graphing.htm. Charting and Graphing Data Presentation.

Edward Tufte's website – www.edwardtufte.com/tufte/newet. Examples of clear presentations of data and of misleading or confusing presentations of the same data.

Indiana University Southeast Basic Business Statistics, Chapter 2: Presenting Data in Tables and Charts (8th ed) – homepages.ius.edu/SRAUSC01/Chap02.ppt#286,29.

New York State, Department of Health: Tips for presenting data – www.health.state.ny.us/statistics/chac/tips.htm.

Presenting Data: Tabular and graphic display of social indicators – lilt.ilstu.edu/gmklass/pos138/datadisplay/sections/goodcharts.htm. Website created by Gary Klass, Illinois State University (2002). Good do's and don'ts.

See also Chapters 6A.3 and 1B.8.

CHAPTER 6C

BOOKS

The Economist (2005) *Style Guide: The bestselling guide to English usage*. London: Profile Books.

Jessop E (2001) Writing to effect change. In: Pencheon D, Guest C, Melzer D, Muir Gray JA (eds) *Oxford Handbook of Public Health Practice*. Oxford: Oxford University Press, 428–35.

WEBSITES

How to write medical information in plain English – www.plainenglish.co.uk/medicalguide.pdf. Useful tips for clearer writing and a list of medical terms explained in non-technical language.

References

4S Study (1994) Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *The Lancet* **344**:1383–9.

Acheson D (1998) *Independent Inquiry into Inequalities in Health Report*. London: TSO. Available online at: www.archive.official-documents.co.uk/document/doh/ih/ih.htm.

Administration for Children and Families (2006) Head Start General Information website – www.acf.hhs.gov/programs/hsb/.

Amabile TM (1998) How to kill creativity. *Harvard Business Review* **76**:76–87.

Anderson G, Sotir Hussey P (2001) Influencing government policy. In: Pencheon D, Guest C, Melzer D, Muir Gray JA (eds) *Oxford Handbook of Public Health*. Oxford University Press, Chapter 3.3.

Anon (2006) Preventing coronary heart disease: Does Rose's population prevention axiom still apply in the 21st century? (editorial) *BMJ* **332**:617–8.

Armstrong BG (1998) Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occupational and Environmental Medicine* **55**:651–6.

Armstrong D (2002) Clinical autonomy, individual and collective: the problem of changing doctors' behaviour. *Social Science and Medicine* **55**:1771–7.

Avent ND, Chitty LS (2006) Non-invasive diagnosis of foetal sex: utilisation of free foetal DNA in maternal plasma and ultrasound. *Prenatal Diagnosis* **26**:598–603.

Baker R, Robertson N, Farooqi A (1995) Audit in general practice: factors influencing participation. *BMJ* **311**:31–4.

Barker D (2003) The midwife, the coincidence, and the hypothesis. *BMJ* **327**:1428–30. David Barker's account of the circumstances that led to the recording of a unique set of birth records at the beginning of the twentieth century, which became known as the Hertfordshire Cohort. It was access to this dataset that led to the development of the Barker hypothesis.

Barringer BR, Harrison JS (2000) Walking a tightrope: Creating value through interorganizational relationships. *Journal of Management* **26**:367–403.

Bandura A (1977). *Social Learning Theory*. New York: General Learning Press.

Barsky AJ, Borus JF (1995) Somatization and medicalization in the era of managed care. *JAMA* **274**: 1931–4.

Beattie (1991) Knowledge and control in health promotion: a test case for social policy and social theory. In: Gabe J, Calnan M, Bury M (eds) *The Sociology of the Health Service*. London: Routledge.

Beauchamp TL, Childress JF (2001) *Principles of Biomedical Ethics* (5th ed). Oxford: Oxford University Press.

Belbin RM (1996) *Team roles at work*. London: Butterworth-Heinemann.

Belsky J, Melhuish E, Barnes J et al (2006) Effects of SureStart local programmes on children and families: early findings from a quasi-experimental, cross-sectional study. *BMJ* **332**:1476.

Benzeval M, Judge K, Whitehead M (1995) *Tackling Inequalities in Health: An agenda for action*. London: King's Fund.

Benzeval M (1995) *Tackling inequalities in health*. London: King's Fund.

Binney G, Williams C (1995) *Leaning into the Future: Changing the way people change organizations*. London: Nicholas Brealey Publishing Ltd.

Black D, Morris JN, Smith C et al (1980) *Inequalities in Health: Report of a research working group* (Black Report). London: Department of Health and Social Security. Available online at: www.sochealth.co.uk/history/black.htm.

Blackburn H (1999) *On the Trail of Heart Attacks in Seven Countries* – www.epi.umn.edu/research/7countries/index.shtm.

Blake R, Mouton J (1964) *The managerial grid: key orientations for achieving production through people*. Houston: Gulf.

