WHO Regional Publications, Eastern Mediterranean Series 30

A Practical Guide For Health Researchers

World Health Organization
Regional Office for the Eastern Mediterranean

WHO Regional Publications Eastern Mediterranean Series 30

A Practical Guide for Health Researchers

Mahmoud F. Fathalla

Professor of Obstetrics and Gynaecology
Assiut University, Egypt
Chairman, WHO Global Advisory Committee on Health Research
Former Director, UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction

Mohamed M. F. Fathalla

Department of Obstetrics and Gynaecology
Assiut University, Egypt

World Health Organization
Regional Office for the Eastern Mediterranean
Cairo, 2004

WHO Library Cataloguing in Publication Data

Fathalla, Mahmoud F.

A practical guide for health researchers / by Mahmoud F. Fathalla and Mohamed M.F. Fathalla.
p. (WHO Regional Publications, Eastern Mediterranean Series; 30)
ISBN 92-9021-363-9
1. Research–Scientific 2. Writing–Guidelines 3. Publishing–Guidelines 4.WHO Regional Office for the Eastern Mediterranean I. Fathalla, Mohamed M.F. II. Title

[NLM Classification: Q 179.9]

© World Health Organization 2004

All rights reserved.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

The World Health Organization does not warrant that the information contained in this publication is complete and correct and shall not be liable for any damages incurred as a result of its use.

The named authors alone are responsible for the views expressed in this publication.

Publications of the World Health Organization can be obtained from Distribution and Sales, World Health Organization, Regional Office for the Eastern Mediterranean, PO Box 7608, Nasr City, Cairo 11371, Egypt (tel: +202 670 2535, fax: +202 670 2492; email: DSA@emro.who.int). Requests for permission to reproduce WHO EMRO publications, in part or in whole, or to translate them – whether for sale or for noncommercial distribution – should be addressed to the Regional Adviser, Health and Biomedical Information, at the above address (fax: +202 276 5400; email HBI@emro.who.int.)

Cover design by Ahmed Hassanein

Printed in Cairo, Egypt, by Dar Mourad for Printing

Contents

Foreword 7

Preface 9

Acknowledgements 10

Chapter 1. Introduction and overview 11
References and additional sources of information 19

Chapter 2. Ethics in health research 20
2.1 Introduction 20
2.2 General ethical principles 21
2.3 Responsibility for ethics in health research 22
2.4 Ethics committees 22
2.5 Ethical considerations throughout the research process 23
References and additional sources of information 24

Chapter 3. What research to do? 25
3.1 Introduction 25
3.2 Selection of a field for research 26
3.3 Drivers for health research 29
3.4 Participation in collaborative international research 32
3.5 Participation in pharmaceutical company research 34
3.6 Where do research ideas come from? 36
3.7 Criteria for a good research topic 39
References and additional sources of information 41

Chapter 4. Planning the research 43
4.1 Introduction 43
4.2 Types of research design 44
4.3 Selecting a research design 47
4.4 Defining and refining the research question 49
4.5 Generating the research hypothesis 50
4.6 Study sample 50
4.7 Sample size 52
4.8 Measurement 54
4.9 Planning qualitative research 55
4.10 A note on questionnaire design 57
4.11 A note on research in health economics 58
4.12 Ethics in research design 59
References and additional sources of information 62

Chapter 5. Writing the research protocol 65
5.1 Introduction 65
5.2 Format for the protocol 66
References and additional sources of information 71

Chapter 6. Submitting a research proposal 72
6.1 Introduction 72
6.2 How to get your research project funded 72
6.3 Components of a research proposal 75
References and additional sources of information 78

Chapter 7. Implementing the research project 79
7.1 Introduction 79
7.2 Scientific rigour 79
7.3 Pre-testing the protocol 80
7.4 Monitoring of the study 80
7.5 Periodic tabulations and reports 82
7.6 Validation of results in qualitative research 82
7.7 Good clinical practice 82
7.8 Research on new pharmaceutical products 83
7.9 Termination of the study 84
7.10 Changes in the protocol 84
7.11 Ethical issues in the implementation of the study 85
References and additional sources of information 86

Chapter 8. Describing and analysing research results 87
8.1 Introduction 87
8.2 Descriptive statistics 87
8.3 Tabulation 88
8.4 Calculations 89
8.5 Graphs/figures 91
8.6 Correlation 91
8.7 Inferential statistics 93
8.8 What statistical tests tell us 95
8.9 Selection of statistical test 98
8.10 Examples of some common statistical tests 101
8.11 Description and analysis of results of qualitative research 102
References and additional sources of information 105

Chapter 9. Interpreting research results 106
9.1 Introduction 106
9.2 Interpreting descriptive statistics 106
9.3 Interpreting “statistical significance” 107
9.4 Bias 107
9.5 Confounding 109

9.6 Making the case for causation ... 110
9.7 Interpreting end points to measure the outcome ... 112
9.8 Interpreting studies of risk factors ... 112
9.9 Interpreting studies of diagnostic tests ... 114
9.10 Interpreting studies that report the results of interventions ... 116
9.11 Interpreting results of qualitative research ... 116
References and additional sources of information ... 117

Chapter 10. Communicating research ... 119

10.1 Introduction ... 119
10.2 Communicating to scientists ... 120
10.3 Communicating to funding agencies ... 123
10.4 Communicating to health professionals ... 124
10.5 Communicating to policy-makers ... 125
10.6 Communicating to patients ... 127
10.7 Communicating to the community ... 127
10.8 Communicating to the public ... 127
10.9 Communicating to the public media ... 128
References and additional sources of information ... 129

Chapter 11. Writing a scientific paper ... 130

11.1 Introduction ... 130
11.2 Selecting a title for the paper ... 131
11.3 Writing the abstract and key words ... 131
11.4 Article structure ... 132
11.5 Writing the Introduction ... 132
11.6 Writing the Methods section ... 132
11.7 Writing the Results ... 134
11.8 Writing the Discussion and Conclusions ... 137
11.9 Acknowledgements ... 138
11.10 Citation of references ... 139
11.11 Steps in the process of writing a paper ... 140
11.12 Revision of the manuscript for scientific content ... 141
11.13 Revision of the manuscript for style ... 142
11.14 Writing a case report ... 144
11.15 Writing a secondary scientific paper ... 145
11.16 Writing a paper on qualitative research ... 147
11.17 The dissertation or thesis ... 147
References and additional sources of information ... 149

Chapter 12. Publishing a scientific paper ... 151

12.1 Introduction ... 151
12.2 How to get your paper published ... 151
12.3 Uniform requirements for manuscripts submitted to biomedical journals ... 153
12.4 Summary of technical instructions for submission of papers ... 154

12.5 Sending the manuscript to the journal 155
12.6 After submitting the manuscript 155
12.7 Authorship in scientific papers 156
12.8 Patents and publication 157
12.9 Ethics in scientific publication 157
References and additional sources of information 161

Chapter 13. Making a scientific presentation 162
13.1 Introduction 162
13.2 Planning of the presentation 162
13.3 Preparation 163
13.4 Presentation 167
13.5 Guide to how to give a “bad” presentation 169
References and additional sources of information 170

Chapter 14. Assessment and evaluation of research 172
14.1 Introduction 172
14.2 Assessment and evaluation by researchers 173
14.3 Assessment and evaluation by health professionals 175
14.4 Assessment and evaluation by policy-makers 177
14.5 Assessment and evaluation by investors in research 180
References and additional sources of information 183

Annex 1. World Medical Association Declaration of Helsinki Ethical principles for medical research involving human subjects 185

Annex 2. International ethical guidelines for biomedical research involving human subjects 191

Annex 3. Searching the literature 201

Annex 4. Guidelines on how to write references for scientific papers 210

Annex 5. Bangkok Declaration on Health Research for Development 214

Glossary of terms in health research 216

Foreword

The central role of health research in improving health and stimulating national economic growth is now well established. Health research supports health systems in the delivery of better, fairer and more equitable health care to people. It does so by identifying challenges and providing best solutions, monitoring how health systems perform and producing new knowledge for better technologies and improved approaches to public health. The World Health Organization (WHO) has, time and again, affirmed that all national and international health policies should be based on valid scientific evidence; that such evidence requires research; and that the application of the knowledge, information and technology emanating from health research has enormous potential in promoting health.

Shifting epidemiological trends in disease patterns, rapid increase in populations, new and emerging health problems, increasing commercial interests of the private health sector and ever shrinking financial resources all contribute to the global inequity in health care. It is therefore extremely important that research addresses priorities and focuses on the most important health issues, conditions and determinants. Health research must serve as a driver for health policy and practice. For this to happen, the health research systems not only have to be fully accountable for the sake of transparency, but also have to be capable of delivering the desired returns, to justify the allocation of scarce resources to research and development.

Inadequacy in capacities for research and development remains a major impediment for the developing world. Despite over three decades of efforts to build capacities, during which thousands of scientists from developing countries have been trained, most of the expected breakthroughs have not happened. Large numbers of trained scientists are not working in their countries of origin. Therefore, building indigenous capacity for health research must move to centre stage, as this is vital for sustainable development. The WHO's Regional Office for the Eastern Mediterranean recognizes this acute need. Supporting health research for better health and building regional capacities for better quality research is an important priority.

The literature on research methodologies is vast. Researchers and scientists worldwide nowadays have access to enormous, and growing, information resources which provide in-depth knowledge, training and education to enhance and improve research. This *Practical Guide for Health Researchers* is, however, quite unique. It is different in that it is not a classical textbook on research methodology, but focuses directly on those who carry out health research or aspire to do so in the future. It embodies the seriousness, the sincerity and the passion of the authors as they try to guide and direct

the reader in her or his pursuit of research to seek new knowledge, identify problems and provide answers. The authors, with great skill, have articulated and shown the way forward for anyone who seeks the value of research, desires to undertake good quality research and aspires to draw benefits from it.

The book begins with a very strong message: health research is not a luxury, but an essential need that no nation can afford to ignore. The authors first reason out why it is so important, especially for the developing countries, to do research and explain the consequences of ignoring research as a tool for evidence on which to base planning, practice and actions. They describe the research process, beginning with the selection of a research topic, the narrowing down of specific objectives and how best to achieve the stated objectives. They describe the characteristics of a good research proposal—one that has potential for obtaining the required financing, is feasible and will produce valid information and knowledge that will ultimately have an impact on health. The various options are discussed with regard to research methodologies and strategies, and invaluable guidance is provided on data collection and its analysis.

There are other unique aspects to this book. In describing the different research approaches and methods, it underscores the merits (as well as demerits) of both quantitative and qualitative research methods, and reminds the reader of how and under what situations one or the other strategy (or both) can be helpful to the research question in mind. Another strength is the authors' emphasis on the ethics of health research. Throughout, the reader is reminded constantly of the ethical principles that govern health research, and the need for upholding and defending moral and ethical values in such practices. Some key international guidelines are indexed for the benefit of the reader.

The book provides useful tips for the health researcher, which are ordinarily absent in classical textbooks on health research methodology. It tells them how to find information pertinent to their research and how to seek funds for their research. It discusses the various ways of communicating research results to different audiences, as well as preparation of manuscripts for submission to medical journals and presentations, with an overall reminder that the culmination of the research effort should be in its application in order to bring about the required changes in policies, actions and practices.

The greatest strength of this book is that it reflects the first hand experience of the authors, especially Prof. Mahmoud Fathalla. It is heartening to note that he chose to share his expertise and the richness of his experience in health research in this manner. The book is easy and simple to follow. It demystifies health research. It is a book that every health researcher will treasure, and a ready reference that he or she will want to keep close by.

Hussein A. Gezairy MD FRCS
Regional Director for the Eastern Mediterranean

Preface

The intended audience of this book, as indicated in its title, is health researchers. Health researchers are not limited to scientists pursuing a research career. Health research can and should be pursued by a broad range of people. Health research can be simply defined as the process for systematic collection, description, analysis and interpretation of data that can be used to improve the health of individuals or groups. Health professionals, health administrators, health policy-makers and nongovernmental organizations, among others, can and should use the scientific method to guide their work for improving the health of individuals and communities. Even if they do not pursue much research themselves, they need to grasp the principles of the scientific method, to understand the value and also limitations of science, and to be able to assess and evaluate results of research before applying them.

Most textbooks on the subject of health research are written in a language that is highly technical, and for an audience of trained scientists. There is a need to demystify the research process for a broader community of health researchers. The research process is largely about good sense and reason. We have tried to make this book as reader-friendly as possible, but not at the expense of scientific accuracy.

We have attempted in this guide to cover the broad spectrum of the research process. The research process is not simply about the methodology of research design. Before considering research design, researchers need to know how to define and refine the research question. After settling on research design, they need to be able to write a research protocol, submit a proposal for funding, properly conduct the research, describe, analyse and carefully interpret the research results, and finally communicate the findings to all who stand to benefit from the research, through writing and publishing papers and making scientific presentations. Researchers need also the skills to be able to assess and evaluate the research done by others. Beginners in health research have to consult different sources if they want to get a complete grasp of this whole spectrum of the research process. Our objective was to provide a concise practical guide to cover these areas, rather than a comprehensive manual. To be able to obtain more technical detail and information on the issues discussed, we have provided a list of useful sources for each chapter, as well as a number of annexes.

We hope that this guide will help in expanding the community of health researchers, beyond the traditional groups of trained scientists. We hope it will help health researchers to plan, conduct and disseminate good research.

Mahmoud F. Fathalla
Mohamed M.F. Fathalla

Acknowledgements

We are grateful for Dr Mohamed Abdur Rab, Regional Adviser, Research and Policy Coordination, WHO Regional Office for the Eastern Mediterranean for encouraging us to complete the task of writing this guide, for providing useful comments and suggestions and for bringing it before the regional publications committee. Ms Jane Nicholson, Editor at the Regional Office for the Eastern Mediterranean, was extremely helpful throughout the process of preparing the book for publication.

Many colleagues persuaded us about the need for such a guide, and provided us with constructive comments. It would be difficult to try to name them all. The sources written by previous authors and listed under the references and additional sources have been very helpful in putting together the material for this book.

Finally, we are grateful to our students who taught us how to teach, and to our families, on whose time this book was written.

Chapter 1

Introduction and overview

The health gains in the last century have been unprecedented. Advances made in health research account for a significant part of these health gains. New scientific frontiers, now opening up, promise to transform medical practice in ways never imagined before, and to contribute to further improvements in health. However, health research is not only about the development of new tools and advancing our understanding about health and disease. Health research is important to inform policy and decision-making in health systems.

Health research is not a luxury, to be conducted only by countries with the resources to spare. When India gained independence, the country faced the problem of how to allocate its scarce resources to areas of most need. Jawahar Lal Nehru, in this context, made the following statement: "Because we are a poor country, we cannot afford not to do research". The participation and contribution of developing countries in scientific research has been well expressed by the Pakistani Nobel Laureate Abdul Salam, as follows: "Science and technology are a shared heritage of all mankind; East and West, South and North have all equally participated in their creation in the past, as, we hope, they will in the future—the joint endeavour in science becoming one of the unifying forces among the diverse peoples on the globe." (Salam, 1989.)

