Italian College of Anaesthesiologists (ICA)

Teaching and Training in Anaesthesiology and Critical Care

S.B. Martins
W.A. Zin

Planning and Designing Clinical Research

Vol. 1

Continuing Medical Education (CME)

The Basics

Springer

Planning and Designing Clinical Research

Springer-Verlag Italia Srl.

S.B. Martins
W.A. Zin

Planning and Designing Clinical Research

Series edited by
Antonino Gullo

Springer

S.B. Martins
Fellow in Health Services Research
University of California at San Francisco
San Francisco, USA

W.A. Zin
Professor of Physiology
Chief, Laboratory of Respiration Physiology
Carlos Chagas Filho Biophysics Institute
Federal University of Rio de Janeiro
Rio de Janeiro, Brasil

Series Teaching and Training in Anaesthesiology and Critical Care edited by
A. Gullo, MD
Department of Anaesthesia, Intensive Care and Pain Therapy
University of Trieste, Cattinara Hospital, Italy

Originally published by Springer-Verlag Italia,Milano in 2001

ISBN 978-88-470-0139-8 ISBN 978-88-470-2919-4 (eBook)
DOI 10.1007/978-88-470-2919-4

Library of Congress Cataloging-in-Publication Data: applied for

Cover design: Simona Colombo, Milan
Typesetting: Photolife, Milan

SPIN: 10790128


Italian College of Anaesthesiology (ICA)

Forewords

I am greatly pleased and honored to introduce to the National and International Scientific Community a new publication by the Italian College of Anaesthesiologists. The first issue of this series is entitled *Teaching and training in anaesthesiology and critical care* and is devoted to a topic of fundamental importance: **planning and designing clinical research**, the contribution of two prestigious colleagues in basic disciplines, Susan M. Martins and Walter A. Zin.

The Italian College of Anaesthesiologists includes teachers of anaesthesiology and critical care and was set up in 1986 in order to gain more information about how the different Italian schools operate. Its other, deeper objective was a further strengthening of the collaboration regarding the training of future specialists and refresher courses for all those who need them and intend to maintain a close relationship with the school where they trained. Of course, the college of teachers has a very difficult task, which is harmonizing the basic aspects of teaching and also supporting its pupils in their methodological research in order to optimize the clinical approach.

Teachers, specialists, and scholars are aware that the learning process is not accomplished in just a few years, because the range of topics is wide and the technology complex. Anaesthesiology yesterday and perioperative medicine today, just like intensive care yesterday and critical care today, need the contribution of various experiences and skills, which have the main objective of training specialists in order to provide patients in any acute or chronic condition with the highest possible standards of professional care. This is one of the main goals and the *raison d'être* of the Italian College of Anaesthesiologists.

Twelve years have passed since we discussed, together with many famous colleagues, the institutional goals the College had to represent. This idea has matured, developed in other countries by various professional colleges in other disciplines to create a philosophy within research, teaching, and clinical medicine. It is hoped that this initiative may also provide food for thought for us teachers and an impetus to continue our journey in such a fascinating discipline as anaesthesiology. I thank you all for your support for this project – in certain respects an ambitious one – which better aligns the Italian College of Anaesthesiologists with its institutional tasks.

Florence, October 2000

Gian Paolo Novelli
Founding President of the
Italian College of Anaesthesiologists

It is with great pleasure that I present the first volume of the new series *Teaching and training in anaesthesiology and critical care*, published by Springer-Verlag Italia on behalf of the Italian College of Anaesthesiologists.

This initiative, to which my colleague Professor Antonino Gullo has significantly contributed, falls under the College's objective of optimizing teaching methods and training techniques in anesthesiology disciplines.

This first volume, **Planning and designing clinical research**, was written by Susan M. Martins and Walter A. Zin, two highly qualified scholars in the field of basic disciplines, and is extremely important for those who want to develop their studies and research on the basis of clinical activity.

My wish is that this series becomes a means of communication between teachers, young and more experienced specialists as well as scholars of the discipline, facilitating fruitful debate on the various aspects of education and training in anaesthesiology disciplines.

I would also like to thank Springer-Verlag for their help in making this publication possible.

Padua, October 2000

Giampiero Giron
President of the
Italian College of Anaesthesiologists

Contents

Planning and designing clinical research

What is the meaning of research?

In a broad sense the word *research* may be used for the act of informing oneself about what one does not know, generally by retrieving bits of information here and there. Research is in fact the systematic process of collecting and analyzing data in order to increase the available knowledge on a specific field of interest. The whole process presents seven distinct characteristics:

1. question or problem;
2. aim;
3. specific plan of procedure;
4. division of the principal problem into subproblems;
5. acceptance of critical assumptions;
6. collection and interpretation of data;
7. cyclical nature.

Question or problem

A question, problem, or hypothesis should originate one's research. In all fields there are so many unanswered questions, it is just a matter of asking oneself: "Why?" "What is the meaning of this?" "What has caused that?". Accidents may happen, of course, but an experienced researcher will never start data collection expecting that these data will show him/her the way to follow. The problem establishes the origin of formal research.

Aim

A clear and unambiguous statement of the aim of the research is imperative. One must start up a project knowing exactly what is intended. "What is the goal?" represents the question to be answered at this step of the project.

Specific plan of procedure

Research requires discipline: the plan of procedure must be carefully and explicitly decided in advance. In other words, planning and designing are fundamental steps. Clearly, depending on the specific project, different designs and methods will be chosen in order to generate data of importance to the research problem. The question to be answered is: "How to reach the goal?"

Division of the principal problem into subproblems

Sometimes it is wise to reduce the main problem to a series of logical subproblems. The independent solution of these subproblems will lead to the unravelling of the overall question. When applicable, this technique may change a cumbersome and unwieldy research project into a series of easy-to-handle smaller studies.

Acceptance of critical assumptions

Assumptions are self-evident truths that must be valid or else the research can not proceed. In other words, an assumption is a condition that is taken for granted, without which the research could not be carried on. It is fundamental that the reader knows what is assumed with respect to one's project. If the assumption is not self-evident the author should mention it explicitly.

Collection and interpretation of data

Data collection must be done very carefully. If data are gathered in a sloppy way, the meaning of the results may be distorted. Usually a series of experiments is carried out in order to train the members of the research team, as surgeons train their fellows, until the group knows exactly what needs to be done. Notes must be taken as frequently as possible, so that later on anyone can understand the meaning of a sudden drop in arterial blood pressure during the experiment, for instance. Moreover, another researcher may want to repeat the experiment, and these precisely taken notes will help him in his task.

Interpretation of the results allows the discovery of the meaning of the data. Without it there is no research. It must be stressed that the interpretation of the data may be different among researchers. An analogy to this evidence was magnificently created by Luigi Pirandello, an Italian dramatist, novelist, and Nobel prize winner in his play *Così è si vi pare* (1918), in which the same facts are presented to different characters, who report back non-coincident views of them. The interpretation of the same set of data may change with time too. For instance, the realities of time and space have not changed; the way we interpret them has.

Cyclical nature

The research process follows a cycle. It begins with the initial questioning, then follows the identification of a problem, data are collected and interpreted, a conclusion is reached, and the initial hypothesis is either supported or not supported (the question is partially/completely answered or not).

Although the circle seems neatly closed, this is a deceptive conclusion. Research is never conclusive, because in exploring an area, one comes across additional problems that need solving. Thus, research begets research, and the circle becomes a spiral.

Introduction to the general tools of research

Every professional needs particular tools in order to carry out his work. Without scalpel and scissors, the surgeon finds himself in rough waters to perform his job safely. Naturally, researchers have their own kit of tools to achieve their aims. Furthermore, the specific tools required by a physiologist, for instance, will be quite distinct from those needed by an astronomer. However, the great majority of researchers will share a set of general tools for research.

Four general tools of research are frequently identified:

1. the library and its resources;
2. the computer and its software;
3. statistics;
4. facility with language.

The library and its resources

The library has evolved from a repository of writings, books, manuscripts, and so on into a very dynamic space which may be visited even electronically. As in the past, ease and speed of access to information are priorities, but with the avalanche of new knowledge produced daily, researchers became more demanding.

In order to keep up to date with the researchers' present needs and desires, the ambience of a modern library has been changing. Besides the card catalogue and shelf after shelf of books and periodicals, one will find rows of computer terminals and keyboards. The final goal is to access the whole realm of global knowledge rapidly and comprehensively.

The university library offers both printed material as well as electronically accessed information. Thus, it will provide the reader with: (1) periodicals and books, together with a card catalogue; (2) CD-ROM (Compact Disk - Read Only Memory) stored databases; (3) local computer-stored databases; (4) remotely accessed databases; (5) commuting facilities.

Typically, the card catalogue indicates if the desired periodical/book is part of the library collection. Then, it is just a matter of retrieving the volume from its shelf for consultation. CD-ROMs contain a great deal of information (650 megabytes) and are still frequently used in libraries. Many databases are stored in CD-ROMs. Depending on the computer network available at the university, they can be used at the library or from one's office. In many universities information is stored in local computers or networks. They replace the catalogue card and also provide a way of searching for terms, authors, and keywords. Remotely accessed databases are physically located somewhere else, even in another country. The users purchase the right to access them electronically from specified computers.

CD-ROMs, locally and remotely located databases present characteristics that must be specified by the users when purchasing them. One may be given only the complete reference of a paper, its abstract, or the full paper. Furthermore, some databases allow search of the literature by author's name and/or terms, thus providing a fast and comprehensive way of finding the relevant papers in a specific field of knowledge. Finally, if one desires to have a copy of a certain paper that can-

not be found in the library, it may be forwarded in a matter of hours by the commuting system of the library. The librarian places an order to another library that has the required periodical. The librarian at the latter scans the paper and sends it electronically to the former library, where it will be printed on a high-quality machine. In summary, the modern library can astonishingly shorten the process that begins with consulting the literature and ends with the availability of the information for immediate reading.

The computer and its software

The computer is perhaps the workhouse of the generalized tools of research. It can perform a myriad of tasks at an incredible speed. However, one must take into account that it is not a human brain, and, indeed, it is a very limited machine that depends on a person at the keyboard to direct its actions. As a research assistant, the computer may play a role in: (1) planning the study; (2) literature review; (3) reference managing; (4) data gathering; (5) data analysis; (6) statistical analysis; (7) graphic production; (8) reporting.

During the planning phase of the study software can be used to help generate and organize ideas for the research, as well as to assist in budget planning. Computers can be used to review the literature by using technology such as CD-ROM and online database information acquisition. They are also able to provide telecommunication assistance (e-mail, electronic bulletin boards, list servers, Internet, World Wide Web, and so on). Software can easily assist with electronic storage and retrieval, so that valuable bibliographic citations and comments can be readily sorted and accessed by title, author, keyword, etc.

By means of analogue-to-digital (A-D) converters, computers can sample and store experimentally generated data. This is a fundamental step in setting the stage for data analysis. Paper recording of information is long gone thanks to A-D boards. Data points can be sampled at very high frequencies, increasing the accuracy and precision of the results. Data analysis can be performed off line by means of general purpose and/or dedicated software. The latter usually suit better the neophyte, whereas the former yield a more-elastic and comprehensive analysis. Data analysis can also be performed on line by dedicated programs that run in computer-controlled machines, such as ventilators, spirometers, or capnographs. Different statistical and spreadsheet software packages can operate on a personal computer, thus allowing a quick analysis of various types of data sets. Some of these programs even guide the user towards the right statistical test to be used.

In the late 1980s plotting a curve or drawing a graph meant sharpening up one's pencils, and later calling on the artist. Nowadays a fairly sophisticated figure can be prepared in no time by means of powerful software.

The preparation of the manuscript is facilitated by the use of word processing to write and edit many versions and updates of the final report. Furthermore, desktop publishing software may help to produce professional-looking documents that can be presented at conferences and elsewhere. In this same line, figures and text can be combined in a slide presentation by other software.

In summary, the computer and its software represent an extremely powerful

tool of research that takes part in practically all the steps of a research project. Last but not least, laptop computers add the mobility factor to the "old" desktop personal computer, while maintaining the same degree of precision and usefulness.

Statistics

Statistical analysis is a powerful tool when used correctly. However its limitations ought to be considered. Statistical values are never the final answer to the research problem. One must always bear in mind that behind statistics lie the data to which the tests refer. Statistics provides information about the data, but an experienced researcher will never be satisfied until the meaning of the data is unveiled. Statistics has two principal functions: descriptive and inferential. Descriptive statistics provide information such as mean, standard deviation, median, etc. of the data, and, in the case of two or more groups of data, their relationship. Inferential statistics seeks to fit data to the ideal form of a statistical model.