Bland M (2000) *An Introduction to Medical Statistics* (3rd ed). Oxford: Oxford University Press, 128.

Bonnefoy XR, Annesi-Maesano I, Moreno LM et al (2004). Review of evidence on housing and health. Fourth Ministerial Conference on Environment and Health, Budapest, 23–25 June 2004 (EUR/04/5046267/BD/1). Available online at: www.euro.who.int/document/H0H/ebackdoc01.pdf.

Bradshaw JR (1972) The taxonomy of social need. In: McLachlan G (ed) *Problems and Progress in Medical Care*. Oxford: Oxford University Press, 69–82.

Brager G, Specht H (1973) quoted in World Health Organization (2002) *Community Participation in Local Health and Sustainable Development: Approaches and techniques*. Available online at: www.euro.who.int/document/e78652.pdf.

Breslow NE, Day NE (1987) *Statistical Methods in Cancer Research*, Volume II: *The Design and Analysis of Cohort Studies*. Lyon: International Agency for Research on Cancer.

Bunton R (2002) Health promotion as social policy. In: Bunton R, Macdonald G (eds) *Health Promotion: Disciplines and diversity* (2nd ed). London: Routledge.

Buse K, Mays N, Walt G (2005) *Making Health Policy*. Maidenhead: Open University Press.

Carstairs' Deprivation Index for Scottish Postcode Sectors, 1991 and 2001 – www.datalib.ed.ac.uk/EUDL/carstairs.html.

Carstairs V, Morris R (1989) Deprivation, mortality and resource allocation. *Community Medicine* **11**:364–72.

Cavanagh S, Chadwick K (2005) *Health Needs Assessment: A practical guide*. London: National Institute for Clinical Excellence. Available online at: www.nice.org.uk/page.aspx?o=513203.

Centre for Evidence-Based Medicine (2007) *Levels of Evidence* – www.cebm.net.

Centers for Disease Control and Prevention (2003) *Radiation Measurement Factsheet* – www.bt.cdc.gov/radiation/pdf/measurement.pdf.

Chadwick E (1842) *The Sanitary Conditions of the Labouring Population*. Self published.

Chapman S (2004) Advocacy for public health: a primer. *Journal of Epidemiology and Community Health* **58**:361–5.

Cochrane AL (1972) *Effectiveness and Efficiency: Random reflections on health services*. London: Nuffield Provincial Hospitals Trust. Reprinted in 1999 for Nuffield Trust by the Royal Society of Medicine Press, London.

Cohen D (1994) Marginal analysis in practice: an alternative to needs assessment for contracting health care. *BMJ* **309**:781–4.

Comptroller and Auditor General (2004) *The Drug Treatment and Testing Order: early lessons*. London: National Audit Office. Available online at: www.nao.org.uk/publications/nao_reports/03-04/0304366.pdf.

Cornwell J (1984) *Hard-earned Lives: Accounts of Health and Illness from East London*. London: Routledge.

Council of Europe (2006) *Social Cohesion Strategy*. Available at: www.coe.int/T/E/social_cohesion.

Covey SR (1989) *The Seven Habits of Highly Effective People*. London: Simon & Schuster.

Dahlgren G, Whitehead M (1991) *Policies and Strategies to Promote Social Equity in Health*. Stockholm: Institute for Futures Studies.

Davis RL, Khoury MJ (2006) A public health approach to genomics and gene-based diagnostic tests. *Pharmacogenomics* **7**:331–7.

Degeling P, Maxwell S, Kennedy J, Coyle B (2003) Medicine, management and modernisation: a 'danse macabre'. *BMJ* **326**:649–52.

Department of Health (1998) *A First Class Service: Quality in the new NHS*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Heath (1999a) *Patient and Public Involvement in the New NHS*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Health (1999b) *Saving Lives: Our healthier nation*. London: The Stationery Office. Available online at: www.archive.official-documents.co.uk/document/cm43/4386/4386.htm.

Department of Health (2000a) *Coronary Heart Disease National Service Framework*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Health (2000b) *The NHS Plan: a plan for investment, a plan for reform*. London: The Stationery Office. Available online at: www.dh.gov.uk.

Department of Health (2002) *Getting ahead of the curve: a strategy for combating infectious diseases (including other aspects of health protection)*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Health (2003) Recognising the potential. In: *'Our Inheritance, Our Future: Realising the potential of genetics in the NHS'*. London: The Stationery Office. Available online at: www.dh.gov.uk.

Department of Health (2004a) *Choosing Health: Making healthy choices easier*. London: The Stationery Office. Available online at: www.dh.gov.uk.

Department of Health (2004b) *A matron's charter: An action plan for cleaner hospitals*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Health (2004c) *National Standards, Local Action: Health and social care standards and planning framework 2005/06–2007/08*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Health (2005a) *Resource Allocation: Weighted Capitation Formula* (5th ed). Leeds: Department of Health. Available online at: www.dh.gov.uk.