Health research may be pursued as a career in a public or private research organization. Research may be done in pursuit of prestige or under the pressure of the threat of "publish or perish" when climbing the ladder of a successful academic career. A strong argument can, however, be made that all health professionals should do some research, or at least get enough knowledge about the research process, even if they wish to spend the rest of their lives dealing with patients or health administration. A scientific approach is essential for health professionals. As the practice of medicine is advancing rapidly, the need for critical evaluation of new developments becomes more urgent. The medical past is littered with examples of possible major advances eventually being shown to be of no value, or even to be harmful. Research helps to develop a scientific critical attitude. A clinician will find that the faculties developed by doing research are those most needed in clinical diagnosis.

Health policy-makers, particularly in developing countries, may not appreciate the contribution which research can make. There is still a divide between the universe of research and the universe of policy-making. The stereotype of the researcher in her or

his ivory tower still prevails. In fact, health managers and policy-makers may be doing research without knowing it. Research can be defined as the systematic collection, description, analysis and interpretation of data to answer a certain question or solve a problem. Health research can also be defined as the process for systematic collection, description, analysis and interpretation of data that can be used to improve the health of individuals or groups. The research process changes "information" into "knowledge", through critical assessment and relating it to other existing human knowledge. As they go through this research exercise, health managers and policy-makers need to understand more about the process of research.

There is a need to demystify the scientific process. Scientific inquiry is basically a potentiation of common sense, which is probably one of the most equitably distributed human gifts. Einstein said, "The whole of science is nothing more than a refinement of everyday thinking." In a sense, most of us may be conducting some research in our daily life. When we, for example, want to buy a car in a proper way, we collect information about models and dealers, analyse it, then try to reach a "scientific" conclusion on which car to buy. The use of complex instrumentation is not a necessary requirement for good research. Key attributes of good research are proper planning, accuracy in data collection and proper unbiased interpretation.

There is only one type of research: good research. Bad research does not deserve the name of research. Badly done research is not only a waste of time, money and effort. It can be considered unethical if it exposes research subjects to the inherent risks of experimentation with no reward to them, to others or to their communities. This book is about how to do health research, and how to do it well.

The research process begins with selecting a field and topic for research, then planning the research, writing up the plan as a research protocol, and, where appropriate, submitting it as a research proposal for funding. Implementation of the research project is followed by describing and analysing the research results. The research results then need to be carefully and objectively interpreted. Research is not complete until it is communicated to those who may benefit from it. This commonly involves, but is not limited to, writing and publishing a scientific paper, and/or making a scientific presentation. The research process not only involves doing the research, but also assessing and evaluating research done by others. Throughout the research process, and particularly where the research involves human subjects, rules of ethical conduct must be carefully observed. All these steps in the research process are dealt with in detail in the different chapters of the book.

Because of the importance of ethics in health research, the next chapter of the book outlines the concerns about ethics in health research, general ethical principles, responsibility for maintaining ethical standards, and the duties of ethics committees which review and approve research. After this introduction to ethics, ethical considerations are

discussed in more detail in subsequent chapters dealing with different stages of the research process. In addition, Annexes 1 and 2 provide documentation on the topic: The World Medical Association Declaration of Helsinki, in its latest version; and International Ethical Guidelines for Biomedical Research Involving Human Subjects prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO).

The first decision a researcher needs to make is what research to do. This is dealt with in Chapter 3. There are different fields of health research, all of which can make a contribution to improvement of health, and all are needed. In these days of specialization and sub-specialization, the investigator may have already landed in one of the disciplines. But it is important that s/he should be aware of the other disciplines and what they can contribute. Collaboration between researchers from different disciplines is one of the most effective mechanisms in advancing health research. The distinction between basic and applied research is probably more a function of time. Basic research provides the pool of knowledge from which leads for applied research can be picked up. Also, the strong interest in quantitative research should not lead us to ignore the potential contribution of qualitative research. Qualitative research can provide insights that will not be apparent from quantitative methodologies.

The selection of a topic for research is influenced by what drives the research. Research is driven by curiosity, health needs, profit and/or opportunity. Scientists, on the one hand, are happy to pursue their own lines of interest, enjoy academic freedom and follow scientific curiosity. They can say, and they are right, that many significant discoveries in the field of health were made by serendipity, and not through targeted research. Policy-makers and funders of research, on the other hand, would like to see research targeted to respond to priority health needs. Private industry, now a major actor in health research, is driven by profit, and pursues lines of research that are likely to lead to the development of products that sell in large profitable markets. Governments in developed countries often encourage and support research for wealth creation, not just for health. The selection of topics for research may be driven by the opportunity for funding. A major concern in health research today is the 10/90 gap. Of the total funds spent worldwide each year on health research by both the private and public sectors, it has been estimated that only about 10% are devoted to the health problems of 90% of the world's population. Opportunities for research may arise through participation in collaborative international research. For developing country researchers, this is a good opportunity but not without concerns. Concerns include distortion of country priorities in research, internal brain drain where the brains of researchers in the country are working for the problems of other countries, and ethical considerations that need to be addressed. Participation in pharmaceutical company research is another opportunity. Collaboration between academia and industry is to be encouraged, but there are concerns that need to be addressed, whether the collaborative research is at the discovery stage, during clinical

testing or after marketing of the product.

Ideas for topics for research come from different sources that need to be pursued by researchers. Familiarity with the research literature is important, not only for identifying where gaps for research are, but also during the planning, implementation and writing up of the research. Annex 3 provides notes on searching the literature, using different sources, in the new information age. As will be discussed in Chapter 3, whatever the topic of research selected may be, it should satisfy the criteria of being feasible, interesting, novel, ethical and relevant.

After deciding on what research to do, Chapter 4 deals with the planning of the research. Time spent on proper planning is never lost. There are different types of research design, whether for observational or experimental intervention studies. All types of research design have their place. The investigator has to select the type of research design that will give the most definitive answer to the research question, and at the same time would be feasible to conduct. In most cases, more than one design will be possible, but a trade off has to be made between the ideal and the possible. In this context, as in others, the best should not be made the enemy of the good. In planning the research, the research topic has to be narrowed down into a well defined research question. The more refined the question, the better will be the plan. Investigators should resist the temptation to broaden the scope of inquiry beyond what can realistically be answered by the research.

With a well defined and refined research question, a research hypothesis can be generated. In proper scientific methodology, we do not develop a research hypothesis in order to prove it; we develop a hypothesis in order to test it. Scientists doing research adopt a sceptical attitude. They start with the assumption that the research hypothesis is not true, using the term “null hypothesis”. If the results do not support the “null hypothesis”, then the research hypothesis is more likely to be true. Probability is another feature of the scientific methodology. There is usually no certainty about the validity of scientific results. It is only a high level of probability that is sought. This level depends on the magnitude of the finding, as well as the size of the study. Analytical statistical methods help in assessing the level of probability.

In planning the research, a crucial question is the type and size of sample to be studied. We cannot study all the population. We need to define a target population, as well as an accessible population. The term population in scientific methodology does not necessarily refer to people; it refers to the material for the research, be it people, animals or non-animates. There are different ways of sampling. The sample size appropriate to provide the answer to the research question has to be defined. A larger sample size than needed is an inappropriate use of research resources. A smaller sample size than needed is a waste of effort and money on a study that will not provide definitive answers. The types of measurements to be used have to be carefully identified in the planning stage of the research to ensure validity and reliability. The methodologies for qualitative research

need to be appreciated and applied, as appropriate, by researchers. A research topic may be better addressed by quantitative research, qualitative research, or both.

The planning phase is also the time to think carefully about ethical considerations. Different categories of health research have their ethical implications. There are important considerations in research designs involving experimentation on human subjects, in epidemiological, field and qualitative studies, and in research involving experimentation on animals.

Chapter 5 deals with writing the research protocol. After developing the plan for the research, it has to be written down as a protocol. This is particularly important if the study is done by a team of investigators, but is also important if there is only a single investigator. It helps to clarify the thinking about the plan, and is necessary for getting approval from ethics review committees. There is a traditional format for writing research protocols. It starts with a title and a summary. The project description should then include the rationale for the study, its objective, and methodology, including statistical methods used for sample size calculation and for data management and analysis. Ethical considerations should be spelt out, where appropriate, using an ethics checklist. Where relevant, gender issues should also be addressed. The protocol should include a small number of recent and relevant references to previous work on the topic.

Chapter 6 deals with the question of how to get funding for the research project. Investigators must make themselves familiar with potential sources of funding, their lines of interest and their procedures. A research proposal has to be prepared and submitted. It should include, in addition to the protocol, information to persuade the funding agency about the importance of the project, the relevance of the research to the priorities of the agency and the capacity of the investigators to undertake it. It should outline a timetable, and any problems anticipated. A budget should be submitted, properly itemized and justified. Information about the research institution, the curriculum vitae of the investigators, and any previous work on the topic will be needed to show the capability for carrying out the research.

In Chapter 7, we move to the question of how the project should be implemented with scientific rigour. The protocol may need to be pre-tested. Elements must be in place for monitoring the study during implementation, including record-keeping and handling of data, quality assurance and quality control, periodic tabulations and reports, checking of laboratory procedures, and checking the accuracy of data. In clinical trials, the principles of Good Clinical Practice (GCP) should be observed, and the trial may be subjected to auditing. Research on new pharmaceutical products should proceed in consecutive phases, as the safety and efficacy of the product is progressively established. Once the product is shown not to be safe or effective, the trial should be terminated, and not allowed to continue. In the implementation of any study, the protocol once approved should not be changed. Particularly in multi-centre studies, violations of the protocol

cannot be tolerated, and will affect the validity of the study.

Ethical considerations are important in the implementation of the study, whether involving human subjects or experimentation on animals. The study should be monitored for adherence to ethical principles. In addition, scientific honesty in recording the results and fiscal honesty in research expenditure are basic ethical principles.

Chapter 8 deals with description and analysis of research results. Descriptive statistics are useful to summarize and present the data in a way that allows subsequent analysis. Tools of descriptive statistics include tabulation, calculations, graphs, and correlation. Tabulations include frequency distribution tables, and cross tabulations. Calculations estimate the central tendency in numerical data (the mean, median and mode), the variability (range, standard deviation and percentiles), as well as ratios and rates. Different ways are available to display the data visually in graphs. The frequency distribution curve is particularly important to show how the data are distributed, with implications for subsequent statistical analysis. A scatter diagram will show whether there is correlation between the variables, for which a correlation coefficient and a regression equation can then be calculated.

Inferential statistics try to answer the questions of whether we can infer with a good probability from the study findings, whether the findings from the sample of the study can be generalized beyond the population studied, and whether differences or associations found can be possibly explained by chance. Statistics are based on principles of common sense, which need to be understood, more than on mathematics. The investigator may not do the elaborate mathematics, but must fully grasp the underlying concepts behind the statistical method, and must make decisions about the questions that need to be answered by statistical analysis, and the degree of uncertainty that can be acceptable. The chapter includes a description of the concept of “standard error”, tests of statistical significance, the use of “confidence intervals”, the concept of “statistical power”, as well as a note on some common statistical methods.

Description and analysis of research results is much easier and less tricky than their interpretation. Chapter 9 deals with the many pitfalls, shortcomings, and misconceptions in the interpretation of results. It describes pitfalls in the interpretation of descriptive statistics, whether they deal with the mean, graphs or correlation. The term “statistical significance” should be understood only for what it stands for. It simply means that the finding or difference is unlikely to be due to chance. It does not necessarily mean that the finding is important. Bias, whether in selection or measurement, and confounding factors must be excluded before drawing any conclusions. Association of two variables should not be taken to mean a causal relationship. Scientific criteria for making the case for causation must be fulfilled. Care should be taken in trying to extrapolate from results using other end points, as a surrogate for the outcome in question.

Special studies need careful interpretation to avoid any misconceptions about the

results. When studies of risk factors are interpreted, we need to understand the concepts of basic risk, relative risk, confidence intervals, attributable risk, as well as the need to balance risks and benefits. In reporting studies of diagnostic tests, the investigators must report on sensitivity, specificity, predictive value, and efficiency. A trade-off may need to be made between sensitivity and specificity. Studies reporting the results of interventions need careful interpretation, including cost considerations. The concept of "the number needed to treat" in order to achieve the advantage of the intervention is a useful tool that is not always considered.

Research is not complete before its results are communicated. This is dealt with in Chapter 10. Most of the communication done is to fellow scientists. But the beneficiaries of health research are much broader than the scientific community, and they are entitled to the information. Communication to scientists is commonly in the form of publication in peer reviewed scientific journals. The availability of expensive scientific journals is limited, particularly in developing countries. Thanks to the internet, new initiatives are underway to allow researchers to communicate their research findings to a much wider audience, with the ultimate hope that scientific information will be made freely available to all who want it. The age of paperless papers is now speeding up the process of submission, review and publication, making scientific information more up to date (apart from saving many trees in the process). Presentation of the results of scientific research in scientific meetings is another approach for exchanging scientific information, with advantages and disadvantages relative to publication. Research results should also be communicated periodically to the funding agency. Release of funding is generally contingent on receipt of satisfactory progress and financial reports.

For health research to influence the way health professionals practice, it should be communicated in a user-friendly but accurate way. The research findings may need to be synthesized in systematic reviews. Practice guidelines, developed after rigorous review of various studies, can be very useful. For many studies, it is more important to communicate the results to policy-makers. Submission of a report is generally not enough. Guidelines on how to make a presentation to policy-makers are given. If the research was based on a community study, the community involved has a right not only to know, but also to discuss the research results.

The millennium of cybermedicine promises a revolution in the availability of health information, not only for health professionals, but also for patients and for the public at large. Scientists should also learn how to communicate with the public, who are often paying the bill for the research. A constructive dialogue between scientists and the public is becoming increasingly important. A favourable scientific public environment is essential for science to thrive. Recently, there have been increasing signs of public mistrust in science. This has to be overcome through better communication between scientists and the public. The public needs to be adequately informed to make appropriate decisions.

Because of the importance of the scientific paper as a way of scientific communication, detailed guidelines are provided in Chapter 11 on how to write it. Guidelines deal with the selection of the title, writing the abstract, and following the classical article structure of introduction, methods, results (including the use of tables and illustrations) and discussion. Annex 4 provides detailed instruction on how to cite the references, from different sources, in the paper. The steps in writing the scientific paper should start before the research is implemented. The process should continue during the research, and is to be completed after the research. After writing, the manuscript should be revised for scientific content, using a checklist. After revision of the manuscript for scientific content, it should be carefully revised for style, revising paragraphs, sentences, and words. Revision for style is particularly important for those writing in a language that is not their first language, but should not be ignored by those writing in their first language. Writing a case report, and writing secondary scientific papers (narrative review, systematic review and meta-analysis) requires different formats. There are also special considerations for writing a paper on qualitative research, and for writing a dissertation or thesis.

After writing the scientific paper, comes the task of getting it published. Chapter 12 gives advice on how to get a paper published, based on defining its message, matching the topic with the interest of the journal, checking scientific validity of the results, and ensuring quality of the manuscript. The International Committee of Medical Journal Editors has agreed on uniform requirements for manuscripts to be submitted for publication. A summary of the technical requirements is given, as well as guidance on how to send the manuscript and how to deal with reviewers' comments. The International Committee of Medical Journal Editors has also recommended guidelines on authorship, emphasizing that intellectual input in the study is a requisite for qualifying for authorship. Issues of potentially patentable findings need to be addressed, where appropriate, preferably by a special office in the institution, before submitting the findings for publication to be available in the public domain. Ethical considerations apply to research communication, and include questions of credit, conflict of interest, redundant or duplicate publication, protection of patient's rights to privacy, release of information to public media before the publication, and the serious accusation of scientific fraud.