The fundamental problem in statistical analysis lies in the choice of the right test to be applied in order to reveal the meaning of the data. To start with, the experimented researcher finds out which major set of tests may be applied to his data: parametric or non-parametric. A great deal of scientific reports erroneously forward a conclusion based on parametric statistical analyses when the distribution of the data points is clearly not normal. Presently powerful software (that run on personal computers) guide the novice researcher to find the most-suitable test to answer his/her question. In summary, statistics is a powerful tool of research that can help understand the meaning of the data. However, it must be applied cautiously to avoid misleading conclusions.

Facility with language

A successful research project ends as a published document. To generate such a paper, the author must be able to use language with a degree of clarity that will make easy the reading of the document in all its aspects. As a rule of thumb, after several drafts the author should hand the manuscript to a colleague not involved in the project, so that its clarity can be checked. Many good studies are not accepted for publication simply because the reviewers cannot understand their meaning. An experienced writer will always make good use of a well-thumbed dictionary. The spelling of every single word must be checked. In summary, facility with the language paves the way to a well-written document. For this purpose, experienced help is always needed.

Literature search

When developing a protocol for a research project it is crucial to write a background overview of the current evidence to support the need for the study. The background overview does not have to be an extensive review of the topic, but it does need to highlight the current issues that support the specific research question.

The development of the research question will guide the search for the evidence, and by increasing your knowledge of the topic you will be able to further develop your hypothesis. Initially review papers are a good starting point to expand your comprehension of a given subject, but one needs to be aware of the potential for bias in traditional reviews.

Traditional reviews often are not clear about the question that elicited the review or the selection of primary studies. This makes it difficult for readers to evaluate how comprehensive the author was in his/her search for primary studies, the quality of included and excluded studies, and consequently the conclusions reached. They have been criticized as haphazard and biased, and subject to the idiosyncratic impressions of the individual reviewer [1].

Systematic reviews are explicit on the question they are trying to answer, the methodology used, i.e., the selection of primary studies, the evaluation for quality markers, and the statistical methods used to combine results (if this was possible). The combination of results of different studies addressing the same question is called meta-analysis. Unique advantages of quantitative systematic reviews or meta-analyses are increased power and precision in estimating effects and risks. Qualitative and quantitative systematic reviews, with their explicit methods, may limit bias and improve the reliability and accuracy of recommendations [1]. "Researchers use the systematic review to identify, justify, and refine hypotheses; recognize and avoid pitfalls of previous work; estimate sample sizes; and delineate important ancillary or adverse effects and covariates that warrant consideration in future studies" [2].

Today searching for evidence has become an easier task due to the Internet. Electronic access to databases and journals on the Internet is increasing every year; and much of this information is free, such as that of PubMed and The Canadian Medical Journal, some is through university privileges, and some only by subscription.

However computer skills are required, but the development of user-friendly software makes the experience less "technical" and more easily accessible.

The decision as to which databases will be searched to look for the evidence will depend on the topic of interest. Evidence from searches in different topics have shown that there is much to be gained from searching different databases, as journal coverage and indexing terms differ [3-7]. Talk to your librarian and other researchers in the field to find out about the available databases. Language can be a constraint, so evaluate the benefit of looking for primary studies in other languages, for example LILACS (Latin American and Caribbean Health Science Information Database).

The quality of indexing of bibliographic databases will affect the success of your search strategy. The National Library of Medicine together with the Cochrane Collaboration is updating the indexing of clinical trials to improve researchers' access to them. The Cochrane Collaboration's database of clinical trials is regularly updated, and is not limited by journal of publication, as are commercial databases.

Efficient literature searching is a skill acquired through experience and use of effective search strategies. Evidence is published in a range of sources, including

journals, books, and research reports. For example, 20,000-30,000 biomedical journals and about 17,000 new biomedical books are published annually, and Medline already holds over 7 million articles [8].

Although the quantity of materials may seem overwhelming, in practice there are four sources of evidence that can help make your search more efficient.

1. Specialist organizations that are dedicated to summarizing research findings and creating databases to aid dissemination.
   A. The Cochrane Collaboration – an international organization that aims to help people make well-informed decisions about health care by preparing, maintaining, and ensuring the accessibility of systematic reviews of the effects of health care interventions. Italian Cochrane Center – *www.areas.it/index.htm*; Australian Cochrane Center – *www.cochrane.org.au.*
   B. NHS Centre for Reviews and Disseminations (University of York), established in January 1994 to provide the National Health Service (United Kingdom) with important information on the effectiveness of treatments and the delivery and organization of health care. It is a good resource for systematic reviews – http://www.york.ac.uk/inst/crd/welcome.htm.
2. Specialist journals which provide systematic reviews or review high-quality primary search.
   A. *ACP Journal* – bimonthly publication of structured abstracts of selected studies published in over 50 journals related to internal medicine that have met explicit criteria for scientific merit and clinical importance. A commentary from experts in the subject area is included. Published since 1991. http://www.acponline.org/journals/acpjc/jcmenu.htm.
   B. *Evidence-Based Medicine* – articles in the area of internal medicine, general and family practice, surgery, psychiatry, pediatrics, and gynecology and obstetrics are summarized into value-added abstracts and commentated by clinical experts if studies meet explicit criteria for scientific merit and clinical importance. Published since 1995. http://www.bmjpg.com/data/ebm.htm.
   C. *The Cochrane Library* – an electronic publication designed to supply high-quality evidence about health/care related issues, contains the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effectiveness, and the Cochrane Controlled Trials Register; http://www.cochrane.de/cc/cochrane/cdsr.htm.
3. General and specialized databases which you can search using specific techniques to find relevant articles.
   A. MEDLINE – National Library of Medicine (USA) bibliographic database of biomedical indexed journals. Free internet access through PubMed; http://www.ncbi.nlm.nih.gov/entrez/query.fcgi.
   B. EMBASE – published by Elsevier Science, it is a comprehensive pharmacological and biomedical database.
   C. Cancer-CD – a compilation by Silver Platter of Cancerlit and EMBASE cancer-related records from 1984.
   D. Psychlit – produced by the American Psychological Association, this database covers the areas of psychology, psychiatry, and related subjects. Journals from 1974 and books from 1987 are included.

   E. LILACS – Latin American and Caribbean Health Science Information Database, published since 1982. The database indexes 670 journals in the region, with abstracts in English, Portuguese, or Spanish. Only 41 journals overlap with MEDLINE or EMBASE.
4. The Internet offers direct links to organizations that may hold the evidence you need.
   A. Medical associations;
   B. Government Organizations;
   C. Pharmaceutical industry.

The search strategy

A structured question will help you focus the search, deciding on which sources and search strategies to use. When deciding on a search strategy, it is recommended to start with a broader search approach. Although this may generate a lot of irrelevant material, limiting the search too early may cause you to overlook a vital piece of information that is pertinent to answering your question.

Deciding on the best words with which to begin a search is critical to identifying key references. You can start by using a "text word", known as "natural language". This "text word" can be used to search the database directly or to identify keywords by which related articles are indexed. Most databases have their own indexing terms. For example, MEDLINE has Medical Subject Headings (MeSH). MeSH is a controlled vocabulary where authors make a deliberate choice of the terms that describe the content of the study. In this way synonyms and spelling differences can be grouped into a single group to improve retrieval of relevant papers. Identify the indexed vocabulary of the database that best describes your keywords and include them in your strategy. For example, in MEDLINE a search for "kidney diseases" using this as a "text word" would retrieve only articles in which this exact term appears in the title or the abstract. If used as a MeSH term, articles that contained "renal disease" would also be retrieved.

Search strategies have been developed to enhance your ability to retrieve relevant articles in MEDLINE. For example, clinical trials for therapy effectiveness, cohort studies for questions about prognosis, etc. Use the Boolean operator "AND" with the keywords as "textwords", or MeSH terms you have identified together with the search strategies described in Table 1.

Not one strategy will pick up 100% of the articles relevant to your question. Checking the references of the papers, asking experts in the area, and the use of different databases are ways of expanding your bibliography and making sure you are not missing an important piece of work.

Levels of evidence

The strength of the evidence of a result will be determined by the fulfilment of quality indicators in the research report. The research question will determine the type of study design in which the best evidence can be found. For example, when evaluating therapy efficacy, a double-blind randomized controlled design is the

Table 1. Search strategies for MEDLINE

Area of interest	Best single term search	MEDLINE
Articles on therapy [9]	Clinical trial	Field: publication type
	Random	Textword
Review articles [10]	Meta-analysis	Field: publication type
	Review	Field: publication type
Article on prognosis [11]	Explode cohort studies	
Article about harm [9]	Risk	Textword
Articles on diagnosis [9]	Sensitivity	Textword

gold standard, and when evaluating prognosis, a prospective cohort study design is the gold standard.

The highest level of evidence is usually found in reviews of primary research conducted in a scientific manner, i.e., the systematic review. A systematic review will explicitly answer a focused question, clearly defining the search criteria for the selection of primary papers (published and unpublished), evaluate these for quality markers, extract the data systematically, and analyze the data using a validated method. Homogeneity (the direction and degree of the results of the primary studies) is evaluated to further strengthen the level of the evidence.

Meta-analysis is the technique through which data from primary studies are pooled and re-analyzed. The data can be abstracted from the primary research reports or from the original patient data requested from the authors. Unpublished data can also be collected for the meta-analysis. This technique allows for an increase in power by enlarging the "sample size", i.e., adding several study samples.

Interventions receive grades of recommendation that are determined by the strength of the evidence. Table 2 is a summary of the grades of recommendations for therapy as displayed by the Centre for Evidence-Based Medicine [9]. Clinical practice guidelines are now expressing their recommendations with the grade of the evidence [10-12].

Databases

Electronic databases are very important tools for finding information about healthcare for research or patient management issues. We will describe "PubMed", "EMBASE", and "The Cochrane Library".

PubMed

PubMed is the National Library of Medicine search service that provides access to over 11 million citations in MEDLINE, PreMEDLINE, and other related databases, with links to participating online journals, through the Internet – http://www.ncbi.nlm.nih.gov/entrez/query.fcgi. PubMed provides access to bibliographic information, which is drawn primarily from MEDLINE, PreMEDLINE,

Table 2. Levels of Evidence and Grades of Recommendations

Grade of recommendation	Level of evidence	Therapy/prevention
A	1a	Systematic review (with homogeneity)
	1b	Individual RCT (with narrow confidence interval)
B	2a	Systematic review of cohort studies
	2b	Individual cohort study (including low quality RCT, e.g., 80% follow-up)
	3a	Systematic review (with homogeneity) of case-control studies
	3b	Individual case-control study
C	4	Case-series
D	5	Expert opinion without explicit critical appraisal, or based on physiology, bench research, or "first principles"

RCT, Randomized controlled trial

HealthSTAR, as well as Publisher-Supplied citations. On the left-hand-side menu "Journal Browser" (Fig. 1) will link you to a list of Web-based journals. New journals are regularly added. Web-based journals usually contain the full text of the original article, but this is not always the case. It varies according to publisher and journal. Some sites may require that you register, subscribe, or pay a fee in order to view the full text of an article. Contact the journal publishers as noted on their individual Web sites for specific access information.

MEDLINE. MEDLINE is the National Library of Medicine premier bibliographic database covering the fields of medicine, nursing, dentistry, veterinary medicine, the health care system, and the preclinical sciences. MEDLINE contains bibliographic citations and author abstracts from more than 4,000 biomedical journals published in the United States and 70 other countries. The file contains over 10 million citations dating back to 1966. Coverage is worldwide, but most records are from English language sources or have English abstracts.

PreMEDLINE. PreMEDLINE, the National Library of Medicine's in-process database for MEDLINE, provides basic citation information and abstracts before the citation is indexed with NLM's MeSH heading and added to MEDLINE. New records are added to PreMEDLINE daily and appear in PubMed with the tag

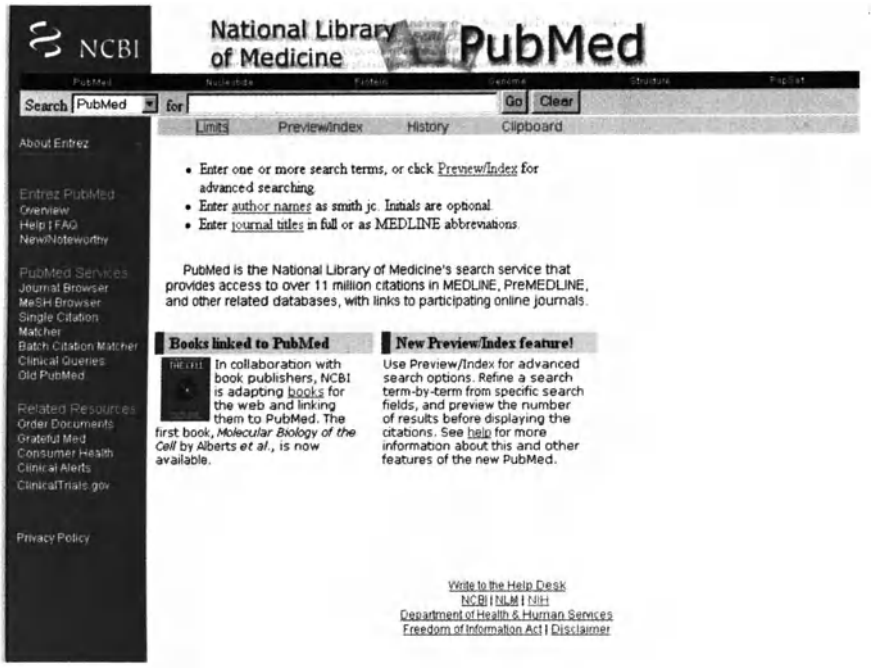

Fig. 1. PubMed initial screen (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed)

"[MEDLINE record in process]". After MeSH terms, publication types, GenBank accession numbers, and other indexing data are added, the completed citations are added weekly to MEDLINE. PreMEDLINE citations are incorporated into PubMed on a daily basis

Searching Pubmed. The opening screen of PubMed is quite user friendly. The horizontal menu includes links that will help you develop and focus your search strategy. This menu includes "Limits", "History", "Preview/index", and "Clipboard".