Department of Health (2005b) *Saving Lives: a delivery programme to reduce healthcare associated infection including MRSA: Skills for implementation*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Health (2006) *NHS Allocations*. London: Department of Health. Available online at: www.dh.gov.uk.

Department of Health, South Africa (2004) *Ethics in Health Research: Principles, structures and processes. Research ethics guidelines – 2004*. Pretoria: Department of Health. Available online at: www.doh.gov.za/docs/factsheets/guidelines/ethnics/index.html.

Department of Health Expert Group (Chairman, CMO) (2000) *An Organisation with a Memory. Report of an expert group on learning from adverse events in the NHS chaired by the Chief Medical Officer*. London: The Stationery Office.

Department of Health, National Service Framework for Coronary Heart Disease (2000). *The National Service Framework for Coronary Heart Disease*. London: Department of Health. Available online at: www.dh.gov.uk.

Diderichsen F, Hallqvist J (1998) Social inequalities in health: some methodological considerations for the study of social position and social context. In: Arve-Parès B (ed) *Inequality in Health – a Swedish perspective*. Stockholm: Swedish Council for Social Research, 25–39.

Doll R, Hill AB (1954) The mortality of doctors in relation to their smoking habits: a preliminary report. BMJ **228**:1451–5.

Donabedian A (1966) Evaluating the quality of medical care. *Millbank Memorial Fund Quarterly* **44**:166–206.

Donaldson C, Gerard K, Mitton C et al (2004) *Economics of Health Care Financing: The visible hand*. Basingstoke: Palgrave MacMillan.

Dorak MT (2007) *Genetic Epidemiology* – www.dorak.info/epi/genetepi.html.

Dosani S, Cross P (2004) How to pass MFPH. *BMJ Career Focus* **329**:5.

Drucker P (1950) The New Society: The Anatomy of Industrial Order. New York: Harper.

Drucker P (1954) *The Practice of Management*. New York: HarperCollins.

Dunbar RIM (1992) Neocortex size as a constraint on group size in primates. *Journal of Human Evolution* **22**:469–93.

Durkheim E (1897). *Le suicide*. Paris: Alcan.

Egger M, Davey Smith G, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**:629–34.

Elbein SC (2002) Perspective: The search for genes for type 2 diabetes in the post-genome era. *Endocrinology* **143**:2012–8.

Emberson J, Whincup P, Morris R et al (2004) Evaluating the impact of population and high-risk strategies for the primary prevention of cardiovascular disease. *European Heart Journal* **25**:484–91.

Etziou A (1973) Mixed scanning: A third approach to decision-making. In: Faludi A (ed) *Reader in Planning Theory*. Oxford: Perganum Press, 218–20.

Evans R, Stoddart G (1990) Producing health, consuming resources. *Social Science and Medicine* **31**:1347–63.

Ewles L, Simnet I (1995) *Promoting Health: A practical guide* (3rd rev ed). London: Ballière Tindall.

Ewles L, Simnet I (2003) *Promoting Health: A practical guide* (5th ed). London: Ballière Tindall.

Eyers JE (1998) How to do (or not to do) … Searching bibliographic databases effectively. *Health Policy and Planning* **13**:339–42.

Fagan TJ (1975). Nomogram from Bayes's theorem (letter). *New England Journal of Medicine* **293**:257.

Fiander M, Burns T (2000) A Delphi approach to describing service models of community mental health practice. *Psychiatry Services* **51**:656–8.

Field MJ, Lohr KN (eds) (1990) Institute of Medicine Committee to Advise the Public Health Service on Clinical Practice Guidelines. *Clinical Practice Guidelines: Directions for a new program*. Washington DC: National Academy Press.

Fight Crime: Invest in Kids (2006) *Head Start reduces crime and improves achievement*. New York: Fight Crime: Invest in Kids. Available online at: www.fightcrime.org/reports/headstartbrief.pdf.

Fleiss JL (1981) *Statistical Methods for Rates and Proportions* (3rd ed). New York: John Wiley & Sons, 77.

Foundation Trust Network (2005) *New Voices, New Accountabilities: A guide to wider governance in foundation trusts*. London: Foundation Trust Network. Available online at: www.nhsconfed.org/ftn.

Framingham Heart Study (1948) *A Project of the National Heart, Lung and Blood Institute and Boston University* – www.framinghamheartstudy.org/about/index.html.

French JRP, Raven BH (1960) The bases of social power. In: Cartwright D, Zander A (eds) *Group Dynamics*. New York: Harper and Row.

Friedson E (1959) Specialties without roots: the utilization of new services. *Human Organization*, **18**:112–16.