Chapter 13 provides detailed guidelines on how to make a presentation to a scientific meeting, by good planning, good preparation (including preparation of text and visual aids, as well as rehearsal), and presenting in style (getting ready, speaking well, managing the visual aids, keeping to the time and answering questions).

Researchers need to acquire the ability to assess and evaluate science, and to develop a critical attitude. Science is not to be admired; science is to be questioned. Chapter 14 provides guidance on how to read and review a scientific paper, how to evaluate the scientific evidence, how to assess scientific reviews and meta-analyses,

how to apply evidence to practice, and how to assess the appropriateness of health technologies. Evaluation of the investment in research should take into consideration, not only the impact on the advancement of science and the impact on wealth creation, but also, importantly, the impact on health promotion. The scientific quality of research, as assessed by scientists, does not necessarily go hand in hand with the impact on health promotion.

It should also be recognized that health is wealth, and that health research is important for overall development. Annex 5 provides the Bangkok Declaration on Health Research for Development.

Not all issues about health research can be covered in detail in this short guide. The book ends with a list of sources for each chapter for those who want to get more information on the particular subject.

References and additional sources of information

Commission on Health Research for Development. *Health research: essential link to equity in development*. Oxford, Oxford University Press, 1990.

Einstein A. http://www.memorablequotations.com/einstein.htm (accessed 30/3/2004).

Global Forum for Health Research. *The 10/90 report on health research 2000*. Geneva, Global Forum for Health research, 2000.

Medawar PB. *Advice to a young scientist*. New York, Basic Books, 1979.

Investing in health research and development: Report of the Ad Hoc Committee on Health Research Relating to Future Intervention Options. Geneva, World Health Organization, 1996.

Sustaining the strength of the UK in healthcare and life sciences R&D: competition, cooperation and cultural change. A report on the third SmithKline Beecham R&D Policy Symposium. Oxted, ScienceBridge, 1996: 17.

Salam A. *Notes on science, technology and science education in the development of the south. Prepared for the fifth meeting of the South Commission, May 27–30, 1989, Maputo, Mozambique*. Trieste, The Third World Academy of Sciences.

Chapter 2

Ethics in health research

2.1 Introduction

> No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment. In particular, no one shall be subjected without his free consent to medical or scientific experimentation
>
> International Covenant on Civil and Political Rights, Article 7, 1966

A number of developments have brought the subject of ethics in medical research to the front line of concerns of the health profession and the society at large. These include a major expansion in health research, the significant public investment in research, the increasing need for experimentation on human subjects, publicized cases of ethical violation, internationalization of research, and the expanding role of private industry. This century has witnessed a major expansion in health research. Medical research has opened new areas for investigation, for which society has not yet been fully prepared morally, legally and socially. These include areas such as organ transplantation, assisted conception, advances in fertility regulation, and the new era of genomics. Societies make a significant investment in health research. They have become shareholders and thus have a say in how their investment is made.

Advancement of medical knowledge depends, to a large extent, on expansion of research involving experimentation on human subjects. With the increasing acceptance and appreciation of individual human rights, and of the need to respect and protect them, it is not acceptable that the welfare and the respect of the individuals be compromised in the pursuit of benefits that may accrue to science and society. Instances of violation of ethical principles for the sake of advancement of science have occurred. The most outrageous cases were revealed in the Nuremberg trials after the Second World War. These resulted in the elaboration of the Nuremberg Code in 1947, for regulating experimentation on human subjects. The medical profession then took charge and the World Medical Association, starting in 1964, developed, adopted and updated the Helsinki Declarations which today provide guidance to the medical research community (Annex 1).

Internationalization has been a recent phenomenon in medical research. Research now knows no national frontiers. There is a need for agreement on the basic values that

govern medical research, so that the same standards apply to subjects participating in the same research in different countries. It is feared, sometimes for good reason, that advantage may be taken of countries that do not have, or do not enforce, high ethical standards, in order to advance medical knowledge, and particularly if the benefit will go primarily to other populations.

Medical research is now a major investment for private industry. Economic gains are anticipated. The strong drive to make health research an engine of economic development runs the risk of pushing research beyond acceptable ethical standards.

This chapter provides only a brief general introduction of the subject of ethics in health research. Ethical considerations are discussed in more detail in subsequent chapters dealing with what research to do, planning of the research, writing the research protocol, submitting a research proposal, implementing the research, as well publication ethics.

2.2 General ethical principles

Ethics are principles of right conduct. There are generally no disagreements on the ethical principles in themselves, since they represent basic human values. There can, however, be differences on how they are interpreted and implemented in specific cases. Basic principles include beneficence, non-maleficence, respect and justice.

Where research involves experimentation on human subjects, every effort should be made to maximize the benefits to the subjects (beneficence), and the subjects should suffer no harm (non-maleficence). The principle of respect implies that participation in the research should be completely voluntary and based on informed consent. Where research involves collection of data on individuals, privacy should be protected by ensuring confidentiality. Respect to the community means respecting its values and having its approval for the research. The principle of justice (distributive justice) implies that participation in the research should correlate with expected benefits. No population group should carry an undue burden of research for the benefit of another group.

Apart from the basic principles of beneficence (non-maleficence), respect and distributive justice, other principles also apply. Where research involves experimentation on animals, mercy is an ethical imperative. For research in general, medical or non-medical, honesty is an indispensable value. International ethical guidelines for biomedical research involving human subjects have been issued by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization. The latest edition was issued in 2000 (Annex 2).

2.3 Responsibility for ethics in health research

Responsibility for ensuring that ethical standards are observed in research rests collectively with the investigators, research institutions, national drug regulatory agencies, editors of medical journals, and funding agencies and organizations. Ethical approval by one does not relieve the others of responsibility.

- Investigators: The primary and ultimate responsibility rests with the investigators who should, as a part of their training, be made aware of and sensitive to the ethical imperatives in research. No research protocol is complete or acceptable if it does not discuss the ethical aspects of a study involving human subjects or experimental animals.
- Research institution: The research institution is responsible for the ethical quality of the research performed by its staff and in its facilities. Any institution involved in research on human subjects should have an institutional ethics review committee. The committee acts as a gathering of the investigators' peers and others to provide advice on ethical aspects of the study and to approve it or disapprove it on behalf of the institution. The membership may include other health professionals, particularly nurses, as well as laymen qualified to represent the community's cultural and moral values. The committee should be completely independent from the investigators. Any member with a direct interest in a proposal should not participate in its assessment. The next section provides more information on ethics committees.
- National Drug Regulatory Agency: New drugs or devices that are not yet approved in the country should not be used on human subjects without approval being obtained for their use under the conditions of the study.
- Editors of medical journals: Reports of research not complying with ethical standards should not be accepted for publication.
- Funding agencies and organizations: No research proposal should be funded by a national or international agency unless it has clearly outlined the ethical aspects of the study and has provided assurances that ethical principles will be observed, including, as appropriate, the approval of an institutional review committee.

2.4 Ethics committees

Countries and institutions should establish ethical review systems to ensure the protection of potential research participants and contribute to the highest attainable quality in the science and ethics of health research. Ethics committees should be established, as appropriate, at the national, regional and institutional levels.

The World Health Organization has issued operational guidelines for ethics committees that review biomedical research, outlining their role, how they can be constituted, procedure for submitting an application, elements for review, decision-making, follow-up, and documentation and archiving (WHO, 2000). The elements of ethical review include scientific design and conduct of the study, recruitment, care and protection of research participants, protection of participant confidentiality, informed consent process and community considerations. Some aspects of the work of ethics committees need to be highlighted.

- Ethics committees should be so constituted as to ensure the competent review and evaluation of all ethical aspects of the research projects they receive and to ensure that their task can be executed free from any bias and influence that could affect their independence.
- Ethics committees should be multidisciplinary and multi-sectoral in composition, including relevant scientific expertise, balanced age and gender distribution, and laypersons representing the interests and concerns of the community.
- Ethics committees should be established in accordance with the applicable laws and regulations of the country and in accordance with the values and principles of the communities they serve.
- Ethics committees should establish publicly available standard operating procedures that state the authority under which the committee is established, the functions and duties of the committee, membership requirements, the terms of appointment, the conditions of appointment, the offices, the structure of the secretariat, internal procedures and quorum requirements. They should act in accordance with their written operating procedures.

It may be helpful to summarize the activities of the ethics committees in a regular (annual) report.

2.5 Ethical considerations throughout the research process

The research process begins with the choice of the research topic, followed by selection of the appropriate research design, development of the research protocol, writing and submitting a research proposal for funding, implementing the study, description and analysis of the research results, interpretation of the research results, and finally communicating the research, including its publication. Ethical considerations apply throughout the research process, and will be discussed in the relevant chapters. The objective of this approach is to demonstrate that ethical considerations are integral components of the research process, and are not a subject to be discussed separately. In fact, scientific assessment of the planned research is an important part of the ethical

review process. It is unethical to expose subjects to research that is not scientifically sound, is not performed by qualified investigators in qualified facilities, and is not likely to provide valid scientific answers.

References and additional sources of information

Angell M. The ethics of clinical research in the Third World. *New England Journal of Medicine*, (editorial). 1997, 337, 847–849.

International ethical guidelines for biomedical research involving human subjects. Geneva, Council for International Organizations of Medical Sciences (CIOMS), 2002.

International guiding principles for biomedical research involving animals. Geneva, Council for International Organizations of Medical Sciences (CIOMS), 1985.

Research ethics training curriculum. [CD-ROM teaching aid appropriate for international biomedical and social science researchers]. Family Health International, 2001. (E-mail: publications@fhi.org).

Fluss S.S. *International guidelines on bioethics: informal listing of selected international codes, declarations, guidelines, etc. on medical ethics/ bioethics/ health care ethics/ human rights aspects of health*. Geneva, Council for International Organizations of Medical Sciences, 2000.

Lansang MA, Crawley FP. The ethics of international biomedical research (editorial). *British Medical Journal,* 2000, 321: 777–778.

Lurie P, Wolfe SM. Unethical trials of interventions to reduce perinatal transmission of the human immunodeficiency virus in developing countries. *New England Journal of Medicine*, 1997, 337: 853–855.

Resnik DB. *The ethics of science: an introduction*. London and New York, Routledge, 1999.

Singer PA, Benatar SR. Beyond Helsinki: a vision for global health ethics (editorial). *British Medical Journal,* 2001, 322: 747–748.

Operational guidelines for ethics committees that review biomedical research. Geneva, World Health Organization, 2000 (TDR/PRD/Ethics/2000/1 available from http: //www.who.int/tdr/publications/publications/pdf/ethics.pdf accessed 22/7/2004).

World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. World Medical Association, 2000 (http://www.wma.net/e/policy/b3.htm accessed 22/7/2004).

Chapter 3
What research to do?

3.1 Introduction

The question of what research to do is not faced by researchers only. Policy-makers and funders also have to make decisions on what research to encourage and support.

Health research can be done in different fields of science, including biomedical sciences, population sciences and health policy sciences. Collaboration is to be encouraged among researchers in these fields of science, which are all relevant to the improvement of health. Multidisciplinary research is becoming a necessity. There is a need for both basic and applied research, as well as for both quantitative and qualitative research.

What drives health research? Health research may be curiosity-driven, needs-driven, profit-driven or opportunity-driven. Scientists like to pursue research out of curiosity, in their own lines of interest, according to traditions of academic freedom. But research is becoming a more and more expensive undertaking. Those who control the purse would like to dictate the type of research to be supported. Governments are responsive to the concerns of their constituencies, and would like to support research that will promote the health of their populations, or will generate wealth. Private industry is becoming the major actor in health research, in terms of funding. Being accountable to their shareholders, companies pursue research for profit. These facts of life lead to a gap between the research needs in developing countries and the level of funding available to address these needs.

As far as the individual researcher is concerned, research may also be opportunity-driven. It may be driven by the opportunity for funding from national or international sources, by the opportunity to participate in multi-centre international research, or by opportunities to participate in industry-sponsored research. These opportunities raise concerns, which need to be considered before undertaking the research.

Good research ideas come from the knowledge, work and attitudes of researchers. They also necessitate an ability to navigate the expanding jungle of already available scientific information. Whatever research topic is selected, it must be feasible, interesting, novel, ethical and relevant, as will be discussed later in this chapter.

3.2 Selection of a field for research

3.2.1 Categories of health research

Health research has been broadly defined as the generation of new knowledge using the scientific method to identify and deal with health problems (Commission on Health Research for Development, 1991). Health research is thus not limited to the biomedical field. Other fields of science can contribute much to the improvement of our understanding about health issues. Broadly speaking, the following categories of science are involved in health research. Under each category, there are a growing number of specialties and sub-specialties.

- Biomedical sciences: These include all biological, medical and clinical research, and biomedical product development and evaluation.
- Population sciences: These include epidemiology, demography and the socio-behavioural sciences.
- Health policy sciences: These include health policy research, health systems research and health services research. Economic analysis studies are now an important sub-category of health policy research.

Researchers in these different fields of science, which are relevant to the improvement of health, are encouraged to collaborate.

It should also be acknowledged that the progress of science in other fields could have significant impact on the health of people. Agricultural and environment sciences are just two such examples, among others.

3.2.2 Multidisciplinary research

With the expansion of science, there has been the inevitable trend for specialization and sub-specialization. This has its merits. It also has drawbacks because cross-fertilization between the different disciplines can benefit the advancement of science. There is an increasing trend for doing multidisciplinary research.

A study by the Wellcome Trust showed that the proportion of papers in biomedical research with a single author decreased in the United Kingdom from 16.6% to 12.9% of papers published between the years 1988 and 1995, respectively (Dawson et al., 1998). The average number of authors per paper rose from 3.2 to 3.8, an indication of an increasing level of collaboration in biomedical research, and an indication that it has become more multidisciplinary. The mean number of addresses per paper rose from 1.7 to 2. There was evidence that both the number of authors and the number of funding organizations on a paper were associated with increased impact: as indicated by the number of subsequent citations of the paper in other publications.

3.2.3 Basic versus applied research

Francis Bacon in the 17th century made the distinction between scientific experiments for light (i.e. knowledge) and experiments for fruit (i.e. results) (Medawar, 1979). We can add to this statement that we need to have "light" in order to be able to search for "fruit". However, in the field of health research, and science in general, the "pure" (basic) versus "applied" debate has raged for decades and shows no signs of abating.

The creation of knowledge has been seen as an end in itself, improving our understanding of the natural world. With the rising cost of research, and the competitive demands for funding, there has been a move to emphasize and promote research that has the potential to improve health or quality of life, i.e. applied research.

It should be recognized, however, that we need a large pool of basic research. Without the availability of this pool, we will have no leads to pursue in our applied research. It can therefore be rightly remarked that there are only two types of science: "applied" science, and "not yet applied" science.

3.2.4 Quantitative versus qualitative research

Clinicians are trained to think mechanistically, and clinicians are therefore most familiar with quantitative research. However, medicine is not only a mechanistic and quantitative science. Patients are not broken down machines or malfunctioning biological systems. Doctors do not treat diseases; doctors treat patients. Health is more in the hands of people than in the hands of health professionals. Qualitative research is needed to provide insights into people's lifestyle behaviour, their knowledge, their feelings and attitudes, their opinions and values and their experience.

Having a good health system structure in place is not enough to ensure good quality health care. How the system functions and the attitudes of health care providers can make all the difference. Quantitative research gives adequate results about the anatomy of the system. Qualitative research gives insights into the physiology of the system. Good anatomy does not always mean good physiology.