Limits. The "Limits" option can be activated with a check mark (Fig. 2). The "Limits" window displays field tags that are available to focus your search. There is an "All fields" option where the choices you can select from include: author name, issue, journal name, language, MeSH major topic, MeSH term, publication type, text word, and title word.

Additional fields exist to further support the focusing of your search: publication type, age group, publication date, languages, human or animal, database subset, and gender. Examples of details of choices for different fields are described in Table 3.

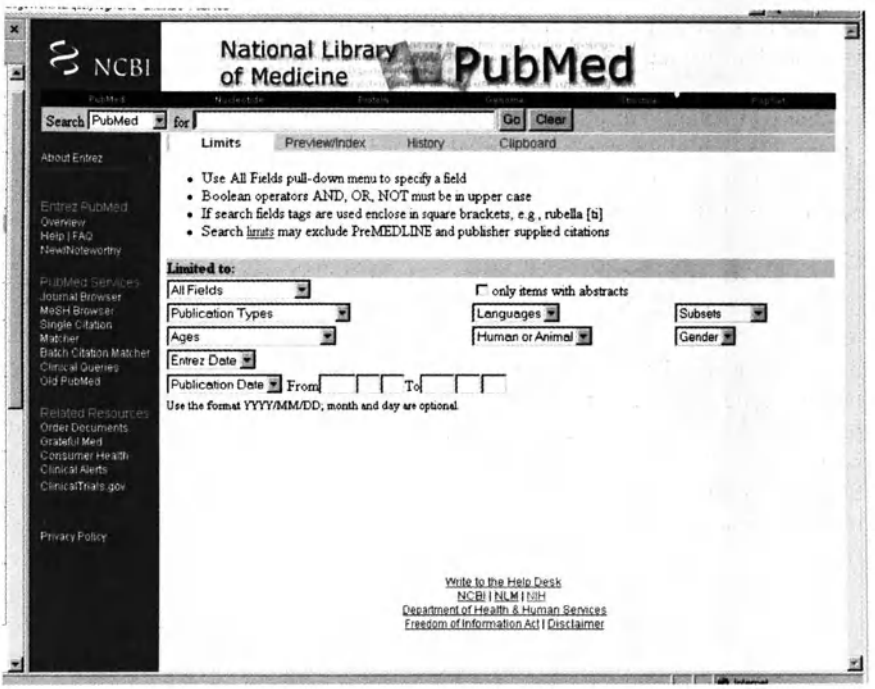

Fig. 2. "Limits" screen in PubMed (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Limits&DB=PubMed)

Preview/Index. In this window you can preview the number of articles retrieved when searching (Fig. 3). This feature helps you decide on the need to improve your search strategy before displaying the abstracts.

History. A complete history of your search strategy is displayed. From this window

Table 3. Description of options within "Field Tags" in the "Limits" section of PubMed

Field: publication type	Field: ages	Field: languages
Clinical trial	Newborn: birth to 1 month	English
Editorial	Infant: 1-23 months	French
Letter	Preschool child: 2-5 years	German
Meta-analysis	Child: 6-12 years	Italian
Randomized controlled trial	Adolescent: 13-18 years	Japanese
Practice guideline	Adult: 19-44 years	Russian
Review	Middle aged: 45-64 years	Spanish
	Aged 65+: 65+ years	
	Aged 80 and over: 80+ years	

you can increment search strategies by using the numbers of the previous searches rather than having to type the whole strategy out.

Clipboard. Clipboard displays the records you have selected (marked) when viewing the abstracts. Records can be added, limited only to a time interval (60 min. of inactivity) or a total number of 500. You can save the records at any time in the available formats (citation, MEDLINE, etc.).

MeSH. MeSH is the National Library of Medicine's controlled vocabulary used for indexing articles in PubMed. MeSH terminology provides a consistent way to retrieve information that may use different terms for the same concepts. The hyperlink MeSH Browser (http://www.nlm.nih.gov/mesh/MBrowser.html) on the left hand side menu of PubMed can be used to find items of interest and see these in relationship to others. MeSH organizes its terms in a hierarchical structure, so that broad searches may include articles indexed more narrowly. This structure also provides an effective way for researchers to browse MeSH in order to find appropriate terms and maximize the search efficiency and efficacy.

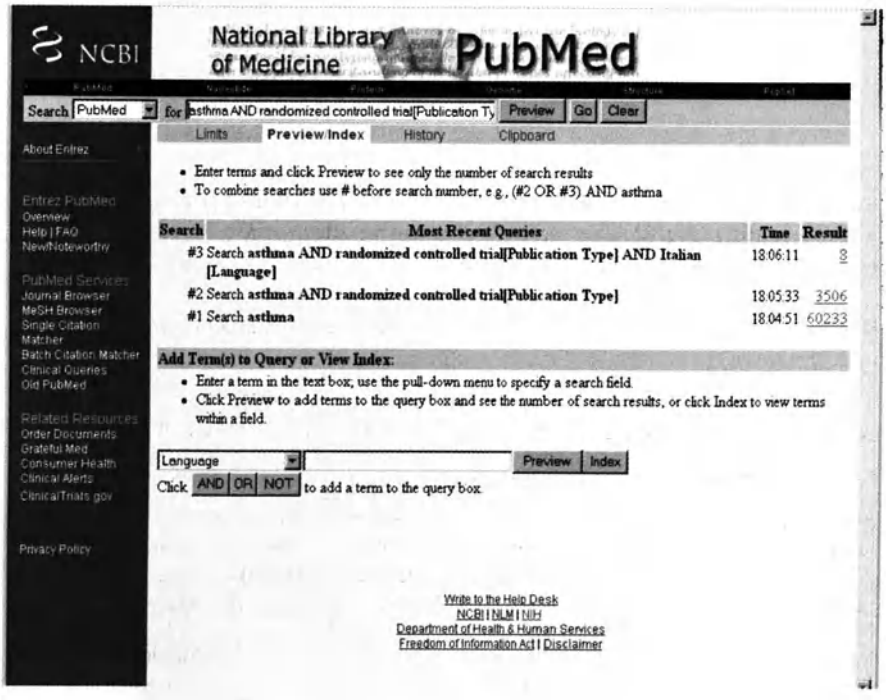

Fig. 3. Preview/index window in PubMed (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Preview&DB=PubMed)

Boolean operators. Boolean logic operators allow you to combine search results. The most commonly used are "AND", "OR", and "NOT". "AND" will narrow your results to articles that include search criteria of different search sets, whilst "OR" will combine all the articles found in different search sets broadening the results. "NOT" will exclude from the search articles with a given characteristic.

Related articles. One useful feature is the "[See related articles]" tag that appears after the citation you retrieve. If your search yields an article that meets your needs and you would like to see more papers on the same topic, clicking on this link would retrieve additional related articles.

EMBASE

EMBASE from Elsevier Science is a comprehensive pharmacological and biomedical database (biology, medicine, psychiatry, and drugs) renowned for extensive indexing of drug information from 3,600 journals published in 70 countries. EMBASE is available only by subscription. Updated monthly, EMBASE is one of the most-current biomedical databases available. Updates of the database appear in less than a month after publication of the journal. The database contains over 3 million records from 1980 to the present, with 375,000 new records added annually. Each record contains the full bibliographic citation, indexing terms, and codes. More than 65% of the records contain abstracts. The database includes EMTREE, a hierarchically ordered controlled thesaurus, which contains 38,000 preferred terms and more than 150,000 synonyms.

The Cochrane Library

The Cochrane Library is an electronic publication designed to supply high-quality evidence about health care-related issues. It is published quarterly on CD-ROM and the Internet (http://www.cochrane.de/cc/cochrane/cdsr.htm; http://*www.cochrane.org/cochrane/revabstr/ccsales.htm*). The Cochrane Library includes three databases plus other resources.

1. *Cochrane Database of Systematic Reviews (CDSR)* – a rapidly growing collection of regularly updated, systematic reviews of the effects of health care, maintained by contributors to the Cochrane Collaboration committed to 'preparing, maintaining and disseminating systematic reviews of the effects of healthcare'. New reviews are added with each issue of *The Cochrane Library*. The Cochrane reviews deal mainly with randomized controlled trials.
2. *Database of Abstracts of Reviews of Effectiveness* (DARE) – includes structured abstracts of systematic reviews from around the world, which have been critically appraised by reviewers at the National Health Service Centre for Reviews and Dissemination at the University of York, England.
3. *Cochrane Controlled Trials Register (CCTR)* – is a bibliography of controlled trials identified by contributors to the Cochrane Collaboration and others, as part of an international effort to hand search the world's journals and create an unbi-

ased source of data for systematic reviews. CCTR includes reports published in conference proceedings and in many other sources not currently listed in MEDLINE or other bibliographic databases.

4. Other resources:
   A. *the Cochrane Review Methodology Database* – a bibliography of articles and books on the science of research synthesis;
   B. a handbook on critical appraisal and the science of reviewing research;
   C. a glossary of methodological terms;
   D. contact details for collaborative review groups and other entities in the Cochrane Collaboration;
   E. Netting the Evidence – a ScHARR Introduction to Evidence Based Practice on the Internet (http://www.shef.ac.uk/~scharr/ir/netting.html).

Managing your references

The development of software to manage references has certainly made a great impact on facilitating the organization of references and writing up of manuscripts. Personal bibliographic software is designed to handle bibliographic information (journal articles, books, book chapters, conference proceedings, magazine articles, audiovisual material, patents, or even your own reference type) in a user-friendly way.

You can develop many databases within the software categorizing them by subject area or by project. It allows you to import references from databases accessed through the Internet (for example PubMed) or from disks of searches done in the library. Of course, you also have the option to type in the individual references.

The powerful search capability helps you find the references you need within the database. Search your reference database as you do your on-line databases by using Boolean criteria, e.g., "Find all the references that contain the keywords 'sleep' and 'snore', with 'Smith' as one of the authors but not 'Jones'".

When writing your manuscript you can automatically insert citations by locating the reference you wish to cite and instruct the software to insert the citation at the cursor position in your word processor text. The software has pre-defined output formats for writing citations (e.g., Harvard or Vancouver formats), or you can define your own. You can then generate bibliography automatically. Some examples of personal bibliographic software are (http://www.bilaney.com): ENDNOTE, PROCITE, and REFERENCE MANAGER.

Critical appraisal skills

Interpretation of scientific studies can sometimes appear to be a daunting task, especially if little training was received in this area. A manageable and user-friendly approach has been developed by the *Evidence-Based Medicine Working Group*, which greatly improves one's confidence and ability to discuss and under-

stand research. These have been published in the *Journal of the American Medical Association* (JAMA) [13-21], and supplemented by a series in the British Medical Journal (BMJ) [22-28]. The Critical Appraisal Skills Program (http://www.public-health.org.uk/casp/) is also an excellent source to help develop these skills.

There are three sets of questions that need to be answered when critically appraising a research report [13]. The first set are screening questions to identify quality markers, when a decision is made whether to continue reading the paper or to drop it. The second set of questions elicits a clear understanding of the results; the third set relates to the application of the results in your particular setting: can we trust the results? (validity of the results); what do the results mean? (what are the results); are the results relevant to our situation? (application of the results).

The critical appraisal of the research design, methodology, quality of the data, and the analysis and interpretation placed on the findings will allow us to answer these questions. A general rule is not to assume that any aspect of research is sound. The search for positive information on quality markers will screen articles on their worth and allow for investing time in reading the relevant papers. For example, if a study claims to provide strong evidence in a given direction, but the appraisal identifies significant doubt about the quality markers of the study then a decision can be made to reject the research or at least to view it with some skepticism.