Fuller E (ed) (2005) *Smoking, Drinking and Drug Use Among Young People in England in 2004*. Leeds: Health and Social Care Information Centre. Available online at: www.ic.nhs.uk/pubs.

Garces GT, Thomas D, Currie J (2002) Longer-term effects of HeadStart. *American Economic Review* **92**:999–1012.

The Global Fund to fight AIDS, Tuberculosis and Malaria website –www.theglobalfund.org/en/.

Goffman E (1963) *Stigma: Notes on the management of spoiled identity*. London: Prentice-Hall.

Goulding N (2003) *How to Prepare and Write for a Scientific Journal*. Available online at: www.wolfson.qmul.ac.uk/epm/students/ssm7/ssm7medauth.pdf.

Hadorn DC (1991) Setting health care priorities in Oregon. Cost-effectiveness meets the rule of rescue. *JAMA* **265**:2218–25.

Hale J (2000) What contribution can health economics make to health promotion. *Health Promotion International* **15**:341–8.

Hall R (2007) Psychology World: Between Subjects One-Way ANOVA example – web.umr.edu/~psyworld/anovaexample.htm; richardhhall.org.

Handy CB (1987) *Understanding Organizations* (Penguin Business) (3rd ed). London: Penguin.

Hanneman RL (1996) Intersalt: hypertension rise with age revisited. *BMJ* **312**:1283–4.

Hart JT (1971) The inverse care law. *The Lancet* **1**:405–12. Available online at: www.sochealth.co.uk/history/inversecare.htm.

He FJ, MacGregor GA (2004) Effect of longer-term modest salt reduction on blood pressure. *Cochrane Database of Systematic Reviews* Issue 3: CD004937.

Health Protection Agency Centre for Infections (2005) *Hepatitis C in England: The First Health Protection Agency Annual Report 2005*. London: Health Protection Agency. Available online at: www.hpa.org.uk.

Health Protection Agency (2006) *Tuberculosis Update, March 2006. The National TB Strain Typing Database – An Update. Information for Action*. Available online at: www.hpa.org.uk.

Henkel GL (undated) North Karelia Project Shows the World How to Reduce Heart Disease – www.kantele.com/nwfwebsite/puska_heart.html. New World Finn website.

Herzberg F (1968) 'One more time: how do you motivate employees?' *Harvard Business Review* **46**:53–62.

Hethcote HW (2000) The mathematics of infectious diseases. *SIAM Review* **42**:599–653.

Hill AB (1965) The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* **58**:295–300.

HM Treasury (2007) Public service agreements. London: HSMO. Available online at: www.hm-treasury.gov.uk/pbr_csr/psa/pbr_csr07_psaopportunity.cfm.

Hochbaum GM (1958) *Public participation in medical screening programs: A socio-psychological study*. PHS publication No. 572. Washington, DC: US Government Printing Office.

Hoffman A, Lowe D, Rudd A et al (2004) *National Sentinel Stroke Audit 2004*. London: Royal College of Physicians.

Humphris D, Hean S (2004) Educating the future workforce: building the evidence about interprofessional learning. *Journal of Health Services Research and Policy* **9**:24–27.

Hunt K, Emslie C (2001) Commentary: The prevention paradox in lay epidemiology – Rose revisited. *International Journal of Epidemiology* **30**:442–6.

Hunt SC, Gwinn M, Adams TD (2003) Family history assessment: Strategies for prevention of cardiovascular disease. *American Journal of Preventive Medicine* **24**:136–42.

Iles V (2005) *Really Managing Health Care*. Oxford: Oxford University Press.

Illich I (1975) *Medical Nemesis: The expropriation of health*. London: Marian Boyars.

Intersalt Cooperative Research Group (1988) Intersalt: An international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion. *BMJ* **297**:319–28.

Irish Council for Bioethics. Website (2004) – www.bioethics.ie.

Jamison J, O'Reilly D, Carr-Hill, R, Stevenson M (2004) Equity of inpatient services by the two communities in Northern Ireland. *Journal of Health Services Research and Policy* **9**:34–8.

Jarman B (1983) Identification of underprivileged areas. *BMJ* **286**:1705–8.

Jefferson T, Demicheli V, Mugford M (2000) *Elementary Economic Evaluation in Health Care* (2nd ed). London: BMJ Books.

Johnston G, Crombie IK, Davies HTO et al (2000) Reviewing audit: barriers and facilitating factors for effective clinical audit. *Quality in Health Care* **9**:23–6.

Kahan B (2005/2006) *Interactive Domain Model Best Practices: The pre-determinants of the social determinants of health* – www.idmbestpractices.ca/idm.php?content=archjottings#brkOctober05.

Kee F, McDonald P, Gaffney B (1997) Prioritising the cardiac surgery waiting list. The angina patient's perspective. *Heart* **77**:330–2.