Qualitative and quantitative research are not alternatives. Rather than thinking of qualitative and quantitative strategies as incompatible, they should be seen as complementary. They may help to answer the same questions. The investigators may start with qualitative research, which will then pave the way for the design of a quantitative study. A quantitative study may be complemented with a qualitative study to provide further insights into the findings. For example, a quantitative study may reveal findings about the prevalence of tobacco smoking among different segments of population. A supplementary qualitative study can then explore, in depth and in smaller groups of people, why they smoke, what they know about the risks of smoking and what was their experience in trying to stop smoking. A study on HIV (human immunodeficiency virus)

infection prevention may show that people know about the methods of prevention, but that many do not practise them. An in-depth qualitative study can explore the reasons behind this attitude. While qualitative and quantitative research may investigate the same topic, each will address a different type of question. For example, adherence to drug treatment can be examined in a quantitative study as well as a qualitative study.

Qualitative research can help in closing the gap between the science of discovery and the implementation of results. Qualitative research is often needed to find out why research results are often not translated into practice. Incorporating qualitative research methodologies into research thinking ensures that the right methodology is brought to bear on the right question.

3.2.5 Action research

Action research is a style of research, rather than a specific methodology. In action research, the researchers work with the people and for the people, rather than undertake research on them. The focus of action research is on generating solutions to problems identified by the people who are going to use the results of research. Action research is not synonymous with qualitative research. But it typically draws on qualitative methods such as interviews and observations.

3.2.6 Research in health economics

It was only recently that economists began to give attention and apply classic economic theory to the issue of the use of health care resources. No matter how rich a nation becomes, the amount of resources it devotes to health is, and always will be, limited and in competition with other possible uses. As resources are scarce, each decision to use resources in one way implies a sacrifice of another opportunity to use the resources in an alternative way. In economic evaluation, costs are regarded as opportunity costs. A common misconception is that health economics is about cutting costs. Health economics is a logic framework which allows us to reach conclusions about the best way that resources can be allocated.

3.2.7 Big science

The nature of health research has been evolving. Relatively small projects initiated by single or small groups of investigators have traditionally been, and continue to be, a mainstay of science. Recent technological advances now allow the exploration of big questions which cannot be answered by small-scale research. The human genome project is the biggest and best-known large-scale biomedical research project undertaken to date. The implications of “big science” for future health research were explored in a report by

the United States Institute of Medicine and National Research Council, under the title "Large-scale biomedical science—exploring strategies for future research" (Nass and Stilman, 2003).

3.3 Drivers for health research

3.3.1 Curiosity-driven research

Scientists enjoy doing research. They are attracted by the fun of the chase. In many types of biomedical research, discovery is the prize in the research game. But hunting for discovery is not a straightforward undertaking.

It is true that many important discoveries in science were not found because they were actively sought; they were found because it was possible to find them. Science is unpredictable. There is no guarantee that research, actively and methodologically pursued, will lead to the discovery of what it set out to discover. It may do; alternatively, something completely different may be found. Many of the drugs we use today have been discovered in research programmes designed for other purposes. Minoxidil (the drug for male baldness) was originally developed and tested for the treatment of hypertension. Sildenafil (Viagra), used for the treatment of erectile dysfunction, was discovered in a cardiovascular research programme.

In fact, serendipity plays an important role in scientific discovery. Serendipity is the faculty of making happy discoveries by accident and is derived from the title of the fairy tale The Three Princes of Serendip (an ancient name for Sri Lanka), the heroes of which were always making such discoveries. Endless examples exist in which chance played the important role in discovery. But three points are important. First, these opportunities come more often to active bench workers and to those involved in research. Second, chance presents only a faint clue that a potential opportunity exists, but the opportunity will be overlooked except by that one person with the scientific curiosity and the talent to grasp its significance. Third, the discovery made by serendipity will need to be rigorously pursued to a fruitful end.

One eminent scientist advised: "Keep on going and the chances are that you will stumble on something, perhaps when you are least expecting it. I have never heard of anyone stumbling on something sitting down." (Heath, 1985.)

Pasteur said: "In the fields of observation, chance favours only the prepared mind." (Roberts, 1989.) It has been said that the seeds of great discovery are constantly floating around us, but that they only take root in minds well prepared to receive them. Alexander Fleming in the summer of 1928, working in St Mary's Hospital in London, was not looking for an antibacterial agent at the time a spore floated into his Petri dish. But he was extremely well read and trained in microbiology and could easily recognize the

meaning of the clear area in the bacterial culture produced by the accidental implantation of the mould. It is possible that many bacteriologists have encountered similar incidents and simply discarded those contaminated cultures. In fact the use of moulds against infections was not totally new. There are records of moulds from bread being used by the ancient Egyptians. Fleming made the discovery in 1928, but it was not until the late 1930s that Howard Florey in Oxford succeeded in concentrating and purifying penicillin (Roberts, 1989).

In 1889 in Strasbourg, while studying the function of the pancreas in digestion, Joseph von Mering and Oscar Minkowski removed the pancreas from a dog. One day later, a laboratory assistant called their attention to a swarm of flies around the urine from this dog. Curious about why the flies were attracted to the urine, they analysed it and found it was loaded with sugar, a common sign of diabetes. But it was only in 1921 that Canadian researchers Fredrick Banting (a young medical doctor), Charles Best (a medical student), and John Macleod (a professor) could extract the secretion from the pancreas of dogs, inject it into dogs rendered diabetic, and prove its effectiveness (Roberts, 1989).

3.3.2 Needs-driven research

Health policy-makers at the national and international level would like to see research driven by the health needs, with a return on the investment that can decrease the disease burden on their people. The relative magnitude of a health problem is determined by its prevalence and its seriousness. A health problem may be prevalent but not serious, and may be serious but not widely prevalent. The tradition in the past has been to consider mortality as the measure for the seriousness of a health problem. This has two drawbacks. First, mortality at a young age cannot be equated with mortality at old age. It is the number of life years lost that counts, rather than the mortality rate. Second, morbidity cannot be ignored. Disability as a result of the health problem should be weighed and taken into full consideration. Mortality does not always go with morbidity. Some disease conditions leave the patient seriously morbid but do not kill. Conversely, some diseases either kill or leave no long-term impairment in health. In the field of international health now, the burden of disease as a result of any health problem is commonly expressed as the disability-adjusted life years (DALYs) lost. This measure expresses both time lost through premature death and time lived with a disability.

The fact that a health problem is of high magnitude does not necessarily mean that it should be a priority for research. The know-how to deal with the problem may be already available, but it is not applied and made available. The need may be for action and not for research. Research should not be an excuse for delaying action.

A health problem may also be of high magnitude, and there may be a need for research to be able to address it. However, before it can be put as a priority for research, other questions need to be asked. Is enough known about the problem now to consider looking for possible interventions? Does the state of the art allow a move forward to develop new interventions? How cost-effective will these interventions be? Can they be developed soon and for a reasonable outlay? This may not always be the case. Finally, is this need for research already being met by currently ongoing research, to which not much can be added?

3.3.3 Profit-driven research

Industry has become a major actor in health research. The research and development share of sales revenues varies among pharmaceutical companies, but is estimated on average to be 13%. In the 1990s, seven countries—United States of America, Japan, United Kingdom, Germany, Switzerland, France and Italy (in decreasing order)—conducted 97% of all worldwide pharmaceutical research and development (Murray et al. 1994). Pharmaceutical industry investments in research and development surpassed public investments in four of the countries (France, Japan, Switzerland and United Kingdom).

The direction for research and development in industry is pushed by the new developments in technology, which provide new leads for developing new drugs. The market pull impacts, however, on the technology push, and thus on the opportunities for research. For example, in the industrialized countries people over 65 years old spend the most on drugs. This aging population is driving new and expanding markets. New drugs are targeting age-related disorders and enhancing quality of life for the elderly (Burrill, 1998). The recent top-selling drugs were mostly in this category, for example Eli Lilly's Evista for osteoporosis, Merck's Propecia for male pattern baldness, Pfizer's erectile dysfunction pill Viagra (with estimated sales of US$ 2 billion by 2000), and Monasto's Celebra for arthritic pain.

Only a very small share of the large research investment by industry is addressed to the health problems of developing countries.

3.3.4 Opportunity-driven research

Selection of a topic for research may be driven by opportunity. The opportunity comes with the availability of funding, the chance to participate in collaborative international research, and working with the pharmaceutical industry. These opportunities provide advantages to the investigator, but they also raise some concerns.

Availability of funding

Research is often driven by the availability of funding, which may or may not correspond to local priority needs or to the curiosity of scientists. Modern research is becoming more and more expensive, and external funding is needed to conduct good research. The trend in research is increasingly moving away from local autonomy and pluralism towards some sort of centralism and dirigism. A study by the Wellcome Trust showed that from 1988 to 1995, there was a reduction from 40% to 33% in the number of research and development papers in the United Kingdom without a funding acknowledgement (Dawson et al., 1998).

Funding for health research basically comes from either public sources, including governments and United Nations intergovernmental organizations, or private sources including for-profit pharmaceutical industry and not-for-profit agencies, such as philanthropic foundations and nongovernmental organizations. Global investment in health research and development in 1998 totalled an estimated US$ 73.5 billion, or about 3.4% of health expenditures worldwide (Global Forum for Health Research, 2001): US$ 34.5 billion or 47% from governments in developed countries; US$ 30.5 billion or 42% from the pharmaceutical industry; US$ 6 billion or 8% from the private not-for-profit sector; US$ 2.7 billion or 3% from governments in developing countries.

Funding has never been more available for health research than it is today. However, there is a gross imbalance in how it is directed. Both the public sector and the pharmaceutical industry are likely to be most responsive to the burden of disease in developed countries. Investment for research by governments of rich countries is driven by the ballot box. They have to be responsive to the needs of their own electorate. Investment for research by industry is driven by market forces.

3.4 Participation in collaborative international research

3.4.1 Models for participation in international health research

Research is an international activity. Knowledge is created and built up incrementally through the work of scientists of different nations. There is no such thing as self-reliance in science. Science is a collaborative effort, involving scientists of the past, present and future. Science is international. There is no national science; there is a national contribution to the pool of science.

There are different models for participation in international health research, including participation in multi-centre clinical trials, the network approach, and the twinning approach.

Participation in multi-centre clinical trials

Multi-centre clinical trials allow recruitment of the required large number of subjects for a trial in a reasonable time. They also allow the perspectives of a number of countries to be taken into consideration. The dispatch of research forms can now be further speeded up through electronic communication.

It is important that centres involved in clinical trials make an intellectual input into the study and not just act as data collectors. The participation of investigators in the collection of the data alone does not qualify them to be authors of the published results.

The trial has to follow a protocol that should not be violated in any of the centres. Many trials, however, allow for some additions to be made by different centres, provided they are relevant to the local context, do not bias the outcome of the study, and are agreed upon.

Data analysis is usually centralized in a coordinating centre. But after completion of the trial, a centre can do further analysis on its own data.

Network approach

In a network approach, a number of centres collaborate in one research project, each centre dealing with one part of the project. One of the best known examples is the very extensive network of centres, in a number of countries, which participated in the human genome project. The project was too vast for one country to consider, but it was successfully achieved with this network approach. Many scientific enterprises are only feasible on a multinational scale. There are currently a number of networks, in both developed and developing countries, collaborating in different research programmes.

Twinning approach

Scientists and research institutions in developed countries should be encouraged to develop healthy partnerships with developing country institutions. In this way, they will not only contribute to solving problems in the developing part of our “global health village”, but they will also learn lessons that can be applied in their own countries. Scientists in developing countries should also be encouraged and supported to participate and make a contribution to the global research effort. Scientists in developing countries can live with their small salary (in a country where small salary is the norm and not the exception), but they dread, as scientists, over and above many things, the sense of isolation.

3.4.2 Concerns in developing countries about international health research

International health research provides good opportunities for developing country researchers. There are, however, certain concerns to consider. Country priorities for research should not be distorted. There is the potential for internal brain drain. There are also valid ethical concerns.

The availability of external funding can distort the national priorities for health research. Each developing country should establish and strengthen an appropriate health research base to understand its own problems, improve health policy and management, enhance the effectiveness of limited resources, foster innovation and experimentation, and provide the foundation for a stronger developing country voice in setting international priorities. This has been given the term essential national health research (Commission on Health Research for Development, 1990).

Another concern is the internal brain drain problem. The brain drain is not simply geographical. Brain drain can take place while the scientists are in their own countries, if their interests and scientific pursuits are completely irrelevant to their country's problems.

The same ethical standards that apply to research in developed countries should apply to research in developing countries. Advantage should not be taken of developing country centres to do research that would not be considered ethical in other countries.

Research should not be done in one country for the benefit of another country. Research subjects and/or their communities, should stand to benefit from the research conducted on them.

There should be no place for so-called "safari research" where expatriate scientists parachute in, do the research they are interested in, and leave, while the local community is left wondering at what was going on. It may be cheaper and faster this way, but it leaves little on the ground. It cannot be ethically justified.

3.5 Participation in pharmaceutical company research

3.5.1 Collaboration between industry and academia

It has been the tradition of pharmaceutical companies in the past to do most of their research and development in-house. Nowadays, a growing number of pharmaceutical companies commission their research to reputable centres in universities. Many companies today outsource more than 30% of their research and development budget and all or part of their clinical research and development (Burrill, 1998).

It must be noted that research for profit is no longer the domain of industry only. The myth about the academia–industry divide is being debunked. The traditional stereotype of scientists working on obscure problems in ivory towers is becoming obsolete. Although some people may still hold a stereotyped view that commercial exploitation is alien to academic research, universities and other public sector research organizations are now working closely with industry, scanning research portfolios for development opportunities.

Collaboration with industry is to be encouraged, because of the important role industry plays in the innovation process. The advantage of participation in industry-sponsored research is that it is usually well funded, and is more likely to be pursued for clinical application. There are, however, important concerns to consider.

3.5.2 Concerns about participation in industry-sponsored research

There are important issues for the independent investigator to consider, when involved in industry-sponsored research. Participation in research sponsored by pharmaceutical companies generally takes place at one or another of the different stages of development of the drug: discovery research, clinical testing and post-marketing research.

- Discovery research: For research at the stage of discovery, agreement must be reached between the research institution and the industry about patent and licensing rights, for any patentable discovery that is made during the research. Most advanced institutions have legal counsels to advise on drafting the language of these agreements.
- Clinical testing: For research at the clinical trial stages, the research should be done according to established guidelines on Good Clinical Practice (GCP) as outlined later in the chapter on implementing the research. Scientists should retain their objectivity in working with industry. As the persons directly responsible for their work, researchers should not enter into agreements that interfere with their access to the data or their ability to analyse it independently, to prepare manuscripts and to publish them (International Committee of Medical Journal Editors, 2003) Prestigious journals require investigators submitting papers for publication to declare who has sponsored the study, and whether they had any non-scientific, for example commercial, interest in the outcome of the study. If the clinical research is partly supported by a public-sector research organization, an agreement should be reached with industry on the benefit in return for the public sector in developing countries, if the research is successful. This usually means concessionary prices for the product.
- Post-marketing research: Post-marketing research sponsored by pharmaceutical companies usually has a promotional objective. It aims at making the clinicians more familiar with the drug. Clinicians involved in this research should do it with scientific

rigour. In particular the drug in question should be compared, in a randomized way, with the currently best available alternative treatment. It should also take aspects other than simple efficacy into consideration. One of these aspects is cost consideration.