The question you have set out to answer and the study design will determine the quality markers you will look for when appraising an article. With practice, you will be able to screen articles in a systematic way and invest your time in interpreting relevant research. We will describe issues relevant to the critical appraisal of clinical trials. Different study designs have different quality markers. The resources for critically appraising different study types have been published in JAMA and the BMJ as described above.

Screening questions

Screening questions [14] can help you to rapidly assess whether it is worth investing more time in analyzing the results. These will help determine the validity of the results. If the validity is compromised in a serious way, then it is best to devote time to search for research of better quality.

Screening question 1			
Did the trial address a clearly focused issue Hint. An issue can be focused in terms of: the population studied the intervention given the outcomes considered	**Yes**	**Can't tell**	**No**

A research report of a clinical trial should have a clear statement about the population studied, the intervention given, and the outcomes considered.

Screening question 2			
Was the assignment of patients to treatments randomized? Hint. Consider if this was done appropriately	**Yes**	**Can't tell**	**No**

The randomization procedure describes how the subjects were allocated to control or experimental intervention. Some methods are not truly random, e.g., alternating between control and experimental groups. Lack of randomization seriously compromises the validity of the results.

Screening question 3			
Were all the patients who entered the trial properly accounted for at its conclusion? Hint. Look for: the completion of follow-up whether patients were analyzed in the groups to which they were randomized	**Yes**	**Can't tell**	**No**

All trials lose patients to follow-up for a variety of reasons. It is important to account for these losses. Completeness of follow-up strengthens the results, and a threshold of 15% is generally accepted. It is important to know the reasons for withdrawal and to reflect whether these could affect the results. For example, if withdrawal is due to side effects this may make the intervention look better and if withdrawal is due to "feeling better" this would distort the results in the direction of the intervention being worse.

Appraisal questions

Appraisal questions [15] address key aspects of the research design that could possibly invalidate the results.

Appraisal question 4			
Were patients, health workers, and study personnel 'blind' to treatment Hint. This is not always possible but consider if it was possible-was every effort made to ensure blinding?	**Yes**	**Can't tell**	**No**

The gold standard of randomized controlled trials is when participants, health care professionals, and researchers are blind to the intervention. This is the best assurance that no intentional or unintentional bias occurs at any phase of the study.

Appraisal question 5			
Were the groups similar at the start of the trial Hint. Think of other factors that might affect the outcome, such as age, sex, social class, etc.	**Yes**	**Can't tell**	**No**

Usually there is a table comparing the baseline characteristics of the participants by intervention group. Evaluate if all relevant characteristics that could affect the outcome have been included.

Appraisal question 6			
Aside from the experimental interventions, were the groups treated equally? Hint. For example, were they reviewed at the same time intervals?	**Yes**	**Can't tell**	**No**

Check for any differences in treatment that could affect the outcomes. For example, using more pain medication or receiving more health care support.

Appraisal question 7			
How large was the treatment effect? Hint. What outcomes were measured?	**Yes**	**Can't tell**	**No**

Reliability of the treatment effect will depend on follow-up rate and completeness of the data collection. Different standard deviations relate to the spread of the data. For a positive effect on reducing the risk of an outcome there are some calculations (Table 2) that will give you a precision of the estimation of the intervention effect.

For example, does 6 months compared with 6 weeks of oral anticoagulant therapy reduce the incidence of recurrent venous thromboembolism in patients who had a first episode of deep venous thrombosis or pulmonary embolism? Follow-up was for 2 years [29].

		Relative risk reduction (RRR)	Absolute risk reduction (ARR)	Number needed to treat (NNT)
Control event Rate (CER)	Experimental Event rate (EER)	CER - EER CER	CER - EER	1/ARR
18.1%	9.5%	48%	8.6%	12

- Relative risk reduction (RRR) expresses the percentage reduction in events in the experimental group (experimental event rate – EER) compared with the control group (control event rate – CER). RRR=[(CER-EER)/CER]×100%.
- Absolute risk reduction (ARR) is the difference in risk between the control group (CER) and the experimental group (EER). ARR=CER-EER.
- Relative risk or risk ratio is the ratio of risk in the experimental group (EER) to the risk in the control group (CER). RR=EER/CER.
- Numbers needed to treat (NNT) is the inverse of the absolute risk reduction. NNT=1/ARR=1/CER-EER. The number needed to treat represents the number of patients you need to treat to prevent one additional bad outcome (death, stroke, etc.). For example, if a drug has an NNT of 10, it means you will need to treat 10 people to prevent 1 additional bad outcome. It is useful when comparing treatments as it takes into consideration the baseline risk of having the bad outcome. Other important issues when comparing treatments are the rate and severity of adverse effects and cost.

Appraisal question 8			
How precise was the estimate of the treatment effect? Hint. Look for confidence intervals	**Yes**	**Can't tell**	**No**

Checking for precision of the outcome data collection is important. A between-observer effect or within-observer effect can give you an idea of how precise the measurement tool was. Usually the smaller the confidence interval the more precise the estimation of effect is, unless other biases or confounders can justify the results.

Appraisal question 9			
Can the results be applied to the local population? Hint. Consider whether the patients covered by the trial are likely to be different from your population?	**Yes**	**Can't tell**	**No**

This includes making an assessment about those included in the trial, as well as those excluded and lost to follow-up.

Appraisal question 10			
Were all important outcomes considered? Hint. If not, does this affect the decision?	**Yes**	**Can't tell**	**No**

Important outcomes should be included in the hypothesis and secondary objectives. Analysis of outcomes not initially forecasted in the study protocol may suffer from bias.

Appraisal question 11			
Are the benefits worth the harms and the costs? Hint. If not, does this affect the decision?	**Yes**	**Can't tell**	**No**

This question involves an overall assessment of the intervention based on the validity of the data and strength of the evidence. Other "real life" considerations can be taken into account.

Study design

"The argument is not about the inherent value of different approaches and worthiness of investigators using them. The issue is which way of answering the specific question before us provides the most valid, useful answer."[30]

Sackett and Wennberg (1997)

Observational research

The researcher observes a population or group of subjects or manipulates data about those subjects. Data can be collected specifically for the purpose of the research or information collected by other agencies can be utilized, such as birth certificates, hospital records, etc. Examples of observational research design include case-control studies and cohorts.

Case-control studies

A case-control study can be used to investigate the relationship between exposure and outcome. If there is a strong association and it meets temporal requirements, i.e., the exposure occurs before the outcome, then causation can be hypothesized. The relationship between lung cancer and smoking [31] is a historical example. Further research is required to support the theory of causation. Case-control studies are also useful to identify adverse effects of treatment.

Quality indicators should be sought for when analyzing case-control studies. In a major review of case control studies, 35 different sources of bias were identified [32]. The selection of subjects and measurement of outcomes of interest were the

main problematic areas for introduction of bias. When appraising a case control study quality indicators include.

1. Do the selection criteria for control subjects match those of the case subjects in every respect except for the presence of disease or risk factor being studied?
2. Were the measurements on the control subjects free from bias?

Advantages of case-control studies include [33]: they are quick and cheap; they are the only feasible method for very rare disorders or those with a long time gap between exposure and outcome; fewer subjects are needed than in cross-sectional studies.

Disadvantages of case-control studies include [33]: reliance on recall or records to determine exposure status; possibility of confounders; selection of control groups is difficult; potential bias in recall and selection.

Cohort study

In a cohort study a group of people is investigated over a period of time and any changes that occur during that period are recorded. It can be prospective (data still to be collected) or retrospective (looking at data already collected). Data can be collected specifically for the study or it can be based on data collected routinely.

A cohort study design can be used to investigate research issues when it is not ethical to carry out a randomized controlled trial to determine the outcome of a treatment. The findings from this type of study enable the elaboration of quality indicators, e.g., readmission rates to be based on evidence. Experiments in the organization and delivery of health services, e.g., the relationship between volume of procedures carried out in a hospital and outcomes such as complication rates or mortality [34], are also ideal for the cohort design.

There are three study design features that are pivotal in the outcome of an appraisal of cohort studies. The first is the recruitment of subjects, which has to be evaluated for its completeness. It can be useful to ask about the subjects not recruited. Also assess the possibility of a selection bias before entry point. For example, the more-severe cases may be referred elsewhere. Secondly, the criteria for evaluating outcomes of care must be valid or validated, e.g., hospital mortality may vary due to variations in length of stay and therefore is not a valid measure. A validated measure would be 30-day mortality. Thirdly, when analyzing the results make sure that some measure of control for co morbidities is taken into account as this can systematically bias the results.

Advantages of cohort studies include [33]: they are ethically safe; subjects can be matched; the timing and directionality of events can be established; eligibility criteria and outcome assessments can be standardized; they are easier and cheaper to manage than a randomized controlled trial.

Disadvantages of cohort studios include [33]: controls may be difficult to identify; exposure may be linked to a hidden confounder; blinding is difficult; randomization not present; for rare disease, large sample sizes or long follow-up necessary.

Experimental research

An intervention is performed as a result of planning by the researcher. The most-representative design is the randomized controlled trial, considered the gold standard in the scientific community for determining causal effect.

Randomized controlled trial

The randomized controlled trial is the gold standard of scientific methodology for evaluating the effect of an intervention. It produces the highest level of evidence available. It is used to assess the effectiveness of an intervention, whether it be drug therapy (risk of myocardial infarction associated with antihypertensive drug therapy), methods of clinical/surgical management (management of patients with venous leg ulcers), or evaluating health services and administration (acute stroke intervention: special stroke unit versus general medical ward).

Advantages of randomized controlled trials include [33]: unbiased distribution of confounders; blinding is more likely; randomization facilitates statistical analysis. Disadvantages of randomized controlled trials include [33]: they are expensive in time and money; volunteer bias; they can be ethically problematic at times.

Clinical trials

A clinical trial is defined as a carefully and ethically planned experiment, which involves a group of subjects, an intervention, and outcomes. The objective is to answer a precisely framed question on the most-appropriate intervention for the given condition under study. The outcome of any clinical trial should be clear, truthful, and precise information, which is relevant to the treatment of future patients [35].

"A good clinical trial design provides the most credible outcomes data from evaluating therapeutic interventions in human subjects." [36]

Chambers and Fairbarn (1998)

Historical development

Lind described the first comparative trial in 1753, studying the most-promising treatments of scurvy in sailors during long voyages. He selected 12 patients with scurvy and proceeded with the following interventions. Two sailors each received either a quart of cider per day, 25 guts of elixir vitriol per day, 2 spoonfuls of vinegar per day, course of sea water per day, or 2 oranges and 1 lemon per day. At the end of 6 days, of the subjects who had oranges and lemon, one was fit for duty and the other was appointed to nurse the sick. Although the trial appeared conclusive, Lind continued to propose "pure dry air" as a first priority with fruits and vegetables as a secondary recommendation. Almost 50 years elapsed before the British Navy supplied lemon juice to its ships.

P.C.A. Louis established a scientific approach to clinical trials and epidemiology. In 1834 Louis discussed the need for: (1) the exact observation of patient outcome; (2) knowledge of the natural progress of untreated controls; (3) precise definition of disease prior to treatment, and (4) careful observation of deviations from intended treatment and he set the foundation for the use of the "numerical method" in assessing therapies. In 1835 he studied the value of bleeding as a treatment for pneumonia (78 cases), erysipelas (33 cases), and throat inflammation (23 cases), and found no demonstrable difference between patients bled and not bled. This finding contradicted current clinical practice in France at the time and instigated the eventual decline of bleeding as a standard therapy.

Greenwood and Yule (1915), in a review of anticholera and antityphoid studies, appear to be the first to suggest that some form of random allocation of patients to treatment is necessary to generate truly comparable treatment groups. Ferguson et al. (1927), in a study of vaccines for the common cold, may have been the first to introduce blinding. Their study was single blind in that research workers, but not the patients, knew who received saline or vaccine injections.

It is generally agreed that the first clinical trial with a properly randomized control group was of streptomycin in the treatment of pulmonary tuberculosis (Medical Research Council 1948). Randomization was made by a system of sealed envelopes. Two radiologists and one clinician who did not know treatment allocation or the others' evaluation assessed the patients' X-ray film independently. A randomized trial of the treatment of the common cold with antihistaminic drugs (Medical Research Council 1950) was notable for using a placebo control in a double-blind manner; neither patient nor investigator knew which treatment the patient was given: placebo or antihistamines [35].

Level of evidence

The randomized controlled trial provides the highest level of evidence for the evaluation of therapy available from a single study. The confidence interval of the results describes how precise these are by giving a 95% or 99% estimation interval as to where the true value lays, so the narrower the confidence interval the more precise the results and viceversa.

Drug development trials

Drug development trials within the pharmaceutical industry are often classified into four main phases of experimentation.

Phase I trials: clinical pharmacology and toxicity

The first experiments in humans are primarily concerned with drug safety and not efficacy. They are usually performed on human volunteers. Drug metabolism and bioavailability studies are conducted, followed by studies of multiple doses to determine appropriate dose schedules for use in phase II. Cohort study designs are generally used.