Keys A (1980) *Seven Countries: A multivariance analysis of death and coronary heart disease*. London: Harvard University Press.

Keys AB, Keys M (1975) *How to Eat Well and Stay Well: The Mediterranean Way*. London: Doubleday.

King's Fund (2006) *Patients at Risk of Re-hospitalisation (PARR) Case Finding Tool* – www.kingsfund.org.uk/current_projects/predictive_risk/patients_at_risk.html.

Kirkwood B, Sterne J (2003) *Essential Medical Statistics* (2nd ed). Oxford: Blackwell Science.

Knott L (2006*) Patient UK: Records and computers* – www.patient.co.uk/showdoc/40000769.

Kohlmann W, Gruber SB (2004, updated 2006) *Hereditary Non-Polyposis Colon Cancer* – www.geneclinics.org.

Koller T (2006) *WHO Interview with Prof Goran Dahlgren* – www.euro.who.int/socialdeterminants/socmarketing/20060221_6.

Kono T (1984) *Strategy and Structure of Japanese Enterprises*. London: Macmillan.

Kotler P, Zaltman G (1971) Social marketing: an approach to planned social change. *Journal of Marketing* **35**:3–12.

Künzli N, Kaiser R, Medina S et al (2000) Public-health impact of outdoor and traffic-related air pollution. A European assessment. *The Lancet* **356**:795–801.

Lakhani A, Coles J, Eayres D et al (2005) Creative use of existing clinical and health outcomes data to assess NHS performance in England: Part 1 – performance indicators closely linked to clinical care. *BMJ* **330**:1426–31

Lakhani A, Olearnik H, Eayres D (eds) (2007) *Compendium of Clinical and Health Indicator*s. London: The Information Centre for Health and Social Care/National Centre for Health Outcomes Development. Available online at: www.nchod.nhs.uk.

Lalonde M (1974) *A New Perspective on the Health of Canadians: A working document* – www.phac-aspc.gc.ca/ph-sp/phdd/pdf/perspective.pdf.

Larson (1977) *The Rise of Professionalism: A sociological analysis*. Berkeley, CA: University of California Press.

Last JM (1963) The iceberg: 'completing the clinical picture' in general practice. *The Lancet* **2**:28–31.

Lemert EM (1967) *Human Deviance, Social Problems and Social Control*. London: Prentice Hall.

Lewis G (2000) *A Failure of Treatment*. Oxford: Oxford University Press.

Lewis R, Hunt P, Carson D (2006) *Social Enterprise and Community Based Care: Is there a future for mutually-owned organizations in community and primary care?* London: King's Fund.

Lincoln P, Nutbeam D (2005) WHO and international initiatives. In: Davies M, Macdowell W (eds) *Health Promotion Theory*. London: McGraw-Hill.

Lindblom C (1959) The science of muddling through. *Public Administration Review* **19**:79–88.

Lo YM, Corbetta N, Chamberlain PF et al (1997) Presence of fetal DNA in maternal plasma and serum. *The Lancet* **369**:440–2.

London Health Commission (2002) *Health in London: 2002 Review of the London Health Strategy high-level indicators*. London: London Health Commission. Available online at: www.londonshealth.gov.uk/pdf/hinl2002.pdf.

Manuel DG, Lim J, Tanuseputro P et al (2006) Revisting Rose: Strategies for reducing coronary heart disease. *BMJ* **332**:659–62.

Marshall T (2000) Exploring a fiscal food policy: the case of diet and ischaemic heart disease. *BMJ* **320**:301–4.

Marshall T, Kennedy E, Offutt S (2000) Exploring a fiscal food policy: the case of diet and ischaemic heart disease. Commentary: Alternative nutrition and outcomes using a fiscal food policy. *BMJ* **320**:304–5.

Marshall T, Rouse A (2002) Resource implications and health benefits of primary prevention strategies for cardiovascular disease in people aged 30 to 74: mathematical modelling study. *BMJ* **325**:197.

Maslow AH (1943) A Theory of Human Motivation. *Psychological Review* **50**:370–96.

Maxwell RJ (1984) Quality assessment in health. *BMJ* **288**:1470–2.

Maynard A (2001) Economic-based medicine: an evolving paradigm. *Journal of the Royal College of Physicians of Edinburgh* **31** (Suppl 9):16–7.

McCarthy MI (2004) Progress in defining the molecular basis of type 2 diabetes mellitus through susceptibility-gene identification. *Human Molecular Genetics* **13**:R33–41.

McConnachie A, Hunt K, Emslie C et al (2001) 'Unwarranted survivals' and 'anomalous deaths' from coronary heart disease: prospective survey of general population. *BMJ* **323**:1487–91.

McKeown RE, Weed DL (2002) Ethics in epidemiology and public health II. Applied terms. *Journal of Epidemiology and Community Health* **56**:739–41.

Mechanic D (1968) *Medical Sociology*. New York: Free Press.