3.6 Where do research ideas come from?

3.6.1 Searching the medical literature

For an investigator to be able to conceive good research topics, s/he is advised to:

- read the medical literature, including reviews which outline gaps in research;
- attend scientific meetings;
- teach—questions asked by students can often give ideas for research;
- be a team player—ideas can come from colleagues or mentors, in the same or different disciplines;
- acquaint herself/himself with the lines of interest of funding research organizations;
- develop specific areas of scientific interest—it is a good idea to be an expert in a small field, it is better to be a big fish in a small pond than a small fish in a large lake.
- get new ideas out of her/his own previous research;
- be a good observer;
- be imaginative;
- have a sceptical attitude when reading scientific findings—science should not be admired, science should be questioned.

A search of the literature is essential before deciding whether research is worth doing, and what the gaps are that need to be addressed. The current medical literature is a jungle that is not easy to navigate. It is difficult to cope with the information explosion in the literature. There are over 2 million articles published every year in over 20 000 biomedical journals. This has led to the emergence of indexing services and abstracting services. The number of journals that now exist solely to summarize articles probably exceeds 200. While ephemeral literature (literature judged to have a short period of usefulness and only for a small audience) is not normally considered worth indexing or cataloging, it may, however, be important. It includes reports, proceedings of conferences and other types of publication.

English has become the common language of scientific communication and all researchers working in the international arena need to have at least a reading knowledge of it. Computer literacy has now become another requirement, as manual search is being replaced by online search.

The role of libraries has evolved. Modern libraries are no longer repositories of only printed materials. They normally have computerized catalogues of their holdings, filed by subject, author and title. Many college and public libraries are part of a network of libraries. This network expands the holdings of every library, because one library will loan books to other libraries, through an inter-library loan system. Photocopies of articles not available in one library can be requested and sent by fax from another library. A modern library will also provide computer access to resources on the internet, with help from librarians available if needed.

Annex 3 provides a technical note on searching the literature, using the resources of the United States National Library of Medicine (NLM), and the health information available on the internet.

3.6.2 New initiatives for expanding access to the scientific literature

Open access

Open access to scientific information was high on the agenda at the World Summit on the Information Society, held in Geneva in December 2003. Delegates from 176 nations endorsed a Declaration of Principles that included a commitment to "strive to promote universal access with equal opportunities for all, to scientific knowledge, and the creation and dissemination of scientific and technical information, including open access initiatives for scientific publishing" (http://www.biomedcentral.com/openaccess, accessed February 24, 2004). Annex 3 also provides information on organizations which provide free access to scientific journals.

Health InterNetwork Access to Research Initiative (HINARI)

Health problems in developing countries are more likely to be solved by researchers in those countries, who better know the right questions to ask, and who can look for feasible solutions. For this, they need access to the global pool of scientific knowledge. Until very recently, most health institutions in developing countries had little or no access to international scientific journals. The few that were available were often out of date. Institutions could not afford the cost of the subscriptions.

The World Health Organization (WHO) gives high priority to improving access to scientific information. HINARI began as a voluntary partnership between WHO and

five leading publishers—Blackwell, Elsevier (including Harcourt), Springer Verlag, John Wiley and Wolters Klumer—to provide institutions in developing countries with free access to journals. The first phase was launched on 31 January 2001, supplying 68 countries with free access, on the internet, to 1400 journals. A total of 438 institutions in 56 countries have registered, and more than 100 institutions are accessing the journals regularly. The number of institutions is growing, and the number of journals has increased to over 2000 since 18 further publishers have joined HINARI.

In January 2003 access was extended to another 42 middle-income countries. Institutions in these countries must pay US$ 1000 for access to about 2000 electronic journals (which would buy subscriptions to only about three journals at normal prices), and the publishers are donating the revenue to WHO to use for training librarians in using HINARI.

Improved functionality has provided a direct link to the HINARI journals from PubMed (the database for the United States National Library of Medicine). Annex 4 provides information on how to search the literature through HINARI. More information on HINARI is available from the website http://www.healthinternetwork.net.

Eastern Mediterranean Region Virtual Health Sciences Library

The WHO Regional Office for the Eastern Mediterranean started an initiative to link libraries in the Region in a virtual network. The objective of the network is to make available and/or accessible the widest range of health and biomedical literature to potential users in a cost-effective way in the Region. The internet, now available in most Member States in the Region, allows the operation of the network as a virtual network. A core group of libraries have already expressed interest to participate in the network. Researchers can access the services at http://www.emro.who.int/HIS/VHSL/Index.htm.

PubMed Central

This initiative of the United States National Library of Medicine (NLM) provides free online access to the full text of life science research articles (http://pubmedcentral.nih.gov). As a public web-based archive, it offers barrier-free access to peer-reviewed primary research reports in the life sciences, and provides the worldwide scientific community, and users of the World Wide Web in general, the opportunity to search the life sciences literature and retrieve not only article titles and abstracts, but entire research reports free.

PubMed Central can be looked at as a logical extension of MEDLINE, which offers the bibliographic details of articles and their abstracts. It depends on publishers and scientific societies transferring peer-reviewed articles to PubMed Central, which, like

MEDLINE, is funded by the US National Institutes of Health. Its LinkOut capability allows easy navigation to the full text content available by hyperlinking to the hosted content of many publishers of science, technology and medicine

Eastern Mediterranean Region Index Medicus

The Eastern Mediterranean Region Index Medicus project started in 1987 with indexing of the health and biomedical journals published in the countries of WHO Eastern Mediterranean Region from 1984 onwards. The database is now current and as up-to-date as the journals themselves, and can provide a current awareness service to what has been published in the Region. The Index is distributed in three forms: in a print version of the current contents on a quarterly basis; online through the Regional Office web site on the internet (http://www.emro.who.int./library); and in a CD-ROM update on a six-monthly basis.

3.7 Criteria for a good research topic

A good research topic should be feasible (can be done), interesting, novel, ethical and relevant (has an implication). These criteria have been collectively called the FINER formula (Hulley et al., 2001). The investigator can test how good the proposed research question is by using these five criteria.

Feasibility

Before deciding on a research topic, the investigator must be sure that the research can be done and completed. The following are examples of factors to be considered, depending on the category of research.

- It should be possible to recruit the number of subjects required to provide the answer to the research question within the timeframe of the planned research.
- The research facility available to the investigators should have the equipment, supplies and other requirements to undertake the research.
- The investigators must have the required expertise.
- The cost of doing the research must be affordable and the financial resources available.
- The research objectives must not be too many or too ambitious. It is always advisable to establish a single primary objective around which to focus the development of the study plan. This can be supplemented with secondary objectives that may also produce valid conclusions.

Sir Peter Medawar, a British Nobel Laureate, used to describe scientific research as "the art of the soluble", in an analogy to Otto von Bismarck's description of politics as "the art of the possible" (Medawar, 1979). He was careful to point out that he was not advocating the study of easy problems yielding quick solutions. What he meant was that the art of research is about making a problem soluble by finding out ways of getting at it, and by defining research questions that can be answered.

Interest

The research topic must be of interest to the investigators and to the scientific community. If the investigators are not excited about the topic, or cannot get colleagues interested in it, the project is probably not worth doing.

Novelty

It is essential that the investigator is familiar with the up-to-date literature on the planned topic for the research. The research must be expected to contribute new information. Novel does not necessarily mean that the research has not been done before. The prefix "re" in the word research implies searching again. Most good studies are neither original nor simple duplication of other studies. The progress of science is incremental, with knowledge gradually building up from different studies. The question should not be about whether the study has been done before, but whether it will add to the existing body of knowledge. The addition to previous studies may be confirmatory (especially if there was weakness in the original reports), contradictory, or extend previous findings.

Ethics

Ethical issues must be addressed at the early stage of selecting the research topic. Other ethical issues will need to be addressed in planning the research. Some ethical problems may indicate that the research should not be considered from the beginning.

If the research topic involves experimentation on human subjects, the following issues should be considered.

- If the topic is about testing a new therapy or procedure, evidence should already be available to suggest that it can be superior to currently available alternatives.
- Adequate data must be available from animal studies and from studies on a small number of human subjects to confirm safety and to suggest effectiveness, before subjecting patients to a new drug or procedure. The ethically acceptable practice is to step up clinical trials in successive phases, starting first with a small number of subjects, and only moving to the next phase after the successful completion of the previous phase.

- It is unjustifiable to do clinical trials on therapies that are unlikely to become available to people in the country or community. For example, drugs that are likely to be non-affordable or non-marketable should not be tested in a given population. This applies in particular to pharmaceutical company research and to international research.
- The research should not conflict with the society's cultural, moral, religious and legal values.

If the research involves experimentation on volunteer human subjects, for whom the research has no immediate benefit, the research should only be carried out if the information needed is likely to advance scientific knowledge and medical practice, and if the information cannot be obtained otherwise, e.g. through animal experimentation.

Research involving experimentation on animals should be justified. In-vitro biological systems or computer simulation models should be considered, wherever possible, as substitutes to animal research. The animal experiments must be relevant to the advancement of knowledge, or are an essential step before human experimentation.

Relevance

This criterion can be called: the "so-what?" test. For the research to be considered relevant, it must have the potential to advance scientific knowledge, influence clinical management, influence health policy, or guide further research.

References and additional sources of information

Burrill GS. *Biotech 98*: *tools, techniques and transition*. San Francisco, Burrill and Company LLC, 1998: 36.

Commission on Health Research for Development. *Health research: essential link to equity in development*. Oxford University Press, 1990: 13; 20–22.

Cummings SR, Browner WS, Hulley SB. Conceiving the research question. In: Hulley SB, Cumming SR, eds. *Designing clinical research: an epidemiologic approach*. 2nd edition. Philadelphia, Lippincott Williams & Wilkins, 2001: 17–23.

Dawson G, Lucocq B, Cottrell R, Lewinson G. *Mapping the landscape. National biomedical research outputs 1988-95. Policy report No.9*. London, The Wellcome Trust, 1998: 21: 39.

Fathalla MF. Promotion of research in human reproduction: global needs and perspectives. *Human Reproduction*, 1988: 3;7–10.

Global Forum for Health Research. *Monitoring financial flows for health research*. Geneva, Global Forum for Health Research, 2001.

Heath DA. (Quoting Kettering, the automotive engineer). Research: Why do it? In: Hawkins C, Sorgi M. *Research–How to plan, speak and write about it*. Berlin, Springer-Verlag, 1985: 2.

International Committee of Medical Journal Editors. *Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication*. Updated November 2003. (http://www.icmje.org accessed 24/2/2004).

Jefferson T, Demicheli V, Mugford M. *Elementary economic evaluation in health care*. 2nd edition. London, British Medical Journal Books, 2000.

Marcondes CH and Sayao LF. The SciELO Brazilian Scientific Journal Gateway and Open Archives, *D-Lib Magazine,* 2003. 9:1–12 (http://www.dlib.org/march03/macondes/o3marcondes.html accessed 24/2/2004).

Medawar PB. *Advice to a young scientist*. New York, Basic Books, 1979: 18; 47.

Murray CJL, Govindaraj R, Musgrove P. National health expenditures: a global analysis. *Bulletin of the World Health Organization*, 1994, 72: 623–637.

Nass SJ, Stillman BW, eds. *Large-scale biomedical science: exploring strategies for future research*. Washington, DC, The National Academies Press, 2003.

Investing in health research and development: Report of the Ad Hoc Committee on Health Research Relating to Future Intervention Options. Geneva, World Health Organization, 1996.

Roberts RM. *Serendipity: accidental discoveries in science*. New York, John Wiley & Sons, Inc, 1989: 159–164; 123–125; 244.

Chapter 4

Planning the research

4.1 Introduction

After deciding on the research topic, the investigators have to think carefully about the plan of the research. In this process, they consider the options they have about different ways in which the research topic can be investigated, i.e. a research design. In making this choice, they have to weigh two factors. They should try to choose a design that will give most definitive answers about the research topic. But they have to weigh this against the feasibility of doing the study. They have to consider, among other things, their own capabilities, the availability of material or subjects for the research, and the availability of resources. Often, a trade-off has to be made between the ideal and the possible. The best should not be made the enemy of the good.

After deciding on a research design that is appropriate to deal with the research topic and that is feasible, they have to look again at the broad research topic, and define and refine it into a research question which can be answered by the research design. For many studies, this will involve generating a research hypothesis that can be tested.

Among the issues the investigators have to deal with in designing the research is the question of sampling. Since the study cannot include all the target population, they have to depend on the accessible population, and select a sample that is as representative as possible of this population. The size of the sample is an important decision to make. If, on the one hand, the sample is too small, the results obtained will not be reliable, the resources for the research will be wasted and, if human subjects are involved, it would have been unethical to subject them to research that does not give useful results. If, on the other hand, the sample is too large, it prolongs the study and makes it more expensive, with no added scientific value. The investigators also have to give attention to how the study results will be measured, by choosing methods that are reliable and valid.

The design of qualitative research needs different approaches from that of quantitative research. These approaches include observation, in-depth interviews and focus group discussions. If a questionnaire is used to collect information from respondents, there are a number of options for the investigators, and there are guidelines to follow.

Last but not least, planning is the time to think carefully about ethical implications before the study is implemented.

All these topics will be discussed in the next sections. For more detail, the references and additional sources listed for the chapter can be consulted.

4.2 Types of research design

The study type may dictate certain research designs. More commonly, the study objectives can be achieved through a number of alternative designs. The investigators have to select the most appropriate and most feasible design.

Generally, there are two main categories of research design: observational study, and experimental or intervention study. In the observational study, the investigators stand apart from events taking place in the study. They simply observe and record. In the experimental or intervention study, the investigators introduce an intervention and observe the events which take place in the study.

Observational studies

An observational study may be descriptive or analytical. A descriptive study is an observational study that simply describes the distribution of a characteristic. An analytical study is an observational study that describes associations and analyses them for possible cause and effect.

An observational study may be cross-sectional or longitudinal. In a cross-sectional study, measurements are made on a single occasion. In a longitudinal study, measurements are made over a period of time.

A longitudinal observational study may be retrospective or prospective. In a retrospective study, the investigators study present and past events. In a longitudinal prospective study, the investigators follow subjects for future events.

Case–control studies are a type of observational-analytical-retrospective studies over time in which a group of subjects with a specified outcome (cases) and a group without that outcome (controls) are identified. Investigators then compare the extent to which each subject was previously exposed to the variable of interest, such as a risk factor, a treatment or an intervention. Case–control studies are useful for studying rare conditions and conditions with long intervals between exposure and outcome such as, for example, risk of developing neoplasia. In such situations, a prospective study will be difficult. Case–control studies can be efficient and economical, but do not have the strength of evidence of a prospective study.

In clinical and epidemiological research, a longitudinal observational study is usually called a cohort study. The word cohort was the ancient Roman term for a group of soldiers who marched together into battle. The prospective cohort design is generally considered to be the “crème de la crème” of observational methodologies for the following reasons.

- Data are gathered prospectively.
- Recall bias is not a problem (research subjects are not asked to recall past events).
- Time–order relationships are clear (it is easy to decide that an outcome followed, rather than preceded, a possible cause).
- Investigators have much more control on the quality of the data.

There are, however, some drawbacks.