Phase II trials: initial clinical investigation for treatment effect

These are small-scale investigations into effectiveness and safety of a drug, and require close monitoring of each patient. Cohort study designs are most frequently used, although clinical trials can be implemented.

Phase III: full scale evaluation of treatment

After a drug is shown to be reasonably effective, it is necessary to compare it with the current standard treatment(s) for the same condition in a large clinical trial.

Phase IV: post-marketing surveillance

After a drug is approved for marketing, there remains a substantial enquiry still to be undertaken regarding monitoring for adverse effects and additional large-scale, long-term studies of morbidity and mortality. These studies can utilize case-control, cohort, or clinical trial study designs

Types of clinical trials

Explanatory trials

Subjects are carefully screened to exclude those who may suffer side effects, may not comply with treatment schedules, have other illnesses, or are taking other medications. These trials are high on exclusion and focus strongly on internal validity by optimizing conditions for testing interventions.

Efficiency trials (pragmatic trials)

Pragmatic trials are conducted to evaluate the intervention in a real life setting. For example, evaluating laparoscopic surgery against conventional surgery. The first trials are performed in reference centers developing the experimental procedure, the next step is to evaluate the effectiveness in a general hospital setting.

Crossover trials

This is a combination of the time series design and the randomized controlled trial design. Half the participants are randomly assigned to start with the placebo treatment and then switch to intervention, and the other half follows the other way around. This permits within-group (each subject's response to treatment against another) and between-group analyses. There may also be a need for a washing-out period during which the response is not being assessed. This approach has the advantage of needing smaller sample sizes.

Advantages of cross-over trials include [33]: they control confounding variables

and effectively double sample size; they are a good choice when subjects are few and difficult to recruit and the carry over effect is not likely to be a problem; all subjects receive treatment (at least some of the time); statistical tests assuming randomization can be used; blinding can be maintained.

Disadvantages of cross-over trials include [33]: all subjects receive placebo or alternative treatment at some point; the duration of the study is doubled and the analysis is more complex; the wash-out period lengthy or unknown; they cannot be used for treatments with permanent effects.

Run-in design

All patients are first put on placebo. After a few weeks those who comply are randomized. Two advantages of this design are: it allows researcher to check who is likely to comply or not and allows for any previous medication to get washed out of the system before baseline measurements are made. A variant is to use an active drug instead of a placebo; those who do not show side effects are randomized for treatment with a new related drug.

Factorial design

Two types of intervention are compared simultaneously. The subjects would first be randomized to intervention A and within each group to intervention B. For example: behavioral therapy.

Matched pair randomization

Subjects are selected in pairs matched by factors like age, sex, and severity of disease, and then randomly assigned to interventions.

Group randomization

Groups are randomized, e.g. factories, villages, rather than individuals. The study takes into account naturally occurring clusters.

Time series design

Serial measurements are performed sequentially during treatment and control periods. Each subject serves as its own control. Factors like age, sex, residence, social class, and similar others get eliminated as confounders. The sample size is doubled because each subject is both control and experimental. The disadvantage is that a carryover effect in the form of residual influence of the intervention can occur during the control period. To solve this problem, a repeated measure strategy may be used. This is where an intervention is repeatedly started and stopped. If the start and stop periods show matching patterns of outcome than the influences of confounding factors can be ruled out.

N-of-1 trials

Here the patient undergoes pairs of treatment periods (one period with active treatment and one with placebo or old treatment), assigned at random. Both patient and clinician are "blind" to allocation. The treatment effects are monitored, by designing quantitative measures of the patient's symptoms by means of diaries or questionnaires. Pairs of treatment are continued until effectiveness is proven or refuted. There should be at least three replications of alternative treatments. These are best suited to stable chronic conditions. The therapy being assessed should have a rapid onset of action, and cease to act soon after it is discontinued. The N-of-1 trial represents a more accurate test of efficacy of treatment compared to open uncontrolled trials.

Designing a clinical trial

Firstly define the purpose of the trial and the research question the trial aims to answer. The question should be refined to a clear and measurable hypothesis. A well-built question has three or four elements [30, 37]:

1. the problem - the patient, population or condition you are dealing with;
2. the intervention you are considering;
3. a comparison intervention (if appropriate and relevant) - an alternative intervention with which you would like to compare your intervention, i.e., placebo, standard treatment;
4. the outcome or outcomes you are interested in.

Problem	Intervention	Comparison	Outcome

For example:

A. Does high-frequency oscillatory intervention (intervention) compared with conventional ventilation (comparison) reduce mortality (outcome) of preterm infants with respiratory distress syndrome (population)?
B. In patients with small abdominal aortic aneurysms (population), does early elective surgery repair (intervention) reduce mortality (outcome) better than regular ultrasonographic surveillance of the aortic diameter (comparison)?

Framing a question is by no means a straightforward process, it demands rigor of thought. You must make careful judgements about what aspects of the problem you will focus on in your research. You must also be prepared to revisit your initial judgements, and refine or redirect your question. This process of fine-tuning of the research question is circular and iterative. This structure can be adapted to build questions about diagnosis, prognosis, efficiency, and cost-effectiveness.

Primary hypothesis and secondary objectives

Identify and test for a single primary hypothesis. Avoid multiple primary questions within a single study. Secondary objectives are defined in the protocol and usually are hypotheses generating only. Future studies are required to test the new hypothesis, due to limitations in sample size and study design.

Outcome measures

The classic outcome measures are morbidity and mortality. Other issues that have been increasingly appraised include health-related quality of life and cost analysis. Use valid outcome measures rather than surrogate outcomes, e.g., using heart ejection fraction as a predictor of heart failure.

Statistical significance

Define the kind of association of outcome between study groups. Is this study testing for statistical significance or for equivalence between groups?

Statistical power

Will difference between intervention group and control group be missed? The ability to avoid missing differences is called power of the study.

The written protocol

The protocol of a research study is a formal written document specifying the manner in which a trial is to be conducted. It is the blueprint of the study and will serve as a guideline throughout the implementation and analysis. Start by writing a precise definition of the research objective in terms of specific hypotheses regarding intervention efficacy and safety. The researcher must define exactly the type of patient, the interventions to be compared, and the methods of evaluating each patient's response to treatment, in a manner that is easily replicable and is resistant to drift during the course of the study.

The next step is to develop a detailed design for a randomized trial and document one's plan in a study protocol. This design needs to fulfill scientific, ethical, and organizational requirements, so that the trial may be conducted efficiently and according to plan, standardizing the procedures for everyone to follow. Two fundamental issues must be considered: (1) the trial must recruit enough patients to obtain a reasonably precise estimate of response on each intervention; (2) the selection, ancillary care, and evaluation of patients should not differ between interventions, so that the intervention comparison is not affected by factors unrelated to the interventions themselves. The purpose of the trial and the setting will determine the professionals who will provide input to the design of the study protocol. These can include subject area experts, epidemiologists, and statisticians.

Protocol components

Title page. The title page should include: study title, name of intervention, protocol number, site of study, conduct or protocol chair name and affiliation, sponsor, and final or draft date.

Table of contents. The table of contents lists the titles for the different sections and the page where they are located. Use the word processor facility by choosing the "style" for the titles, i.e., heading 1, heading 2, etc. and generate the table of contents.

Summary. The protocol summary should contain the following topics: background on study area and intervention, objectives, subject eligibility criteria, intervention(s), study design, measurements, and outcomes.

Introduction. Background information on intervention and options available should be described and a summary of clinical studies relevant to the hypothesis included. Justify and support your research question.

Hypothesis. In clinical trials the hypothesis is a statement of the expected outcome of the study. It is a clear and interpretable answer to the research question. For example, "drug A will have a greater effect on blood pressure than drug B".

Objectives. The objectives should be directly derived from the hypothesis, which is being tested. For example, "to compare the effect of drug A with placebo on systolic and diastolic blood pressure with moderate hypertension". Single primary objectives should be stated that relate directly to the primary efficacy variables. Secondary objectives are sometimes included.

Subject eligibility. The objective of defining eligibility criteria is to narrow the variability of intervention responses, therefore allowing trial outcomes to be stated with confidence. Inclusion and exclusion criteria should be clearly defined. Characteristics frequently taken into consideration include: age, sex, weight, function of major systems, previous health, and drug treatment. Focus on criteria related to disease or condition under study.

Study design. A statement with a summary of design elements should open the study design description. Use standard terminology when labelling study design, i.e., randomized, controlled, blinded, etc. Specify the major elements of study design such as the subjects, intervention, and outcome, as well as the temporal aspects relevant to the intervention and evaluation of outcomes. For example: "this protocol describes a randomized, double blind, placebo-controlled study of the safety and efficacy of drug A administered to human immunodeficiency virus – (HIV) infected subjects with HIV RNA levels<50,000 copies/ml with follow up for 1 year."

Sample selection. One key goal of clinical trials is to be able to make inferences from the results of a study to a more general population. The characteristics of patients included in a trial will determine the extent to which the conclusions can be generalized to other patient populations. Factors like age, sex, socioeconomic status, education, and health consciousness can all affect subsequent morbidity and mortality. Results should be applicable to the reference population for which the research question is being posed, i.e., generalizable or externally valid.

Clearly define the condition of interest to select subjects. If a disease, how accurately can a diagnosis be made? Are issues relating to staging the severity of illness clearly described? When the clinical course of a disease is extremely variable, and also varies from one patient to another, assessing treatment effects by clinical observation can be extremely unreliable, e.g., systemic lupus erythematosus. Take into consideration the natural history of the condition of interest and the existence of validated prognostic markers to support the effect of intervention on the outcome variables. Patients who refuse to participate can be systematically different (by social class, severity of disease, or other problems) from those who agree to enter the trial. An account of subject refusals needs to be kept, so a comparison can be made between those who consented and those who did not in the final study sample.

Sample size. The sample size should be determined by a statistician rather than relying on clinical experience or judgment. How many patients need to be enrolled in the trial and how many subjects are required to complete the study protocol in order to generate meaningful results? Important considerations in calculating sample size for clinical trials are the outcome measures. If the outcome variable is a continuous one, then mean and standard deviation are needed for calculating the sample size. These can be obtained from a pilot study or from the literature. In the case of a discrete variable, one needs an estimate of the event rate.

The statistical power of the study will define how certain one wants to be of detecting a significant difference or of accepting the equivalence of the interventions. Generally a power of 80%-90% is regarded as adequate. This translates into having a 10%-20% chance of not detecting a significant difference between interventions, although it may in reality exist (type II error).

How frequently is the outcome expected to occur in the control group and what variations can you anticipate? How small a difference does one consider important and clinically significant? Clinical judgment is valuable in estimating the magnitude of difference between the two treatment groups in the key outcome measures of a trial. The clinically important outcomes would encourage physicians to use the intervention in other patients.

Once you determine the appropriate sample size to detect the clinically important difference, calculate an extra percentage to allow for dropouts or withdrawals. It is almost inevitable that some subjects will drop-out or not comply with the study intervention and others will be lost to follow-up. As a general rule, a loss of up to 10%-15% of subjects is acceptable, and should not bias the result unless there is a strong confounder.

In summary, the choice of the sample size is based on: desired power, magni-

tude of difference to be detected, and magnitude of the parameter in the control group.

Inclusion and exclusion criteria. Have a clear definition of the inclusion and exclusion criteria before recruitment starts. The broader the inclusion criteria, the easier it is to select study subjects and the findings are more generalizable. When defining exclusion criteria make an attempt to exclude extraneous conditions that could distort the outcome, such as subjects that possess contraindications for intervention or who are unlikely to comply with the intervention.

Recruitment. Recruitment procedure is decided with two goals in mind: 1) to recruit a sample that adequately represents your target population and 2) to recruit enough subjects to meet sample size required. If you are recruiting from a clinic population, this limits the generalizability of your results to the community. The response rate will also influence the ability to generalize to a wider population. Improving recruitment methods may improve the response rate, but it is also important to keep track of non-responders. This will enable future comparison and support the interpretation of the results of the study. Subjects will be more motivated to participate with a design that avoids invasive and uncomfortable tests. Language barriers can be overcome by using bilingual staff [38].

The inability to enrol sufficient subjects is a frequent problem in clinical research. The accessible population, from which you are recruiting your study sample, should have a much greater number of probable subjects than you require in your study sample. There is always a tendency to overestimate the number of subjects meeting the study criteria that will agree to participate. Develop contingency plans if recruitment efforts fail in an initial effort [38]. The procedure for handling withdrawals and replacements and other provisions for deviations from the protocol are established.

Measurement of the baseline variables. The purpose of measuring baseline variables is to define and describe the characteristics of the study cohort. This establishes that the characteristics in the intervention and comparison groups are similar before the intervention. It is also important to measure predictors of outcome, i.e., risk factors or existing conditions that can influence the outcome. The baseline measurements allow for the statistical adjustment of the results to reduce the effects of chance maldistribution in baseline factors.