Montori, VM, Smieja M, Guyatt GH (2000) Publication bias: A brief review for clinicians. *Mayo Clinic Proceedings* **75**:1284–8.

Morgan O, Baker A (2006) Measuring deprivation in England and Wales using 2001 Carstairs scores. *Health Statistics Quarterly* **31**:28–33. Available online at: www.statistics.gov.uk/articles/hsq/HSQ31deprivation_using_carstairs.pdf.

Mullins LJ (2005) The nature of leadership. In: Mullins LJ (2005) *Management and Organisational Behaviour* (7th ed). Harlow: Financial Times/Prentice Hall.

Myer L, Morroni C, Link BG (2004). Impact of measurement error in the study of sexually transmitted infections. *Sexually Transmitted Infections*; **80**:318–21.

Naidoo J, Wills J (1998) *Practising Health Promotion: Dilemmas and challenges*. London: Baillière Tindall.

Naidoo J, Wills J (2000) *Health Promotion: Foundations for practice* (2nd edn). London: Baillière Tindall.

National Advisory Committee on Health and Disability (2003) *Screening to Improve Health in New Zealand: Criteria to assess screening programmes*. Wellington: National Advisory Committee on Health and Disability.

National Center for Chronic Disease Prevention and Health Promotion (2006) *Steps of an Outbreak Investigation* – www.cdc.gov/excite/classroom/outbreak/steps.htm.

National Human Genome Research Institute, National Institutes of Health (2007) *Learning about Cystic Fibrosis* – www.genome.gov/10001213.

National Institute for Clinical Excellence (2004) *Guide to the Methods of Technology Appraisal*. London: NICE. Available online at: www.nice.org.uk/niceMedia/pdf/TAP_Methods.pdf.

National Institute for Health and Clinical Excellence (2006) *Methods for Development of NICE Public Health Guidance*. London: NICE. Available online at: www.nice.org.uk/nicemedia/pdf/CPHEMethodsManual.pdf.

National PCT database (2004) *Measures of Need* – www.primary-care-db.org.uk.

Newsom JT (2006) *Distinguishing Between Random and Fixed: Variables, effects, and coefficients* – www.upa.pdx.edu/IOA/newsom/mlrclass/ho_randfixd.doc.

NHS Clinical Governance Support Team (1999) *Seven Pillars of Clinical Governance Model* – www.cgsupport.nhs.uk/downloads/Seven_Pillars.pdf.

NHS Litigation Authority (2007) *About the NHS Litigation Authority* – www.nhsla.com/Channels.

Northern Ireland Statistics and Research Agency (2005) *Northern Ireland Multiple Deprivation Measure*. Belfast: The Stationery Office. Available online at: www.nisra.gov.uk/aboutus/default.asp2.htm.

Nutbeam D, Harris E (2004) *Theory in a Nutshell: A guide to health promotion theory* (2nd ed). Maidenhead: McGraw-Hill Education.

Nutrional Advisory Group (1995) *Recommendations for a Food and Nutrition Policy for Ireland*. Dublin: Stationery Office.

Office for National Statistics (2007) *The National Statistics Socio-economic Classification*. London: Office for National Statistics. Available online at: www.statistics.gov.uk/methods_quality/ns_sec/default.asp.

O'Reilly D, Stevenson M (2004) Our data on the uptake of disability benefits a useful addition to census data in describing population health care needs? *Journal of Health Service Research Policy* **9**:165–70.

Parsons T (1951) *The Social System*. New York: Free Press.

Petrie A, Sabin C (2005) *Medical Statistics at a Glance* (2nd ed). Oxford: Blackwell Publishing.

PHLS Advisory Committee on Gastrointestinal Infections (2004) Preventing person-to-person spread following gastrointestinal infections: guidelines for public health physicians and environmental health officers. *Communicable Diseases and Public Health* **7**:362–84.

Pilowski I, Spence ND (1977) Ethnicity and illness behaviour. *Psychological Medicine* **7**:447–52.

Prochaska JO, DiClemente CC (1984) Self-change processes, self-efficacy and decisional balance across five stages of smoking cessation. In: *Advances in cancer control – 1983*. New York: Alan R. Liss, Inc.

Prochaska JO, Velicer WF (1997) The transtheoretical model of health behavior change. *American Journal of Health Promotion* **12**:38–48.

Puska P, Salonen JT, Nissinen A et al (1983) Change in risk factors for coronary heart disease during 10 years of a community intervention programme (North Karelia project). *BMJ* **287**:1840–4.

Rawls J (1971) *A Theory of Justice*. Cambridge: Harvard University Press.

Richardson SW, Wilson MC, Guyatt GH et al (1999) Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. *JAMA* **281**:1214–9.