- The biggest single problem of these follow-up design investigations is the loss of valuable information through attrition, due to loss to follow-up, or subjects opting out of the study.
- Subjects may change their behaviour over time.
- A bias can occur if there is unequal surveillance of subjects in the two compared groups, during follow-up.

One of the best examples of a prospective cohort study was initiated by Austin Bradford Hill and Richard Doll, to investigate the relationship between smoking and lung cancer. They followed up 40 000 British doctors who were divided into four cohorts: non-smokers, and light, moderate and heavy smokers. Death was the outcome they recorded. They used both all cause death (any death) and cause specific death (death from a particular disease). Publication of their interim 10 year results in 1964, showed a substantial excess in both mortality from lung cancer and all cause mortality in smokers, with a "dose-response" relation (that is, the more the subjects smoked the greater were their chances of getting lung cancer). The study went a long way in demonstrating that the link between smoking and ill-health was causal rather than coincidental. The 20 year and 40 year results of this momentous study (which achieved 94% follow-up of those recruited in 1951 and not known to have died) illustrate the strength of evidence that can be obtained from a properly conducted cohort study (Doll and Hill, 1964; Doll and Peto, 1976; Doll et al., 1994).

Experimental or intervention studies

In the experimental or intervention study, the investigators test the effect of an intervention on the events taking place in the study. An experimental or intervention study may be controlled or non-controlled. Giving a treatment to a patient or group of patients and finding that the treatment works gives only preliminary and non-definitive information. We do not know what would have happened if no treatment or a different treatment was given. For a more definitive answer, we need a "control" group of patients who do not get the treatment under study.

Hawthorne effect: In the late 1920s, a group of researchers at the Western Electric Hawthorne Works in Chicago were investigating the effects of lighting, heating and other physical conditions upon the productivity of workers. Much to the surprise of the researchers, the productivity of the workers kept improving even when the actual physical conditions were not improved. The Hawthorne effect can be manifested in clinical research settings. Even "inert" treatments might result in significant improvements in the patient's condition (Polgar and Thomas, 2000).

A controlled experimental study may be randomized or non-randomized. In testing the outcome in a group of patients who receive the treatment and another group who do not, we are still not sure whether any difference observed is because of the treatment or because the characteristics of the patients in the two groups were different. The best way to be sure is to randomize the allocation of patients to either treatment or to no treatment.

Randomized controlled trials are intervention studies characterized by the prospective assignment of subjects, through a random method, into an experimental group and a control group. In a clinical trial, the experimental group receives the drug or treatment to be evaluated, while the control group receives a placebo, no treatment, or the standard of care. Both groups are followed for the outcome(s) of interest. Randomization is the most reliable method to ensure that the participants in both groups are similar as far as possible with respect to all known or unknown factors that might affect the outcome. With randomization, only chance determines the assignment of subjects to study groups. Random allocation does not mean haphazard allocation. It is a carefully planned method of assigning subjects to similar groups. If important risk factors can be identified at the outset, subjects may be grouped or stratified prior to assignment. Whenever it is ethical and practical, a randomized design should be considered in controlled intervention studies.

Controlled trials without randomization are intervention studies in which allocation to either experimental or control group is not based on randomization, making assignment subject to possible biases that may influence study results.

A crossover study is a special design of controlled intervention study that is sometimes used in drug trials. In this design, half of the participants are randomly assigned to start with the placebo and then switch to active treatment, while the other half does the opposite. It has the advantage of reducing the number of subjects required, since each subject serves as both an experimental subject and a control. It also decreases the biological variability inherent in comparing different subjects by comparing each subject with himself or herself. It has the disadvantage of increasing the duration of the study. There will also be a problem if the treatment has a carry-over effect after it is stopped.

A before-and-after study is a method of control in which results from experimental subjects are compared with outcomes from patients treated before the new intervention was available. These are called historic controls.

A randomized controlled trial may be blinded if participants in the trial are likely to change their behaviour in a systematic way that may influence the outcome of the study when they are aware of which intervention they receive. (Ophthalmologists prefer the term “masking” to the term “blinding”.)

Blinding can take place at a number of levels. At one level, those responsible for assigning the subjects to groups do not know to which group the next subject will be assigned. In another level, research subjects are also not aware of which intervention they are receiving. Then, health workers who take care of patients in the study may not be allowed to know what treatment the different patients are receiving. Lastly, researchers who assess the outcome are also not able to distinguish the subjects in the different groups.

The term double-blind is used when neither researchers not subjects are aware of the type of intervention. A trial in which there is no attempt at blinding may be called open or open label.

The Rosenthal effect: Rosenthal and his colleagues in 1976 performed an experiment involving the training of two groups of rats in a maze learning task. A bright strain and a dull strain of rats especially bred for the purpose were trained by undergraduate student experimenters to negotiate the maze. After a suitable training interval, the relative performances of the groups were compared. Not surprisingly, the bright strain significantly outperformed the dull strain. What was surprising, however, was that the two strains were actually not different. The two groups of rats were actually genetically identical. The researchers had deceived the student experimenters for the purposes of the study, and the students’ expectations of the rats had resulted in different methods of treatment, which had affected the rats’ learning ability. These results have been confirmed time and time again in a variety of experimental settings, and with a variety of subjects. They confirm the need for blinding (Polgar and Thomas, 2000).

4.3 Selecting a research design

A research question may be answered by more than one research design. The researcher has to select the appropriate design for the particular study. All types of research design have a place, and all have advantages and disadvantages. But not all types of design are always possible for a particular study.

For example, the investigators may want to study if there is a relationship between post-menopausal hormone replacement therapy and subsequent development of uterine

endometrial carcinoma. The investigators can design an observational study or an experimental study. If the decision was for an observational study, the investigators may do a descriptive study or an analytical study.

For a descriptive study, they will review the clinical records of all patients diagnosed as having endometrial carcinoma. They will look for a history of post-menopausal hormonal therapy. This study will be useful but cannot be definitive. It shows whether further study is needed to confirm or refute the impression gained from the descriptive study. The information about the strength of the association will also help in the design of further analytical studies. The finding that many of the women who developed endometrial carcinoma had a history of homonal therapy cannot lead to any conclusion. It may simply mean that this therapy is widely used in the community, both by women who develop and who do not develop endometrial carcinoma. This shows the need for further studies.

For an analytical study, the investigators may do a cross-sectional study or a longitudinal study. In a cross-sectional study, the investigators may study all post-menopausal women admitted to hospital over a defined time period. For each woman, they record whether she received or did not receive hormonal therapy, and whether she had or did not have endometrial cancer. The advantage of this study is that it can be done rapidly. It gives more evidence than the simple descriptive study. However, the two groups of patients may not be comparable.

In a longitudinal observational study, the investigators may do a prospective study or a retrospective study. For a prospective study, a cohort of two groups of post-menopausal women is followed up: one group already receiving hormone replacement therapy and another matched group not receiving this therapy. For a retrospective study, a case–control design can be selected. A group of women who have recently developed endometrial cancer (cases) and a group of women with similar characteristics and did not develop endometrial cancer (controls) are identified. The use of hormone replacement therapy in each woman in the case group and in the control group is determined to assess exposure history. The advantage is that the study can be done relatively quickly. The disadvantage is that the two groups may still not be completely similar. Other variables may influence the outcome and may be difficult to exclude.

If the investigators decide on an experimental or intervention study, they may select a randomized or a non-randomized design. In a randomized controlled study, post-menopausal women identified from a population are randomly assigned either to a study group that will receive hormone replacement therapy or to a control group that will be prescribed a placebo. Both groups will then be followed prospectively to determine how many in each group will develop endometrial cancer. This study, if successfully conducted, will provide a more definitive answer to the research question. However, it will raise ethical concerns. Additional difficulties are the large sample size needed

because of the relatively low incidence of the disease, the long follow-up because of the long latent period before the development of the disease and the possibility of poor compliance or loss to follow-up. Alternatively, a non-randomized controlled design may be considered. This may be easier, will allow women to make an informed choice but there will be a need to consider other possible variables that may influence the outcome, since the two groups may not be similar.

Different types of research design are not considered equal in the strength of evidence they provide. In the traditional hierarchy of evidence, randomized controlled studies are generally ranked high, followed by cohort and case–control studies, while observational descriptive studies are ranked at a lower level. The investigators may, however, not be able to select the design that gives a high level of evidence, because it will not be feasible to do, or will not be ethical to do. In this case, their selection of another design will be acceptable and justified.

4.4 Defining and refining the research question

In order to develop the research design, the research topic often has to be changed to a research question, and the research question should be defined and refined so that it can be answered with precision.

If we take again the example of the relationship between post-menopausal hormone replacement therapy and subsequent development of endometrial carcinoma, the research question will be: Does post-menopausal hormone replacement therapy predispose women to develop endometrial cancer?

For the purpose of the research design, the question needs to be better defined. The hormone replacement therapy should be specifically stated. Is it oestrogen alone or oestrogen in combination with a progestagen? Does the duration of therapy need to be defined as, for example, more than one year? Should the diagnosis of endometrial cancer be specified as histologically confirmed?

For the purpose of the research design, the question also needs to be refined. The research will only be able to determine if there is an association or not. The refined question should therefore be: Is post-menopausal hormone replacement therapy, as defined, associated with a subsequent increased risk of endometrial cancer? The association, if found, will need an explanation, but cannot be taken as meaning causation without further questioning.

If we take another example for a research question, “Is passive smoking harmful to the foetus?” the question needs to be better defined and also refined.

The first definition is about passive smoking. What arbitrary definition should be accepted, in terms of number of cigarettes smoked every day? This is called an

operational definition. The operational definition is a statement of how the researchers in a particular study choose to measure the variable in question. It should be unambigious and have only one possible interpretation. Another definition that needs to be made is about effect on the foetus. Could it be defined as effect on intrauterine growth retardation, biophysical profile as determined by ultrasound examination, low birth weight, or the condition at birth (Apgar score for example)? Choice of any of these outcomes will affect the size of the sample to be studied. It will also need control for other variables, which will have to be excluded.

After considering these definitions, there is a need to refine the research question to be, for example, "Are the children born to women whose husbands smoke more than 20 cigarettes a day, of lower birth weight than children born to women whose husbands do not smoke"? This research question is now suitable to turn into a specific hypothesis that can provide a good basis for the development of an appropriate design and calculation of the sample size needed.

4.5 Generating the research hypothesis

If the research question is concerned with relationships between observations or variables, a research hypothesis will need to be developed. The research hypothesis is a tentative statement that can be tested by a scientific research design. Using the previous two examples, the research hypotheses could be as follows.

- Post-menopausal women who received hormone replacement therapy, of a specified type and duration, are more likely to develop endometrial cancer than post-menopausal women who did not receive such therapy.
- Children born to women whose husbands smoke more than 20 cigarettes a day are of lower birth weight than children born to women whose husbands do not smoke.

4.6 Study sample

4.6.1 Target population and accessible population

An important issue in the design of the research is the question of sampling. Ideally, the study design should include all the target population. The term population in scientific methodology refers to the material of the study, whether it is human subjects, animals or inanimate objects. Including all the target population is generally not possible, because of the large numbers, the cost and the time. A subset of the population is studied instead, from which conclusions (or inferences) are drawn as applying to the target population. The sample has to be selected to be as representative as possible of the target population, and in enough numbers to provide valid answers.

The population census is an example of a study in which all members of the population are studied. Even in a small country, it is a very major undertaking. Because of its expense, it is normally carried out every 10 years or so. It normally takes several years to analyse the results. Some countries do an interval census based on subsets of the population in between.

An illustrative example of sampling from another field is that of polls before parliamentary or presidential elections where specialized agencies make predictions based on a relatively small sample representative of the population. Since opinions of voters vary with time before the election, these samplings are commonly done periodically. On the day of the election, samples of exit polls are often accurate in predicting the outcome of the election.

Instead of the "target population", the investigator often depends on the "accessible population". The accessible population must be representative of the target population, in order to draw conclusions about the target population. If we take the above example of voter opinions, a polling agency may use the telephone book as the accessible population from which the sample is drawn. This will be acceptable in a country where practically all people have telephones. It will not, however, be representative in a country where a large segment of the potential voters are not reachable by telephone. This does not necessarily mean that the polling should not have been done in this way. The result, however, should be presented as reflecting the opinion of a segment of the target population who are accessible by phone, and not necessarily representing the whole target population.

In health research, the clinic or hospital may provide the accessible population. This, however, does not necessarily represent the community if not everyone goes to the clinic or hospital for the condition in question. This does not mean that clinic or hospital studies should not be done. They provide useful information but the results should not be presented as reflecting the results for all people who have the condition.

4.6.2 Types of sampling

The sample selected from the accessible population should be representative of the accessible population. It should accurately reflect the characteristics of the population from which it is drawn. It should be a miniaturized representation of the accessible population.

Random sampling is not haphazard sampling. It is sampling done in a systematic way to ensure, as far as possible, complete objectivity in the selection of the sample. Random sampling is a way of ensuring that all members of the population have an equal chance of being selected. It does not guarantee that the sample will not be different in characteristics from the accessible population. Rather, it eliminates a possible reason that they should be different.

As discussed in section 4.3, random assignment is important when two interventions or more are compared. It minimizes group differences due to biased selection. Randomization was commonly done manually using a table of random numbers. Now, it is usually done using a computer program.

Stratified random sampling is a special type of sampling to ensure that all subgroups in the accessible population are represented in the sample. This is particularly important if certain subgroups are present in small numbers in the population, or are important to be included. In stratified random sampling, key subgroups are defined, for example by sex, social class, income groups, geographic locations, etc. and samples are drawn at random from each of these "strata". The computer program can be adjusted to draw disproportionately from one or more groups, to ensure their adequate representation.

Cluster sampling is another way of random sampling. It is based first on the random selection of certain subgroups, from which the sample can be taken. For example, in a community survey certain streets or blocks are selected at random first. Then a random sample is selected from each randomly selected cluster. In a health services study, a number of districts are randomly selected. Then a random sample of health service units is selected from each.

Systematic sampling is done by a simple periodic process, for example selecting every second or third patient.

Consecutive sampling involves taking every subject who presents herself/himself over a specified time period. These are not strictly random techniques, but they avoid bias in the selection.

4.7 Sample size

The desired sample size is now easily calculated with the help of computer statistical programs, but the principles underlying the calculation, and the limitations must be clearly understood by investigators.

It is not necessarily true that the bigger the sample, the better the study. Beyond a certain point, an increase in sample size will not improve the study. In fact, it may do the opposite, if the quality of the measurement or data collection is adversely affected by the large size of the study. It is also better to ensure that the sample is representative, rather than being very large.

The statistical concept behind calculation of the desired sample size is simple. When we study a representative sample, we aim to generalize from the sample findings to the population from which the sample was drawn. We cannot be completely certain about this. Unless we study the whole population, the sampling error cannot be brought down to zero. Analytical statistics helps us to define the degree of probability that a finding, a

difference or a relationship can be generalized to the population from which the sample is drawn. This is called the statistical significance of the finding. The size of the sample is an essential element in making this statistical probability calculation. The smaller the size of the sample, the less likely that the findings can be generalized. For calculating the desired sample size before beginning the study, we do the exercise in reverse. We decide beforehand on a level of probability or uncertainty that we are willing to accept for the study, and then we find the desired sample size to provide that level of statistical probability. Traditionally, most studies set this level of statistical significance at 0.05, that is accepting a chance of 5% of finding an association that is not actually there. It must be recognized, however, that this value is arbitrary, and other values can and are sometimes used. In general, the investigator should aim for a lower probability of error when it is particularly important to avoid making a false-positive statement about a finding.