Intervention. A detailed description of the intervention should include: randomization and intervention assignment strategy, schedule of study intervention (e.g., drug administration – "subject will be randomly assigned to take the study drug at dose level 1 or 2, or placebo, to be taken three times daily following meals"), dispensation (drugs), interval of intervention, supplies (e.g., bottles with capsules), labelling, and duration. A clear account of the management of forecasted situations can include: handling dose reduction, changes allowed in background therapy (e.g., use of other chemotherapeutic agents), and temporary study discontinuation.

Several types of control can be used in clinical trials: concurrent placebo, con-

current active drug, no treatment, different dose of the same drug, and concurrent use of standard care. Historical comparison of data can be obtained from the same patients on no therapy, the same therapy or different therapy, or other patients on no or some different therapy.

The presence of competing interventions needs to be ascertained, many times subjects use concurrent interventions inadvertently and this can go unrecognized amongst study personnel. In the protocol describe what kind of concurrent therapy, i.e., medications, physiotherapy, etc., will be accepted and which shall result in exclusion of the subject. This can happen before as well as during the study, so it is important to elicit such information throughout the study period.

Compliance bias occurs when there are differences in the interventions being compared and consequently there are differences in patient adherence to treatment (e.g., difficult diet, bad taste). The Hawthorne effect is the tendency for people to modify their behavior because they are receiving special interest and attention. The subjects can make changes in habits or life-style in a manner that reduces undesirable outcome, although unrelated to the intervention. The two most-important elements in eliminating bias in clinical trials are randomization and blinding.

Randomization procedure. The goal of randomization is to maximize the probability that groups receiving differing interventions will be comparable. When an adequate randomization procedure is used, the potential for bias in allocation to study groups is removed. The assignment of subjects to the intervention is performed in a manner determined by chance alone. Study groups tend to be comparable with respect to all potentially confounding variables, as well as other unsuspected confounders. However there is no guarantee that differences will not arise by chance between the two groups. This risk is much greater with smaller samples.

The randomization method should provide true random allocation. Computer-generated random numbers and tables of random numbers are the most commonly used methods of randomization. Alternate allocation and allocation by birth date are not true randomization procedures. Allocation concealment is an important aspect of the randomization process, usually achieved by using sequentially numbered, sealed, opaque envelopes.

Central randomization procedures are preferable to local randomization; i.e., carried out by a statistician or a researcher not involved in the recruitment and follow-up of subjects. A blocking procedure can be used when there is a need to balance the number of subjects in different strata of subject characteristics. Subjects are first classified with respect to variables that will affect the outcome (e.g., age, sex, severity of disease) and then randomized for allocation.

Blinding. Blinding refers to a lack of knowledge of the identity of the intervention. The aim is to avoid bias in the trial execution and in the interpretation of the results. Investigators, clinicians, and subjects are prone to change their behavior if they know who is receiving what treatment. Blinding is about making sure that the several participants are unaware of which treatment is being given to whom, so they do not change their behavior. Blinding is especially necessary when the outcome is easily influenced by the knowledge of the treatment received, e.g.,

pain, nausea, disability. Hard outcomes like death or recurrence are not affected as much.

Ethical issues and practicality need to be taken into consideration when deciding the level of blinding in a clinical trial. Increasing the level of blindness brings higher costs, increased complexities, and longer timelines for trials. Blinding can be described as single, double, or triple, depending on who is unaware of the intervention given.

In a *single-blind* trial patients are unaware of the treatment they are receiving, but the physician evaluating the outcome is informed. In a *double-blind* trial health professionals providing care and subjects are not aware of the intervention being given, whether experimental or control. Those who assess the outcome are not aware of the intervention given. This is the most commonly used blind design. In a *triple-blind* trial health professionals, subjects, and research personnel are unaware of the intervention being given. For example, those allocating patients to the intervention groups do not know which subjects are receiving the experimental or control intervention. Describe clearly in your study protocol whether you will be using a blind or open-label design. Open-label refers to the fact that the type of intervention is known to subjects, investigators, and outcome evaluators.

Variables. Predictor variables are those that will be measured against the outcome variables to evaluate the presence of a relationship. In clinical trials the predictor variables are the experimental intervention and the control intervention.

Outcome variables are those that you have decided will best reflect the effectiveness of the intervention. All the most-important possible outcomes should be included and carefully defined. If some or all of the important parameters cannot be measured, it may not be ethically justifiable to conduct the trial. If possible use measures that have been validated in the literature. Include in your outcome variable data to be collected related to the safety of the intervention.

Continuous variables are those that can be ranked on a spectrum of quantifiable intervals, e.g., body weight and blood pressure. Categorical variables those derived from characteristics that cannot be measured, e.g., sex and blood type. Categorical variables can provide more-informative statistics (mean and standard deviation) than categorical variables (rates, counts, and proportions), adding power to the study and therefore reducing the required sample size [38].

If a new intervention is found to be particularly successful or obviously harmful, then an early termination of the trial and publication of the results is mandatory. Sometimes it is helpful to have an independent monitoring group separate from the investigators, particularly in drug development trials.

Measurement of variables. When deciding on the measurement to assess variables of interest take into account precision and accuracy of the instrument. Precision relates to obtaining the same or very similar results after repeated measures. This can be affected by observer variability (e.g., skill in using an instrument, measuring blood pressure), subject variability (e.g., mood when responding to a questionnaire), and instrument variability (e.g., hearing screening tests affected by background noise). Action can be taken to enhance precision, such as standardiz-

ing the measurement methods, training the observers, refining instruments, automating the instruments and repeating measurements, and using mean values [38].

Accuracy translates if the measurement reflects what it is intended to. Systematic error is reflected by less accuracy. Assess accuracy by measuring a "gold standard" or by comparing with a technique that has had its accuracy accepted. Strategies for enhancing accuracy include standardizing the measurements methods, training the observers, refining the instruments, automating the instruments, developing measures that subjects are not aware of, blinding, and calibrating the instrument [38].

Follow-up. One common pitfall is if the follow-up period is too short and may miss the intervention effect. Describe your strategy to minimize the incidence of drop-outs and withdrawals. Phone calls and house visits are methods to follow subjects that are not complying with the study procedures. Record the reasons why the subject withdrew from the study.

If loss to follow-up is related to the intervention it will affect the outcome, consequently leading to under- or over-estimation of effects of the intervention. The potential bias resulting from drop-outs or withdrawals should be carefully considered. A loss to follow-up in excess of 20% causes serious concern about the validity of the trial. There are strategies that can be used when analyzing the data to control for loss to follow-up, i.e., intention to treat analysis

Adverse events. Thought should be given to how to deal with complications arising from the disease or the intervention and how to respond to new problems. Describe actions to be taken in forecasted scenarios.

Study procedures. This section should provide a detailed description of the organization of the events necessary for the implementation of the trial: (1) screening evaluations and timing; (2) enrolment and randomization; (3) where to send screening samples (if applicable); (4) procedure for enrolling/randomizing subjects; (5) baseline evaluations (6) intervention evaluations (7) discontinuation of evaluation of intervention.

Include a timeline for the events described, allowing time in each step for possible setbacks and unexpected incidents that could delay implementation of the study scheduled procedures.

Follow up evaluations. Case report forms are created to permit accurate and systematic collection of individual subject data. The design of the case report form is extremely important, so they can effectively collect all data required to: (1) meet study objectives; (2) permit investigators to easily and clearly record data; (3) facilitate investigators adherence to the protocol; (4) allow rapid and effective review of the study by monitors; (5) facilitate database generation.

Safety. For clinical trials studying new interventions, active safety actions should be predicted and implemented, such as adverse event reporting. A definition of

adverse event and of serious adverse event is described. Procedures for recording adverse events and requirements for reporting serious adverse events are laid out. The medical monitor contact information (typically a medical doctor familiar with therapeutic area and safety profile of the study drug) is included.

Statistical specification. A statement of the decision of the analysis type to be used when evaluating the primary and secondary outcome variables is included in the protocol. The methods of analysis include population definitions (e.g., intention-to-treat), and if interim analysis is planned, the stopping rule(s) to apply to the analyses. The handling of specific situations is described in advance, i.e., how drop-outs will be dealt with, subgroup analysis, and the inclusion or exclusion of data from ineligible subjects [36].

Interim analyses are conducted at a pre-determined time in the trial. They are most frequently conducted to assure the sponsor that there is no evidence of efficacy, or to assess trends in incidence of adverse events. A decision is made in advance on a strategy for stopping a study early if sufficiently strong evidence of an intervention difference becomes evident in the interim analyses. Ethical issues and cost are the main justifications for planning interim analyses.

References and appendices. Include all references used to support the background and methodology sections. All documents and forms relevant to the implementation of the study are attached in the appendices. These include [39]:

1. schedules of assessments;
2. model informed consent form;
3. study procedural requirements (source documentation, case report forms, data collection and storage methods, study monitoring, study medication accountability, conditions for modifying or terminating the study, disclosure of data/publication and conflict of interest);
4. product information: dose form (e.g., tablets/capsules, active and inactive ingredients, description of bottle in which provided, number in each bottle), rules for sub-dispensing, dose strength(s), shipping, storing and handling, temperature control of product, special handling considerations, return/destruction of used or unused product.

Conducting the trial

Good organization. Before committing time and resources to a study, ensure the feasibility of the trial by making sure there is a written commitment of cooperation from all stakeholders in the clinical trial. A consensus must exist that it is likely that an adequate number of subjects can be enrolled within the specified time period and that the interventions and outcome measures are feasible.

Analysis of the data

Statistical analysis of the data will allow us to define how likely we are to erroneously conclude there is a treatment effect, or to miss a treatment effect that is

there. It provides a mean of making sense out of complex data that in turn enables us to answer important questions.

Type I and type II errors. Type I error (α) is the probability of declaring a statistical difference when in fact there is none. You falsely reject the null hypothesis. The usual risk accepted is 0.05. For example for an α=0.05, the probability of a false-positive result is not greater than 0.05 or 5%. α is the pre-determined level of statistical significance for the study.

Type II error (β) is the probability of not detecting the difference that is looked for when it is present or the chance of missing the real effect. You falsely accept the null hypothesis. The usual risk accepted is a β of 0.1-0.2.

Power (1-β) is the probability of detecting the difference that is looked for when it is present. It is desirable to have a high power or probability of detecting a difference between treatments to justify the efforts and ethical issues involved in conducting a clinical trial.

The investigator usually determines the magnitude of these errors before a clinical trial begins, often in collaboration with a statistician. The goal of any trial is to have a small type I and small type II errors, and this will reflect an increase in sample size.

Descriptive statistical analyses. Summarize the data using non-comparative techniques such as frequency distributions (proportions), description of averages (mean, mode, and/or median), and description of the spread of the values (standard deviation and standard error). Assess assumption for statistical tests, i.e., if the data have a normal distribution or not.

Analytic statistical analyses. Test your hypotheses. Variables are compared to measure the pattern and strength of relationships. The adequate test will depend on a variety of factors (e.g., existence of normal distribution, between-within-subject variation, etc.) and circumstances (e.g., time-line) surrounding your data. Consult a statistician to plan the appropriate methodology for analysis in advance. Hypotheses can be proven wrong, but they can never be proven correct because the investigator cannot test all the potential subjects with the condition of interest.

Analytic: intent-to-treat. An intention-to-treat analysis has as principle the inclusion of all randomized patients in the analysis. Outcome is analyzed by the intervention originally assigned to the subject (including the drop-outs). It is the preferred approach for pragmatic studies. If one starts excluding patients from the analysis for a variety of apparently "justifiable" reasons (subjects later not found to have met criteria, non-compliers, missed study visits, moved, had other illnesses, or dropped out of the trial before completion), this often leads to excluding a large percentage of patients from analysis. The intention-to-treat analyses eliminate the risk of creating a bias due to these exclusions. Differences between treatments tend to be obscured and sample size needs to be increased to compensate.

Analytic: as-treated. In the as-treated analyses the groups compared have been

determined by an algorithm based on a degree to which the patient "complied" with the protocol e.g., if the subject took all the study medications as defined in the protocol. The sample size used to analyze the data will be reduced because not all subjects will meet compliance criteria, compromising the power of the study. Bias can be created, as the subjects that did not comply (e.g., suffered adverse effects) may be different from subjects that did comply. The validity of the statistical testing is undercut, and the possibility of imbalance is created. It is advisable to perform an intention-to-treat analysis as well, so both types of analysis support any difference.

Correlation coefficient. Correlation coefficient is the degree of association between two variables. The P value is the likelihood of an error in the conclusion that two variables are associated. A high correlation coefficient and low P value implies clinical importance. When data are clinically significant but express a low correlation coefficient, reasons might include a relationship between intervention and outcome that is not linear but curved, or a few pairs of observations may vary from the rest. These can be controlled for in the statistical analysis.