Robinson V, Segrott J (2002) *Home Office Research Study 243: Understanding the decision-making of asylum seekers*. London: Home Office. Available online at: www.homeoffice.gov.uk/rds/pdfs2/hors243.pdf.

Rogers E (1995) *Diffusion of Innovation* (4th ed). New York: Free Press.

Rootman I, Goodstadt M, Hyndman B et al (2001) *Evaluation in health promotion: Principles and perspectives*. Copenhagen: WHO Regional Publications, European Series, No 92.

Rose G (1981) Strategy of prevention: lessons from cardiovascular disease. *BMJ* **282**:1847–51.

Rose G (1992) *The Strategy of Preventive Medicine*. Oxford: Oxford University Press.

Rootman I, Goodstadt M, Hyndman B et al (2001) *Evaluation in Health Promotion: Principles and perspectives*. European Series, No. 92. Copenhagen: WHO Regional Publications.

Rosenstock IM, Strecher VJ**,** Becker MH (1998) Social learning theory and the Health Belief Model. *Health Education Quarterly* **15**:175–83.

Salisbury DM (2005) Development of immunization policy and its implementation in the United Kingdom. *Health Affairs* **24**:744–54.

Sandman PM (1987) Risk communication: Facing public outrage. *EPA Journal* **Nov**:21–22. Available online at: www.psandman.com/articles/facing.htm.

Sassi F, Le Grand J, Archard L (2001) Equity versus efficiency: a dilemma for the NHS. *BMJ* **323**:762–3.

Scambler G (ed) (1997) *Sociology as Applied to Medicine* (4th ed). London: Saunders.

Schein EH, Bennis WG (1965) *Personal and Organizational Change through Group Methods: The laboratory approach*. New York: John Wiley & Sons.

Scientific Committee of the Food Safety Authority of Ireland, Nutrition Sub-committee (2007) *Salt and Health: Review of the Scientific Evidence and Recommendations for Public Policy in Ireland*. Dublin: Food Safety Authority of Ireland. Available online at: www.fsai.ie/publications/reports/salt_report.pdf.

Scottish Executive (2004) *Scottish Index of Multiple Deprivation: Overview*. Edinburgh: Scottish Executive. Available online at: www.scotland.gov.uk/Topics/Statistics/SIMD/Overview.

Sihota S, Lennard L (2004) *Health Literacy: Being able to make the most of health*. London: National Consumer Council. Available online at: www.ncc.org.uk/nccpdf/poldocs/NCC064_health_literacy.pdf.

Simon H (1947) *Administrative Behaviour. A study of decision-making processes in administrative organisations* (4th ed). New York: The Free Press.

Smithies J, Adams L (1990) *Community participation in health promotion*. London: Health Education Authority.

Snow J (1855) *On the mode of communication of cholera*. London: John Churchill; 1–38. Available online at: www.ph.ucla.edu/EPI/snow/snowbook.html.

Social Disadvantage Research Centre (SDRC), Department of Social Policy and Social Work at the University of Oxford (2007) *The English Indices of Deprivation 2007: Summary*. London: Department for Communities and Local Government. Available online at: www.communities.gov.uk/publications/communities/indicesdeprivation07.

Sproston K, Mindel J (eds) (2006) *Health Survey for England 2004*. Leeds: The Information Centre. Available online at: www.ic.nhs.uk/pubs.

Sylva K, Melhuish E, Sammons P et al (2004) *The Effective Provision of Pre-School Education (EPPE) Project. EPPE findings from the Early Primary Years*. London: DFES. Available online at: www.surestart.gov.uk/research/keyresearch/eppe.

Tannahill A (1985) What is health promotion? *Health Education Journal* **44**:167–8.

Tuckman BW (1965) Developmental sequence in small groups. *Psychological Bulletin* **63**:384–99.

Tufte ER (2001) *The Visual Display of Quantitative Information* (2nd edn). Connecticut: Graphics Press LLC.

United Nations (2007) *UN Millennium Development Goals*. United Nations. Available online at: www.un.org/millenniumgoals.

Walker P (1999) *Protecting and Using Patient Information: A manual for Caldicott Guardians*. London: Department of Health. Available online at: www.dh.gov.uk/assetRoot/04/06/81/36/04068136.pdf.

Walt G (2001) Global cooperation in international health. In: Merson M, Black RE, Mills AJ (eds) *International Public Health: Diseases, programs, systems, and policies*. New York: Aspen Publishers.

Wanless D (2002) *Securing our future health: Taking a long-term view*. London: HM Treasury. Available online at: www.hm-treasury.gov.uk/Consultations_and_Legislation/wanless/consult_wanless_final.cfm.

Wanless D (2004) *Securing Good Health for the Whole Population*. London: HMSO.

Wanless D (2006) *Securing good care for older people: Taking a long-term view*. London: King's Fund.