When the study is designed to find a difference or an association, we may not find a difference or an association. In this case, we still want to calculate statistical probability that we may have missed a difference or an association that exists in the population, but was not found in the sample. This so-called statistical power of the study depends also on the size of the sample. The larger the size of the sample, the higher the power of the study. For calculating the sample size before the study begins, the investigators have to make a decision on the level of statistical power they are willing to accept for the study. Traditionally, most studies set statistical power at 0.80, which is accepting a 20% chance of missing a difference or an association that is actually there. It must be recognized, however, that this value is arbitrary, and other values can and are sometimes used. In general, the investigator should aim at a higher statistical power when it is particularly important to avoid false-negative error.

Although a statistician may do the necessary exercise to determine the sample size, s/he can only do it with guidance from the investigator on the level of uncertainty that is considered acceptable. In addition, calculation of the statistical significance and statistical power has to take into consideration some characteristics of the data. These characteristics will thus also be needed for calculating the sample size. Since the data are not available before the study begins, the investigators will have to make some assumptions about the data, and provide these assumptions to the statistician to be able to calculate the desired sample size. The procedure for estimating sample size is not as precise as investigators may be led to think. One such assumption is about the prevalence, incidence or frequency of the condition or event. If the rate of the event is large, statistical power will be high with a smaller number of cases. If the event is rare, a larger sample size will be needed. Also, the larger the variation in the data, the larger the sample size that will be needed to achieve a certain level of statistical significance. For sample size to be calculated, we thus need to make a prior estimate of the frequency of the condition under study, and the degree of variations in the data. Some information may be available

from previous studies to guide the estimates. If not, it is up to the investigators to come up with a tentative estimate which the statistician can use.

The effect size in a study refers to the actual size of the differences observed between groups or the strength of relationships between variables. The likelihood that a study will be able to detect an association between a predictor and an outcome variable depends on the magnitude of the association we decide to look for. Large sample sizes are needed to detect small differences. The choice of effect size is difficult and arbitrary, but it must be set beforehand and must make a meaningful difference. The rule is that the smaller the difference you wish to detect, the larger the sample size needs to be. In designing a study, the investigator chooses the size of effect that is considered important.

In making the final estimation of the sample size, factors such as dropouts, attrition and loss to follow-up should also be accounted for. If the calculated sample size proves to be larger than can be practically obtained, the investigators have a number of options: to increase the effect size they look for; to decrease the power of the study; to modify the design; or to give up the study.

4.8 Measurement

An important question in the research design is the decision on how measurements are made to ensure reliability and validity. Reliability means that the observer repeating the test, or someone else using the same method should be able to obtain the same findings. Validity means that the measurement should actually represent what it is intended to measure.

To ensure reliability or reproducibility of the results the following should be considered.

- Measurements made should not vary by observer or between observers (intra- and inter-observer consistency).
- Instrument or laboratory variability should be taken into consideration.
- Subject variability should be considered if measurements vary according to the time they are made, for example, fasting or after meal, time of the day, or day of the menstrual cycle.

Intra-observer and inter-observer or rater reliability are important issues in measurement. In a study to document them, 29 biopsy slides with suspected Hodgkins disease were presented to three pathologists over an 11-month period (Coppleson et al., 1970). The specimens were unlabelled and over the year of the study were presented on two occasions to each of the three observers. The three observers disagreed with themselves on seven, eight and nine occasions, out of the 29. Overall inter-rater

agreement was calculated at 76% or 54%, according to the particular diagnostic feature described.

Obtaining the same result by the same and different raters ensures reliability and reproducibility, but does not mean validity. The test, itself, may not be accurate in measuring what it is intended to measure. This is particularly apparent in diagnostic tests, as will be discussed in more detail in Chapter 9. The test may be sensitive in detecting people with the disease, but not very specific in excluding people without the condition, or vice versa. To test for validity of the measurement, it has to be compared to a "gold standard". If for example, we are using a diagnostic test as an indicator of breast cancer, it should be compared to the gold standard of a breast biopsy.

4.9 Planning qualitative research

The above sections dealt with planning quantitative research. Qualitative research needs other approaches (Ulin et al., 2002).

One way to keep the design focused on the research problem is to develop a conceptual framework. A conceptual framework is a set of related ideas behind the research design. A conceptual framework helps to outline the research questions, and provides a context for understanding the research.

Three main methods are commonly used in qualitative research: observation, in-depth interviews and group discussion. The investigator has to select which method would be more appropriate to answer the research question, or may use more than one method. The researcher in these different designs plays the role of observer, interviewer or group moderator.

Observation

Depending on the objective of the study, observation can be made from an outsider or insider perspective, or somewhere in between. Outsider observers maintain a distance. Insider observers interact.

As an example of an outsider observation study, the investigator may observe the quality of health care delivery in a clinic, health centre or a pharmacy. A special type of observation study, called "time and motion study" is used to study how health workers use their time. The researcher observes what a health worker is doing over a defined sample of time. S/he may use a beeper that goes off every number of minutes and a checklist to record activities.

A special form of observation is the so-called "mystery client" technique. It is used particularly in client–provider studies where the presence of an outside observer might change the provider's customary behaviour. Trained data collectors present as simulated

clients. The deceptive nature of this technique raises ethical concerns. The decision to use the technique should be made only after careful reflection on the ethical implications. Informed consent may be obtained from the health service to use the technique at unannounced times over a period of time, for example several months.

In participant observation, the investigator interacts. S/he may, for example, ask clients about their perceptions of the health service.

In-depth interviews

Intensive one-on-one interviewing is a classical method in qualitative research. Different from quantitative studies based on a structured questionnaire, the in-depth interview is more of a social encounter, with questions flowing from the answer of the respondent, as a follow-up to the answer, or to probe further into the answer. Open-ended questioning is a basic tool in qualitative research. The interview may take the form of an informal conversation with little or no preparation and sequencing of questions. Alternatively, a topic guide or outline may be used to help in focusing the interview, but without pre-structuring the questions. A pre-determined set of open-ended questions is, however, the most standardized approach for in-depth interviews.

Focus groups

Focus group discussions are the method used when information and insights will be better gained from the interaction of a group than from in-depth interviews with individuals. The two methods may complement each other. A focus group discussion is not a group interview. It is based on the exchange of information, ideas and views among the participants themselves. The researcher is playing the role of a moderator, and not an interviewer. In recent years, focus group methodology has been increasingly used. Certain guidelines need to be observed.

The group should be relatively homogeneous, for example in age and sex and sociocultural background. Anonymity among participants may be desirable, if people feel more comfortable to talk freely with strangers than with people they know and will meet again.

For most purposes, groups of eight to ten participants are adequate for a good and manageable discussion. As to the number of groups, it is generally advised to have at least two groups for each defining demographic variable. If, for example, sex is the variable, two women and two men groups will be needed.

A two-hour discussion is likely to generate 25 to 40 pages of transcript. The role of the moderator is to create a comfortable climate for open exchange, stimulate discussion, keep the discussion focused, and encourage everyone to participate. The moderator should not allow one or two vocal individuals to dominate the discussion.

The rapporteur or note-taker should be recording what people say, but should also be aware of body language.

4.10 A note on questionnaire design

A questionnaire is a document designed for the purpose of seeking specific information from the respondents.

The questionnaire may be self-administered or administered by interviewers. The self-administered questionnaire approach is cheap, less susceptible to interviewer bias and can be administered by mail. At the same time, the rate of non-response may be high, and may bias the results. Also, answers may be incomplete.

There are two major question formats: the open-ended and closed-response types. In a closed-response question, the respondent is provided with a list of pre-determined response options. Open-ended questions elicit more detailed responses, but the responses require more effort to encode for data analysis. A questionnaire may include both question formats.

Closed-response questions may be used to elicit attitudes of the respondents to a certain statement. Two formats can be chosen (Polgar and Thomas, 2000). In the Likert-type format, the respondent chooses from among: strongly agree, agree, undecided, disagree, strongly disagree. In the forced-choice format, responses are limited to: strongly agree, agree, disagree, and strongly disagree. This format does not allow an undecided answer.

Questions should be well worded to avoid any ambiguity. Jargon should not be used. Questions should not be phrased in a way that influences the response in one direction or another. The questionnaire should always be pre-tested in a pilot study before the main survey. Interviewers should be trained to make sure that the questionnaire is administered in a uniform way.

A questionnaire typically includes the following components:

- an introductory statement by the interviewer to introduce herself/himself and explain the purpose of the questionnaire; the respondents should also be informed about the confidentiality of their responses;
- demographic questions to collect relevant information about the background of the respondent;
- factual questions;

- opinion questions: opinion questions require reflection; it is generally easier for the respondent to answer factual questions; putting the factual questions first serves as a "warm up" to the opinion questions;
- closing statement by the interviewer to thank the respondents, and where appropriate to ask if s/he wants to provide any additional comment.

A method commonly used to test for reliability in results obtained by questionnaires is to look for internal consistency, that is the extent to which the responses on different questions correlate with each other. If they tend to be highly correlated with each other, then the test is said to be internally consistent. The computer programme can be built up to detect inconsistency.

There is a tendency among investigators to put too many questions. This has been encouraged by the introduction of computer-assisted analysis. Information collected in a questionnaire should be based on and limited to the objectives of the study.

4.11 A note on research in health economics

All methods of economic evaluation in health care have one principle in common: they examine one (or more) possible interventions and compare the costs of inputs or resources necessary to carry out such interventions with their effects or economically assessed benefits (Jefferson et al., 2000).

In economic evaluation, the cost of an illness generally includes:

- direct costs, which are costs borne by the health care system, community and patients' families in addressing the illness (for example, diagnosis or treatment costs);
- indirect costs, which may be tangible or intangible; indirect tangible costs are mainly productivity losses, caused by the disease condition, and borne by the individual, family, society, or by the employer; indirect intangible costs include the costs of pain, grief and suffering, and the loss of leisure time.

In economic evaluation, resources are estimated as all inputs into health service production, including time, goods, equipment, buildings, specialized knowledge, etc.

Cost-benefit analysis and cost-effectiveness analysis are related analytical methods that compare health care practices or techniques in terms of their relative economic efficiencies in providing health benefits. In a cost-effectiveness analysis, the net monetary costs of a health care intervention are compared with some measure of clinical outcome or effectiveness, such as cases of disease avoided, cases identified in screening procedures, life years gained, or deaths avoided. Cost-benefit analysis compares monetary costs to estimated monetary benefits of an intervention.

Cost-effectiveness analysis is frequently nested within a randomized controlled trial. It is particularly valuable when the compared interventions have widely differing costs or resource consequences. Competing interventions in the trial may show little difference in outcome. The addition of the economic perspective offers a further dimension of evaluation. Prospective economic data collection alongside a trial allows the evaluation to be based on reliable estimates of effectiveness.

4.12 Ethics in research design

4.12.1 Categories of health research

From an ethical standpoint, four categories of health research can be distinguished.

- Research involving human experimentation: This is the research category that raises most ethical concerns. Under this category, two types of medical research can be distinguished: a) research of therapeutic or diagnostic nature that is carried out on patients who may expect a potential benefit from their participation; and b) research of a purely scientific nature for which human subjects volunteer to advance medical science but will not draw any therapeutic or diagnostic benefit. Ethical safeguards are most needed in this category.
- Research involving human subjects but not experimentation: Epidemiological and field studies, as well as qualitative research, fall under this category. Although no experimentation is involved, such studies can be intrusive on the individual's privacy and even on communities.
- Research involving experimentation on animals: Ethics in this category has been receiving increasing attention recently.
- Research not involving human subjects or animal experimentation: This category of research would still be bound by ethical principles that cover research in general, medical and non-medical.

4.12.2 Ethics in research design involving experimentation on human subjects

All research involving human subjects should be conducted in accordance with the ethical principles contained in the current version of the World Medical Association Declaration of Helsinki (Annex 1). All individuals involved in the conduct of any clinical trial must be fully informed of and comply with ethical principles, including beneficence, non-maleficence and respect.

The principle of beneficence implies that:

- a scientific and technically sound design is an ethical requirement; a design that will not provide the answer to the research question is ethically unacceptable, as the patients will be subjected to an unnecessary process;
- the sample size is adequate to provide statistically valid results, but is not larger than is necessary to provide the answers.

The principle of non-maleficence implies that:

- any potential risks are properly evaluated and balanced with potential benefits, are minimized in every way possible, including adequate screening for contraindications, and are carefully monitored;
- where adverse effects are encountered, adequate treatment is provided.

The principle of respect implies that:

- participants are fully informed and give their free consent to participate in the trial;
- research trials on children and persons with mental disability are limited to disease conditions specific to them and the informed consent of parents or a guardian is obtained;
- confidentiality is adhered to.

Confidentiality is an ethical obligation in the practice of medicine. Since in research, information is likely to be handled by other people involved in the research, steps should be taken to ensure the confidentiality of the records either by limiting access or by replacing patient identification with code numbers.

A number of ethical considerations apply when a new therapy is being tested on patients, according to the principle of "do no harm" or non-maleficence.

- Pre-clinical studies that provide sufficient documentation of the potential safety of the pharmaceutical product should be available.
- Information about manufacturing procedures should establish that the product is of suitable quality.
- The data available should be appropriate to the phase, size and duration of the trial.
- Data from previous and ongoing clinical trials should be compiled before the trial.
- The investigators should be well qualified and the trial site adequate.

- All parties involved in a clinical trial should comply fully with the existing national regulations or requirements.

4.12.3 Epidemiological, field and qualitative studies

This research is based mostly on observation, and generally requires no intervention more invasive than asking questions and carrying out routine medical examinations and, sometimes, laboratory tests or X-ray examinations. Such studies do not carry physical risks for the research subjects. However, they can be intrusive. Psycho-social harm may be as or more meaningful to the person than physical harm. Ethical considerations include free informed consent, confidentiality and beneficence.

The principle of free informed consent implies that individual subjects should understand and agree to the reasons for collecting the information. In large community surveys, the community must also agree to the study.

The principle of confidentiality implies that information gathering in qualitative research is based on mutual trust. This trust will be seriously breached by any possibility of break of confidentiality. Information collected about subjects in field studies is generally classified as linked or unlinked (CIOMS, 1991). Unlinked information is information which cannot be linked, associated or connected with the person to whom it refers. Confidentiality here is not at stake. Linked information may still be anonymous, if it is linked to the person by a code or other means, and the investigator cannot know the identity of the person. In other cases, strict adherence to confidentiality should be maintained.

The principle of beneficence implies that:

- The individual has a right to be informed of any health condition revealed during the study, and should be helped to get the appropriate care.
- The community has a right to be informed about the outcome of the study, and any potential implications.
- The investigators have the ethical obligation to play an advocacy role to improve the health condition of the community based on the results of the study.
- Local personnel should be utilized, as far as possible, and they should be trained in the required skills. An ethically conducted epidemiological or field study should leave something behind in the community in which it was conducted. So-called “safari research” should be discouraged.

4.12.4 Ethics in research designs involving experimentation on animals

The animal model chosen must be relevant to the human. The information must be applicable to the human.

The minimum number of animals should be used. Experiments should be designed with proper calculation of the size of the animal sample needed to answer the research question or test the research hypothesis. No more than the minimal number of animals should be used, but a sufficient number of animals should be used to provide a scientifically valid conclusion.