Number of strata. Data may be evaluated with subgroup analyses. The subgroups should always be decided in the protocol with the hypothesis on how different subgroups might react differently to the interventions being studied. There probably will not be enough power, as sample size decreases, to answer many questions. If performed after the data are collected this is designated post hoc analysis and can be helpful only for generating hypotheses for future studies [40].

Subject characteristics are frequently used for subgroup analyses such as: age, sex, race, severity of disease at baseline, or duration of therapy. Within a subgroup, study data can be further divided into strata, e.g., <65 years of age versus 65 or older. The more strata used, the fewer the subjects in each stratum, which results in less power to detect differences.

Interim analyses. Interim analysis is conducted at a pre-determined time in the trial. It is most frequently conducted to assure the sponsor that there is no evidence of efficacy, or to assess trends in incidence of adverse events.

Drawing conclusions

Interpreting clinical trial results. Interpretation denotes the process of discerning the clinical meaning or importance of or providing an explanation for data that is being evaluated. Analyses of data are primarily statistical exercises, whereas interpretation of data is primarily a clinical exercise. The investigator is primarily concerned with patient treatment implication plus issues related to publications and grants. In drug development trials, the sponsor's primary concern will be discerning the clinical significance of the data for future trials.

The goals of interpretation include:

1. to establish the relevance of the data;
2. to report the results of the original objectives of the trial;

3. to compare data from the trial with data obtained from other trials;
4. to develop a hypothesis that may be evaluated in future trials and to extrapolate the data to particular patients, environments, and types of treatments;
5. to assess the clinical significance of the data for diagnosis, treatment, or prevention of disease in patients studied or in other patients groups;
6. possibly to gain insight into interpretation of the characteristics of the disease itself.

Interpretation begins only after the trial is completed. The trial is completed when the data have been collected, edited, and entered into a database, the appropriate statistical tests have been performed to analyze the data, and when the statistical report providing a permanent and detailed record of the entire trial has been written. This report allows another investigator to repeat the trial using identical design or re-examination.

Statistical versus clinical significance. Statistical significance, on its own, does not provide information about whether an intervention is important for patients. In deciding whether an intervention makes a clinically important difference, one must evaluate the importance of the clinical outcomes, the size of treatment effect, the extent to which the benefit is offset by adverse events (risk-benefit), and also if the resources required to provide the treatment are affordable in comparison with other use of resources (cost-benefit) [41].

Overall outcome of results. At the most general level, the outcome of clinical trial results may be viewed as:

1. positive, or proven to do more good than harm;
2. negative, or proven to do more harm than good
3. inconclusive, or of unproven effect. For example, when multiple endpoints of efficacy or safety are used and some yield positive results and other negative results.

Sources of variability. When interpreting the analyses one should look for factors that may bias the data or affect their interpretation. These include characteristics of the clinical trial design, patients, the condition being evaluated, the study intervention, the investigator and staff, and the trial environment. Other factors that can be responsible for bias are the conduct of the clinical trial, the characteristics of the tests and measures used, statistical factors, characteristics of the data report, and errors committed accidentally or willfully before, during, or after the trial.

Ten-step approach to data interpretation [42].

1. Compare the experimental intervention group with the control intervention group to assess differences in baseline characteristics such as demographic data and prognostic factors;
2. evaluate whether the data affirm or refute the primary hypothesis of the research protocol;
3. evaluate whether the data affirm or refute the secondary objectives of the research protocol;

4. consider all the specific factors that may have influenced or biased the data and the interpretation reached;
5. discuss the interpretation(s) with a statistician and other colleagues to determine whether any additional analyses or subgroup analyses should be conducted;
6. adopt the "devil's advocate" perspective and strongly criticize your interpretation. List each criticism and determine how it may be refuted or addressed;
7. compare the interpretations and data with those obtained using standard treatments. The information on standard treatments may be obtained from data collected in the study or from results in the published literature;
8. discuss the interpretation of the results with others and request their critique;
9. extrapolate the interpretations to different patient populations, settings, or levels of organization as may be justified. Identify limitations to generalizability of the results;
10. develop new hypotheses or research questions to be evaluated in future studies.

Ethical considerations

The basic ethical principles of autonomy (respect for human subjects), non-maleficence (do not harm), beneficence (do good), and justice (exclusion) underlie the ethical considerations when developing and implementing a study protocol. Specific clinical trial ethical concerns relate to scientific validity, recruitment procedure, participation procedures, potential harms and benefits, and informed consent [43].

Ethical considerations are important in all aspects of a study, from the development of the research question, decision on the study design through to analysis, interpretation, and publication of the results. Ethical questions are not merely present or absent, but range from the obvious to the strongly debated. Ethical issues in clinical trials concern *individuals* rather than society at large. The objective is that no individual should be exposed to unreasonable risk for the sake of the community. The ethical basis for conducting clinical trials is that there is clearly doubt as to which of two or more interventions is more effective for a given condition. The benefits should outweigh the harm or risks of the given intervention.

Ethical considerations are relevant to all individuals and groups connected with the study. Patients, or parent/legal guardian, have to be informed in an acceptable and understandable way of the risks and benefits relative to receiving the intervention. They should freely provide the informed consent. The investigators have the obligation to make sure the informed consent procedure is followed in a way that subjects have the liberty to ask questions, and that throughout the study their information needs are met and support is given to those who decide to withdraw consent during the study.

The staff involved in carrying out the study procedure have to secure subject confidentiality and comply strictly with protocol procedures. Records should have an identifier, usually a subject number, from which the subject cannot be traced. Transmission of data over the Internet should meet strict standards to

assure that confidentiality is not broken. The sponsoring institution has to meet ethical standards through securing a proper trial design, the investigator selection, site monitoring, and data handling. The institutional review board is a multidisciplinary committee that has as an aim to protect patients' rights and make sure ethical standards are met by the protocol. They review the protocol and any subsequent changes or queries, providing their approval and recommendations. They do not, however, inspect to see if the study procedures described are being adequately applied. This responsibility lies with the principal investigator.

Approval by an institutional review board

The protocol must be submitted to and approved by a properly constituted institutional review board. Include time for review and approval by the review board when deciding on your timeline. Meet all requirements beforehand when writing up the protocol. There are general guidelines to help the investigator meet the requirements. All amendments made during the study period also require review board approval. It is the responsibility of the sponsor to ensure that the institutional review board is properly constituted and to conduct research in accordance with the guidelines.

Informed consent form

The institutional review board will determine the need for an informed consent form and approve the content. It should be obtained from all subjects entering the study and retained for future audit. The informed consent form is written in non-technical language to facilitate subjects' understanding of the potential risks and benefits of entering the study. A checklist for the content of the informed consent form [38] includes the following.

1. Research project - have an explicit statement that the project involves research. Identify the investigators. Describe the purposes of the research and procedure for the selection of subjects.
2. Study procedure - describe the time and commitment required by the subjects, the concept of randomization and blinding, and the assignment of experimental and control intervention. Explain clearly about the probability and magnitude of benefits and harms, mentioning procedures to maximize benefits and minimize harms. Describe alternative treatments, if they are available outside the scope of the study. Inform about potential costs and about the possibility of disclosure of the results. Include statement about procedures to maintain confidentiality.
3. Subject rights - assure that participation is voluntary and that declining to participate or withdrawing at any time during the study will not incur penalties. Offer explicitly to answer questions or provide further information. Clear information about whom to contact for further information, the rights of study subjects, and any adversities related to the study should be stated.

Confidentiality

All study information regarding the subject's state of health is confidential. If disclosure to a third party is required consent must be obtained beforehand.

Compensation

The sponsors are to provide compensation, regardless of legal liability, if the subject suffers deterioration in health or well-being caused by participation in the study. The trial subjects are usually not paid, unless it can be demonstrated that it is not an inducement. Only out of pocket expenses may be reimbursed.

Early trial discontinuation

A clinical trial can be interrupted for a variety of reasons. Problems related to the implementation of the study protocol include impossibility of recruiting an adequate number of subjects or failure of research staff to comply with study procedures. Interim analysis of the results can result in early termination, either because of harmful effects in the experimental group that could possibly be linked to the intervention or because the beneficial effect of the intervention raises ethical concerns for patients not receiving this intervention.

Writing up your scientific paper

A scientific paper is a written and published report describing original research results. However, it should be stressed that a scientific paper must be written in a certain way and be published under strict editorial practice and scientific ethics. An acceptable primary scientific publication must be the first disclosure of original research results containing sufficient information to enable peers: (1) to assess observations; (2) to repeat experiments, and (3) to evaluate intellectual processes.

Writing a research paper

There is no single best way to write a research paper. There are some general rules that should be followed if the production of an acceptable publication is intended, but the whole sequence of events varies from paper to paper. In general authors wait until data analysis is concluded and the results interpreted before starting to write. However, background reading should be a continued activity. It should not be too extensive: some novice authors try to search all the literature and to read all the papers pertinent to the scope of his/her paper. They end up with a huge pile of documents, and are not able to consolidate ideas. As a result there may be hundreds of references to be quoted. When reading, one should make notes, make notes of notes, write down sentences or parts of them. However, reading is not supposed to occur during the time allocated for writing.

Whom writing for?

The paper should persuade the referees of its acceptability. Thus, usually it is necessary to put aside personal gratification, favoring the technical aspects involved in the process of writing an acceptable scientific paper. Before actually writing the first version of the manuscript, the periodical where the paper is to be published must be chosen. Should it be sent to a general or to a subspecialty journal? This is a crucial question and the pros and cons should be weighted carefully by the author.

Sequence

After choosing the journal, a realistic schedule for the progress of the work has to be fixed. A moderate approach represents the ideal. The adherence to the schedule promises a good outcome. Figures, tracings, and tables lead the way, directing the writing of the Methods and Results sections. Discussion, Introduction, and Abstract follow in a very disciplined way.

Structure

Biggest problems come first; after they are dealt with the lesser ones are tackled. Subheadings represent signposts to the reader. In early drafts of the manuscript, subheadings are useful in every paragraph, but in the final version a subheading can never be used over a single paragraph. A scaffold of headings and subheadings makes the understanding of the paper difficult; in fact three levels of subheadings are more than enough for typographical distinction. Each section of the paper (Introduction, Methods, Results, Discussion, Abstract, and References) occupies a separate sheet of paper. The first draft of the manuscript should be written in a telegraphic style, ideas going to the paper in random order.

Writing tools

These are the essential companions for a good writer: a well-thumbed dictionary and a Thesaurus, so that every word, synonyms, and shades of meaning can be checked out carefully. A paper bin can easily accommodate undesired bits of text, and a good use of it is strongly recommended. Pencil, pen, or computer: whichever makes life easier represents the best choice.

Methods section

The Methods section must contain enough information for an experienced investigator to repeat the work being reported. In this line, tiresome detail is to be avoided definitively. Previous work can always provide some bits of text to undergo the process of "cut-and-paste". Methods is the first section of the paper in which subheadings should be used.

Results section

The Results section should be very concise. In order to accomplish this goal always: (1) refer to data (e.g., Fig. X, Table Y), (2) do not repeat numbers listed in Tables in the text, (3) if the text contains a lot of numbers, a Table should be drawn up instead, (4) data from Figures can be given in the manuscript if precision is desired.

Introduction

A well-written Introduction will present two or three paragraphs, no more. The first paragraph introduces the broad area of interest, whereas the second explains the rationale for the work and formulates the hypothesis.

Discussion

In the first paragraph of the Discussion the major findings are usually stated; if necessary paraphrase the Abstract. Each middle paragraph discusses a major result. The last paragragh usually starts with "In summary..." or "In conclusion...". In the former case, two or three sentences will suffice; in the latter the most important message is conveyed, speculation and "need more work" should be avoided, and part of the Introduction can be used. Always focus on your results, never discussing prior work without reference to your own study. Refer to Tables and Figures whenever necessary to clarify and shorten the text.

Abstract

A good Abstract is written in précis format. It must be informative, not descriptive. Some numbers may used, but not in excess. The author must realize that the Abstract is one important determinant of whether the paper will be read, and that abstracts are distributed freely in databases.

Title

The title should maximize information in the least words possible: less than 12 words or 100 characters will do. The title represents the label of one's paper. It should almost never contain abbreviations, but results can be stated. In the form of a question it may convey more impact to the reader.

Figures

Figures should be drawn before writing begins. An experienced writer keeps his/her eyes on the figures while preparing the text. One must bear in mind some simple rules of thumb in order to produce a nice Figure: (1) redraw, redraw, prune clutter; (2) reduce as much as possible non-data-ink; (3) do not use more than four lines, all solid; (4) minimize tick marks and do not number each one; (5) lettering should be uniform, lower case, avoiding bold type, and after reduction it stands 2-

3 mm high; (6) no caption is recommended; (7) reduce the Figure to one column in the journal where it will be published (use a reduced photocopy to check it out); and (8) the original Figure should be smaller than 3 times the final one.