Waterhouse J, Muir C, Correa P et al (eds) (1976) *Cancer Incidence in Five Continents*, Vol 3.3. Lyon: IARC, 456.

Watkiss P, Brand C, Hurley F et al (2000) *On the Move: Informing transport health impact assessment in London*. London: Health Commission. Available online at: www.londonshealth.gov.uk/pdf/onthemove.pdf.

Weber M (1958) *The Three Types of Legitimate Rule* (translation). Quoted in: *Tripartite classification of authority*. Wikipedia website at: en.wikipedia.org/wiki/Tripartite_classification_of_authority. Berkeley: University of California.

Whincup PH, Owen CG, Sattar N, Cook DG (2005) School dinners and markers of cardiovascular health and type 2 diabetes in 13–16 year olds: cross sectional study. *BMJ* **331**:1060–1.

Whitely L (2007) *z, t and F tests* – www.fil.ion.ucl.ac.uk/spm/doc/mfd/FandTtests.ppt.

Wikipedia, the free encyclopedia. (2007) *Molecular biology* – en.wikipedia.org/wiki/Molecular_biology

Willis CE, Kee F, Beverland P, Watson JD (2000) Urban rural differences in total hip replacements. The next stage. *Journal of Public Health Medicine* **22**:435–8.

Wilson JM, Jungner G (1968) *Principles and Practice of Screening for Diseases*. Geneva: WHO.

WHO (1978) *Declaration of Alma-Ata, policy statement*. Geneva: WHO. Available online at: www.euro.who.int/AboutWHO/Policy/20010827_1.

WHO (1980) *International Classification of Functioning, Disability and Health (ICF)*. Geneva: WHO. Available online at: www.who.int/classifications/icf/en.

WHO (1986) *Ottawa Charter for Health Promotion. First International Conference on Health Promotion, Ottawa, 21 November 1986 – WHO/HPR/HEP/95.1*. Geneva: WHO. Available online at: www.who.int/hpr/NPH/docs/ottawa_charter_hp.pdf.

WHO (2000) *Health Systems Performance Report*. Geneva: WHO. Available online at: www.who.int/health-systems-performance.

WHO (2003a) *North Karelia Project: From Demonstration Project to National Activity*. Geneva: WHO. Available online at: www.who.int/hpr/successful.prevention.3.html.

WHO (2003b) *International Migration, Health and Human Rights*. Health & Human Rights Publication Series, Issue No. 4. Geneva: WHO.

WHO (2004a) *Global strategy: overall goal*. Geneva: WHO. Available online at: www.who.int/dietphysicalactivity/goals/en/.

WHO (2004b) Global Strategy on Diet, Physical Activity and Health. Geneva: WHO. Available online at: www.who.int/dietphysicalactivity/en/.

WHO (2005a) *The Bangkok Charter for Health Promotion in a Globalized World from: 6th Global Conference on Health Promotion, Bangkok, Thailand, August 2005*. Geneva: WHO. Available online at: www.who.int/healthpromotion/conferences/6gchp/bangkok_charter/en/.

WHO (2005b) *Framework convention on tobacco control*. Geneva: WHO. Available online at: www.who.int/tobacco/framework/en/.

WHO (2005c) *Statement of the WHO on International Trade and Health*. Geneva: WHO.

WHO (2006) *Healthy Cities and urban governance*. Geneva: WHO. Available online at: www.euro.who.int/healthy-cities.

WHO (2007a) *Ethical, Legal and Social Implications (ELSI) of Human Genomics*. Geneva: WHO. Available online at: www.who.int/genomics/elsi/en/.

WHO (2007b) *Water-Related Diseases*. Geneva: WHO. Available online at: www.who.int/water_sanitation_health/diseases/en.

WHO (2008) International Classification of Diseases (ICD). Geneva: WHO. Available online at: www.who.int/classifications/icd/en/.

Wilkinson M (2006) *Carbon Capture and Storage* – www.co2storage.org.uk.

Wilkinson R, Marmot M (2003) *Social Determinants of Health: The solid facts* (2nd ed). Copenhagen: WHO. Available online at: www.euro.who.int/document/e81384.pdf. A short, readable summary providing evidence of the health implications of social, cultural and economic circumstances.

World Health Annual of Statistics (1991) Available online at: www.who.int/whois/en.

Wright M (2005) Homosexuality and HIV/AIDS prevention: the challenge of transferring lessons learned from Western Europe to Central and Eastern European Countries. *Health Promotion International* **20**:91–8.

Zahra SA (1991) Predictors and financial outcomes of corporate entrepreneurship: An exploratory study. *Journal of Business Venturing* **6**:259–85.

Zborowski M (1952) Cultural components in response to pain. *Journal of Society Issues* **8**:16–30.

Zoonekynd V (2007) *Probability Distributions* – zoonek2.free.fr/UNIX/48_R/07.html.

INDEX