Tables

A Table should be a single unit, understood without referring to the text. Prune, prune, and prune columns and rows: if the Table exceeds one sheet of paper, it should be redrawn. A narrow or broad Table may look better if rotated 90°. Vertical or horizontal lines are to be avoided. Finally, if the Table is too small the data should be moved to the text.

Of writing

Writing should occur at assigned times; during this period do not read (reading comes up at its own assigned occasion). Do not wait for the muses, work hard, writing is a craft not an art. Ideas will come while writing, which must be practiced. Reading good writers, especially non-medical, can help develop your writing skills. Remember that any craft is learned by imitation.

Momentum

Experience advises that a schedule for writing must be fixed beforehand, and monitoring progress may prove really important. Imagine that you write one page a week: can there be a better means of torture? Skip the trouble spots, leave them for later on. Obey your biological clock: some people feel refreshed in the early hours of the day, while others light up in the evening.

Concentration

A good writing session requires a stretch of several hours. When there is a shortage of time something else can be done instead: preparing, revising, etc. Avoid distractions: put away telephone, beeper, and everything else that can interrupt the session. Some people may feel "inspired" by the ocean or a field of corn, but a great deal of writers choose a dull corner to reach and keep the deepest concentration.

First draft

The first draft should be written as quickly as possible, as if the author were thinking out loud. He/she should get everything down, ignoring spelling, grammar, and style. Troublesome words should be skipped.

Good writing

Good writing results from accuracy, clarity, precision, logic, and order of presenta-

tion. The text must be clear, exact (ambiguity, inconsistency, and wooly words should be eliminated), and concise (least words, short words, one word versus many). Always simplify the text, as in the following examples:
- a majority of = most
- at the present time = now
- give rise to = cause
- in some cases = sometimes
- is defined as = is
- on the basis of = by
- pooled together = pooled
- subsequent to = after
- with the result that = so that.

Use and misuse of English

Tense: previously published work is an accepted piece of information, thus it is described and discussed in the present tense. Your own work represents something that was done but is not yet part of the overall "truth", so refer to it in the past tense.

Voice: the active voice is more precise and less wordy than the passive. Name the agent, even "I" or "we".

Singulars and plurals may be very tricky. Take special care with words in Latin.

Bad writing

Bad writing may be suspected when words do not do justice to your ideas. It has to be avoided because if multiple mistakes in spelling and syntax are present the reviewer suspects that similar sloppiness occurred in the laboratory.

Style

The author seeks a clear, orderly presentation that reads comfortably. Writing must be straightforward, linear, and scientific. Science is not literature.

Writing

During writing a great deal of rewriting takes place. Reshape, refine, and tighten the manuscript. Juggle words, change sentences around, keep trying. Use the paper bin frequently. Strengthen transition between sentences, so that no break in the narrative flow is apparent. After several drafts ask for a second opinion, if possible from an experienced researcher that does not belong to your field of interest. If the text is clear to him/her, it is probably fine.

Another important point is to avoid clutter. All first drafts have too many words. In successive drafts vigorous pruning is highly recommended, every sentence will be stripped. Excessive use of adverbs and adjectives contribute to clutter. In fact, writing improves in proportion to deletion of unnecessary words. Abbreviations

and acronyms are liked by authors but hated by readers. Reading should not require a glossary. The use of unwieldy words should be restricted.

Convey only one idea in a sentence, which must be kept short: less than 20 words. The length of the sentences must vary along the text to keep it fluid. Remember that in longer sentences there is a greater risk of grammatical error. Avoid them. The paragraph represents the unit of length in a group of sentences. Normally it looks like a not-too-long solid block of printing.

Your paper should sound like telling a story that the reader follows from start to end. Writing is sequential and logic is the glue. The reader feels like every step is inevitable, if you use smooth transitions between sentences and paragraphs in your text.

Rewriting

The secret of writing is rewriting, and the secret of rewriting is re-thinking.

Typing

A clean typing reflects the profile of a careful professional. Margins should be wide (2.5 cm), and only one side of the paper is used. One looks for adherence to the style of the journal. As a rule to obtain a clean document the experimented researcher proofreads, proofreads, and proofreads.

Benefits of writing

Last but not least let us discuss the benefits of writing. They are greater to the author than to the reader because they enhance clear thinking and lead to invaluable mental discipline. Writing improves one's reading skills and satisfies a creative instinct.

Acknowledgement. The authors would like to thank Dr. Martin J. Tobin for contributing concepts included in the section "Writing up your scientific paper".

References

1. Mulrow C (1987) The medical review article: state of the science. Ann Intern Med 106: 485-488
2. Mulrow CD (1994) Systematic reviews: rationale for systematic reviews. BMJ 309: 597-599
3. Minozzi S, Pistotti V, Forni M (2000) Searching for rehabilitation articles on MEDLINE and EMBASE. An example with cross-over design. Arch Phys Med Rehabil 81: 720-722
4. Watson RJ, Richardson PH (1999) Identifying randomized controlled trials of cognitive therapy for depression: comparing the efficiency of EMBASE, MEDLINE and PsycINFO bibliographic databases. Br J Med Psychol 72: 535-542
5. Topfer LA et al (1999) Comparison of literature searches on quality and costs for health technology assessment using the MEDLINE and EMBASE databases. Int J Technol Assess Health Care 15: 297-303

6. Woods D, Trewheellar K (1998) MEDLINE and EMBASE complement each other in literature searches (letter; comment). BMJ 316: 1166
7. Kleijnen J, Knipschild P (1992) The comprehensiveness of MEDLINE and EMBASE computer searches. Searches for controlled trials of homoeopathy, ascorbic acid for common cold and ginkgo biloba for cerebral insufficiency and intermittent claudication. Pharm Weekbl Sci 14: 316-320
8. Lowe H, Barnett G (1994) Understanding and using the Medical Subheadings (MESH) vocabulary to perform literature searches. JAMA 271: 1103-1108
9. Ball C, Sackett D, Phillips B et al (1999) Levels of evidence and grades of recommendations. Centre for Evidence Based Medicine. http://cebm.jr2.ox.ac.uk/docs/levels.html
10. Feldman RD et al (1999) Canadian recommendations for the management of hypertension. Task Force for the Development of the 1999 Canadian Recommendations for the Management of Hypertension. Can Med Assoc J 161 [Suppl 12]: S1-S17
11. De Mattia D et al (2000) Acute childhood idiopathic thrombocytopenic purpura: AIEOP consensus guidelines for diagnosis and treatment. Associazione Italiana di Ematologia e Oncologia Pediatrica. Haematologica 85: 420-424
12. Boulet LP et al (1999) Canadian Asthma Consensus Report, 1999. Canadian Asthma Consensus Group. Can Med Assoc J 161 [Suppl 11]: S1-S61
13. Oxman AD, Sackett DL, Guyatt GH (1993) Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. JAMA 270: 2598-2601
14. Guyatt GH, Sackett DL, Cook DJ (1994) Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 271: 59-63
15. Guyatt GH, Sackett DL, Cook DJ (1994) Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. JAMA 271: 389-391
16. Jaeschke R, Guyatt GH, Sackett DL (1994) Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 271: 1615-1619
17. Jaeschke R, Guyatt G, Sackett DL (1994) Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 271: 703-707
18. Levine M et al (1994) Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group. JAMA 272: 234-237
19. Laupacis A et al (1994) Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. J Dent Educ 58: 648-653
20. Oxman AD, Cook DJ, Guyatt GH (1995) Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group (see comments). Can J Psychiatry 40: 97-101
21. Richardson WS, Detsky AS (1995) Users' guides to the medical literature. VII. How to use a clinical decision analysis. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 273: 1610-1613.
22. Greenhalgh T (1997) How to read a paper. Papers that report drug trials. BMJ 315: 480-483
23. Greenhalgh T (1997) How to read a paper. Statistics for the non-statistician. II: "Significant" relations and their pitfalls. BMJ 315: 422-425
24. Greenhalgh T (1997) How to read a paper. Statistics for the non-statistician. I. Different types of data need different statistical tests (published erratum appears in BMJ 1997 Sep 13;315:675). BMJ 315: 364-366
25. Greenhalgh T (1997) Assessing the methodological quality of published papers (see comments). BMJ 315: 305-308

26. Greenhalgh T (1997) How to read a paper. Getting your bearings (deciding what the paper is about). BMJ 315: 243-246
27. Greenhalgh T (1997) How to read a paper. The MEDLINE database (see comments). BMJ 315: 180-183
28. Greenhalgh T (1997) Papers that summarise other papers (systematic reviews and meta-analyses). BMJ 315: 672-675
29. Schulman S, Rhedin AS, Lindmaker P et al (1996) Fewer venous thromboembolism recurrences occurred with 6 months than with 6 weeks of anticoagulant therapy (abstract). Evidence-Based Medicine 1: 42
30. Sackett DL, Wennberg JE (1997) Choosing the best research design for each question (editorial). BMJ 315: 1636
31. Doll R, Hill A (1952) The study of aetiology of carcinoma of the lung. BMJ 2: 1271-1286
32. Sackett D (1979) Bias in analytic research. J Chron Dis 32: 51-63
33. Study Design (2000) Centre for Evidence Based Medicine. http://cebm.jr2.ox.ac.uk/docs/studies.html
34. Kelly J, Hellinger F (1986) Physician and hospital factors associated with mortality of surgical patients. Med Care 24: 785-800
35. Pocock S (1983) Clinical trials: a practical approach. Wiley, Chichester, p 266
36. Chambers G, Fairbarn M (1998) Clinical trial design: an introduction to good clinical trial design and the protection of human subjects. Appl Clin Trials 7: 60-66
37. Richardson WS et al (1995) The well-built clinical question: a key to evidence-based decisions (editorial). ACP J Club 123: A12-A13
38. Hulley S et al (2000) Designing clinical research: an epidemiologic approach, 2nd edn Lippincott Williams and Wilkins, Philadelphia
39. Spilker B (1996) Developing clinical trial designs guide to clinical trials. In: Guide to Clinical Trials. Lippincott-Raven, Philadelphia, pp 27-58
40. Detsky A, Naglie I (1995) Subgroup analyses: primary and secondary. ACP J Club 122: A12-A14
41. Cook D, Sacket D (1992) On the clinically important difference. ACP J Club 166: A16-A17
42. Spilker B (1996) Fundamental principles, considerations, and techniques in the interpretation of clinical data. In: Guide to Clinical Trials. Lippincott-Raven, Philadelphia, pp 505-558
43. Hilsden R, Verhoef M (2000) Beginners guide to the research proposal. Centre of Advancement of Health, University of Calgary, Calgary

Further reading

What is research?

Anglin GJ, Ross SM, Morrison GR (1995) Inquiry in instructional design and technology: getting started. In: Anglin G (ed) Instructional technology: past, present, and future. Libraries Limited, Englewood

Barzun J, Graff H (1992) The modern researcher, 5th edn. Harcourt Brace, San Diego

Bouma GD (1994) The research process. Oxford University Press, London

Christensen LB (1997) Experimental methodology 7th edn. Allyn and Bacon, Needham Heights

Davitz JR, Davitz LL (1995) Evaluating research proposals: a guide for the behavioral sciences. Prentice Hall, Upper Saddle River

Howe R, Lewis R (1994) A student guide to research in social science. Cambridge University Press, New York
Mutchnick RJ, Berg BL (1995) Research methods for the social sciences: practice and applications. Allyn and Bacon, Needham Heights
Priest SH (1996) Doing media research: an introduction. Sage, Thousand Oaks

Tools of research

Angel D (1994) The elements of e-mail style: communicate effectively via electronic mail. Addison-Wesley, Reading
Barnes AS (1994) Research skills. Kendall/Hunt, Dubuque, Iowa
Bergan HJ (1995) Where the information is: a guide to electronic research. BioGuide Press, Alexandria
Christensen LB (1997) Experimental methodology, 7th edn. Allyn and Bacon, Needham Heights
DiMartino D (1995) CD-ROM search techniques of novice end-users: is the English-as-a-second-language student at a disadvantage? Coll Res Libraries 56: 49-59
Gates JK (1993) Guide to the use of libraries and information sources, 7th edn. McGraw-Hill, New York
Huck SW, Cormier WH (1995) Reading statistics and research, 2nd edn. HarperCollins, New York
Leedy PD (1997) Practical research: planning and design, 6th edn. Prentice-Hall, Upper Saddle River
Thornburg BJ (1992) CD-ROM and the academic reference librarian: a review of the literature. Electronic Library 10: 219-221

Supported by
School of Critical Care Medicine
APICE (Anaesthesia, Pain, Intensive Care and Emergency Medicine), Trieste, Italy
Department of Clinical Sciences, Section of
Anaesthesiology Intensive Care and Pain Clinic,
Trieste School of Medicine, University of Trieste, Italy