References and additional sources of information

Browner WS, et al. Getting ready to estimate sample size: hypotheses and underlying principles. In: Hulley SB, Cummings SR, eds. *Designing clinical research: an epidemiologic approach*, 2nd edition. Philadelphia, Lippincott Williams & Wilkins, 2001: 51–62.

Byrne DW. *Publishing your medical research paper*. Baltimore, Lippincott Williams & Wilkins, 1998: 5–44.

Carey SS. *A beginner's guide to scientific method*, 2nd edition. New York, Wadsworth Publishing Company, 1998.

Coppleson L, Factor R, Strums S, Graff P, Rappaport H. Observer disagreement in the classification and histology of Hodgkin's disease. *Journal of the National Cancer Institute*, 1970, 45: 731–740.

International guidelines for ethical review of epidemiological studies. Geneva, Council for International Organizations of Medical Sciences, 1991.

International guiding principles for biomedical research involving animals. Geneva, Council for International Organizations of Medical Sciences, 1985.

Hulley SB, Martin JN, Cummings SR. Planning the measurements: precision and accuracy. In: Hulley SB, Cummings SR. eds. *Designing clinical research: an epidemiologic approach*, 2nd edition. Philadelphia, Lippincott Williams & Wilkins, 2001: 37–49.

Devers KJ, Sofaer S, Rundall TG, eds. Qualitative methods in health services research: a special supplement to HSR. *Health Services Research*, 1999, 34 (5) Part II: 1083–1263.

Doll R, Hill AB. Mortality in relation to smoking: ten years' observation of British doctors. *British Medical Journal*, 1964, 1:1399–1414; 1460–1467.

Doll R, Peto R. Mortality in relation to smoking: 20 years' observation on British doctors. *British Medical Journal*, 1976, 2: 1525–1536.

Doll R, Peto R, Wheatley K et al. Mortality in relation to smoking: 40 years' observations on male British doctors. *British Medical Journal*, 1994, 309: 901–911.

Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet* 2002, 359: 57–61.

Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. *Lancet*, 2002, 359: 145–49.

Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet*, 2002, 359: 248–52.

Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. *Lancet*, 2002, 359: 341–45.

Jefferson T, Demicheli V, Mugford M. *Elementary economic evaluation in health care*, 2nd edition. London, British Medical Journal Books, 2000.

Neame R, Kluge E-H. The impact of informatics. Computerisation and health care: some worries behind the promises. *British Medical Journal*, 1999, 319:1295.

O'Brien PMS, Pipkin FB. eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999.

Polgar S, Thomas SA. *Introduction to research in the health sciences*, 4th edition. London, New York, Churchill Livingstone, 2000: 62; 63; 107–114.

Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet*, 2002, 359: 431–34.

Schulz KF, Grimes DA. Generation of allocation sequences in randomized trials: chance not choice. *Lancet*, 2002, 359: 515–519.

Schulz KF, Grimes DA. Allocation concealment in randomized trials: defending against deciphering. *Lancet*, 2002, 359: 614–618.

Schulz KF, Grimes DA. Blinding in randomized trials: hiding who got what. *Lancet*, 2002, 359: 696–700.

Schulz KF, Grimes DA. Sample size slippages in randomized trials: exclusions and the lost and wayward. *Lancet*, 2002, 359: 781–85.

Schulz KF, Grimes DA. Unequal group sizes in randomized trials: guarding against guessing. *Lancet*, 2002, 359: 966–70.

Swinscow TDV, Campbell MJ. *Statistics at square one*. 10th edition. London, BMJ Books, 2002.

Ulin PR, Robinson ET, Tolley EE, McNeill ET. *Qualitative methods: A field guide for applied research in sexual and reproductive health*. North Carolina, Family Health International, 2002.

Varkevisser C, Pathmanathan I, Brownlee. *Designing and conducting health systems research projects. Volume 1: Proposal development and field work. Volume 2: Data analysis and report writing*. Ottawa, International Development Research Centre, 1995.

Wingo PA, Higgins JA, Rubin GL, Zahniser SC. *An epidemiologic approach to reproductive health*. Geneva, World Health Organization, 1994 (WHO/HRP/EPI/1994).

Health research methodology. A guide for training in research methods. 2nd edition. Manila, World Health Organization Regional Office for the Western Pacific, 2001.

Chapter 5
Writing the research protocol

5.1 Introduction

After proper and complete planning of the study, the plan should be written down. The protocol is the detailed plan of the study. Every research study should have a protocol, and the protocol should be written.

The written protocol:

- forces the investigators to clarify their thoughts and to think about all aspects of the study;
- is a necessary guide if a team (not a single investigator) is working on the research;
- is essential if the study involves research on human subjects or is on experimental animals, in order to get the institution's ethical approval;
- is an essential component of a research proposal submitted for funding.

During the process of the development of the protocol, investigators can and should try to benefit from the advice of colleagues and experts in refining their plans. But once a protocol for the study has been developed and approved, and the study has started and progressed, it should be adhered to strictly and should not be changed. This is particularly important in multi-centre studies. Violations of the protocol can discredit the whole study. If the violations are minor, at least that part of the study should be excluded from the analysis.

An additional step, after writing the protocol, particularly in large studies with teams of investigators, is to develop what may be called the operations manual for the study. This will include detailed instruction to the investigators to assure a uniform and standardized approach to carrying out the study with good quality control.

A well-thought out and well-written protocol can be judged according to three main criteria.

- Is it adequate to answer the research question(s), and achieve the study objective?
- Is it feasible in the particular set-up for the study?

- Does it provide enough detail that can allow another investigator to do the study and arrive at comparable conclusions?

The protocol should outline the rationale for the study, its objective, the methodology used and how the data will be managed and analysed. It should highlight how ethical issues have been considered, and, where appropriate, how gender issues are being addressed.

5.2 Format for the protocol

The research protocol is generally written according to the following format.

- Project title
- Project summary
- Project description:
  - Rationale
  - Objectives
  - Methodology
  - Data management and analysis
- Ethical considerations
- Gender issues
- References

Project title

The title should be descriptive and concise. It may need to be revised after completion of the writing of the protocol to reflect more closely the sense of the study.

Project summary

The summary should be concise, and should summarize all the elements of the protocol. It should stand on its own, and not refer the reader to points in the project description.

Project description

Rationale

This is equivalent to the introduction in a research paper. It puts the proposal in context. It should answer the question of why and what: why the research needs to be done and what will be its relevance. A brief description of the most relevant studies published on the subject should be provided to support the rationale for the study.

Objective(s)

Specific objectives are statements of the research question(s). Objectives should be simple (not complex), specific (not vague), and stated in advance (not after the research is done). After statement of the primary objective, secondary objectives may be mentioned. Young investigators are advised to resist the temptation to put too many objectives or over-ambitious objectives that cannot be adequately achieved by the implementation of the protocol.

Methodology

The methodology section has to be thought out carefully and written in full detail. It is the most important part of the protocol. It should include information on the research design, the research subjects, interventions introduced, observations to be made and sample size.

- Research design: The choice of the design should be explained in relation to the study objectives.
- Research subjects or participants: Depending on the type of the study, the following questions should be answered:
  - What are the criteria for inclusion or selection?
  - What are the criteria for exclusion?
  - In intervention studies, how will subjects be allocated to index and comparison groups?
  - What are the criteria for discontinuation?
- Interventions: If an intervention is introduced, a description must be given of the drugs or devices to be used, and whether they are already commercially available, or in phases of experimentation. For drugs and devices that are commercially available, the protocol must state their proprietary names, manufacturer, chemical composition, dose and frequency of administration. For drugs and devices that are still in the experimental stage (or that are commercially available but are being used for a different indication or in a different mode of administration), additional information should be provided on available pre-clinical investigations in animals and/or results of studies already conducted on humans. In such cases, the approval of the drug regulatory agency in the country is generally needed before implementing the study.
- Observations: Information should be provided on the observations to be made, how they will be made, and how frequently will they be made. If the observation is made by a questionnaire, this should be appended to the protocol. Laboratory or other diagnostic and investigative procedures should be described. For established

procedures, reference to appropriate published work is enough. For new or modified procedures, an adequate description is needed, with a justification for their use.

- Sample size: The protocol should provide information and justification about sample size. A larger sample size than needed to test the research hypothesis increases the cost and duration of the study and will be unethical if it exposes human subjects to any potential unnecessary risk without additional benefit. A smaller sample size than needed can also be unethical if it exposes human subjects to risk with no benefit to scientific knowledge. The basis on which sample size is calculated should be explained in the methodology section of the protocol. Calculation of sample size has been made easy by computer software programs. But the principles underlying the estimation should be well understood. These have been explained in Chapter 4.

Data management and analysis

The protocol should provide information on how the data will be managed, including data coding for computer analysis, monitoring and verification. Information should also be provided on the available computer facility. The statistical methods used for the analysis of data should be clearly outlined.

Ethical considerations

As outlined in Chapter 4, section 4.12, ethical considerations apply to all types of health research. These include research involving human experimentation, whether the research is of therapeutic or diagnostic nature that is carried out on patients who may expect a potential benefit from their participation, or is of a purely scientific nature for which human subjects volunteer to advance medical science but will not draw any therapeutic or diagnostic benefit. There are also ethical considerations for research involving human subjects but not experimentation. Epidemiological, field and qualitative studies fall under this category. Although no experimentation is involved, such studies can be as intrusive on the individual's privacy and even on communities. The ethics of research involving experimentation on animals has been receiving proper and increasing attention recently.

All research protocols in the biomedical field, particularly if it involves human subjects, must include a section addressing ethical considerations. This includes two components: The first is a written approval of the appropriate ethics review committee, together with a written form for informed consent, where appropriate. The second is a special section, preferably in the format of a checklist, to address all possible ethical concerns. Simply getting the ethical approval is not enough.

Approval by ethics review committees

For studies in humans (or involving human biological materials), the protocol must be approved by the local, institutional or equivalent ethics committee and/or national ethics committee.

For animal studies approval is required from the animal welfare committee of the institute or its equivalent. If no such committee exists, a statement signed by the principal investigator(s) should indicate that the research will be carried out in accordance with the International Guiding Principles for Biomedical Research involving Animals (see 4.12.4).

Informed decision-making

A consent form, where appropriate, must be developed and attached to the protocol. It should be written in the prospective subjects' mother tongue. The consent form has two parts: a) a statement describing the study and the nature of the subject's involvement in it; and b) a certificate of consent attesting to the subject's consent. Both parts should be written in simple language so that the subject can easily understand the contents. As much as possible, the use of medical terminology in writing up the consent form should be avoided. Special care is needed when subjects are illiterate.

The statement should, as appropriate, explain why the study is being done and why the subject has been asked to participate. It should describe, in sequence, what will happen in the course of the study, giving enough detail for the subject to gain a clear idea of what to expect. It should clarify whether or not the study procedures offer any benefits to the subject or to others, and explain the nature, likelihood and treatment of anticipated discomfort or adverse effects, including psychological and social risks, if any. Where relevant, the statement should include a comparison with risks posed by standard treatments or drugs. If the risks are unknown or a comparative risk cannot be given it should be so stated. Finally, the statement should indicate that the subject has the right to withdraw from the study at any time without, in any way, affecting her/his further medical care.

Ethics checklist

The protocol must describe the measures that will be undertaken to ensure that the proposed research is carried out in accordance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects (Annex 1).

A checklist must address ethical concerns that could be raised about the methodology, including the research design, selection of subjects, the interventions introduced and the observations to be made.

- Is the research design adequate to provide answers to the research question? It is unethical to expose subjects to research that will have no value.
- Is the method of selection of research subjects justified? The use of vulnerable subjects as research participants needs special justification. Vulnerable subjects include those in prison, minors and persons with mental disablity. Particularly in international research, it is important to ensure that the population in which the study is conducted will benefit from any potential outcome of the research. They should not be doing it to the benefit of another population. Justification is needed for any inducement, financial or otherwise, for participants to be enrolled in the study.
- Are interventions justified, in terms of risks/benefits ratio? Risks are not limited to physical harm. Psychological and social risks must also be considered.
- For observations made, have measures been taken to ensure confidentiality?

Gender issues

It was only recently that attention was drawn to the importance of addressing gender issues in research protocols. The Commission on the Status of Women made the above statement. This was in response to several areas of concern. "Ensure, where indicated, that clinical trials of pharmaceuticals, medical devices and other medical products include women with their full knowledge and consent and ensure that the resulting data is analysed for sex and gender differences."

- Women were often excluded from clinical trials on disease conditions that affect both men and women, on the basis of biological variability, and/or vulnerability. But women were given the same drugs, which had not been tested on them, as men if the drugs proved safe and effective for men.
- Drugs and devices intended for use by women only were sometimes tested on them without their proper informed consent, particularly in poor resource settings.
- When women were included with men as research subjects, gender was not always taken into consideration when results were analysed.

It is well known that genetic and hormonal factors modify the prevalence, behaviour and treatment of diseases of body systems in men and women. But what is less known is that culturally evolved gender-related differences in lifestyle behaviour are also powerful determinants of women’s health and account for major differences in the disease burden between males and females, probably more than genetic or hormonal factors. Both biological and gender-related differences can influence the outcome of the research for men and women.

References

The protocol should end with relevant references on the subject.

References and additional sources of information

Commission on the Status of Women Forty-third session. Revised draft agreed conclusions on women and health submitted by the Chairperson of the Commission. New York, United Nations Economic and Social Council, 1999 (E/CN.6/1999/L.2/Rev.1).

Fathalla MF. Gender matters in health research. In: *International Conference on Health Research for Development, Bangkok 10-13 October 2000. Conference report*. Geneva, Global Forum for Health Research, 2001: 105–108.

Kendall MJ and Hawkins C. Planning and protocol writing. In: Hawkins C, Sorgi M, *eds. Research: How to plan, speak and write about it*. Berlin, Springer-Verlag, 1985: 12–28.

UNDP/UNFPA/WHO/World Bank Special Programme of Research, development and research training in human reproduction. *Preparing a research project proposal. Guidelines and forms*. 3rd edition. Geneva, World Health Organization, 2000 (WHO/HRP/PP/2000).

Chapter 6
Submitting a research proposal

6.1 Introduction

A research proposal is a document written for the purpose of obtaining funding for a research project. Researchers should familiarize themselves with the potential sources for funding, and their specific requirements and mechanisms. They should know how to submit a proposal that will have a good chance of getting funded. Grantsmanship is the term used for the ability to secure grants to support research projects. The research proposal includes all the components of the research protocol outlined in the previous chapter. In addition, the proposal has to include additional information to convince the funding agency that the project is worthy of support and can be successfully implemented.

6.2 How to get your research project funded

6.2.1 Sources of funding

Funding for health research basically comes from either public sources or private sources. Public sources include governments and intergovernmental organizations. Private sources include the not-for-profit sector, such as philanthropic foundations and nongovernmental organizations, and the for-profit private industry. Besides these primary sources, there are intermediary agencies/organizations which play a role in channelling funding from the primary sources to the actors in research.

Government funding is provided through publicly funded national research organizations, such as national research councils, institutes of health and universities. Some ministries of health see the value of health research for their work, and allocate a budget for it.

Governments in developed countries may allocate funds for research through their bilateral official development assistance to developing countries. Two countries (Sweden and Canada) provide funding for research through publicly supported semi-autonomous agencies. The Swedish Agency for Research in Developing Countries (SAREC) and the International Development Research Centre (IDRC) in Canada provide a special mechanism for supporting research to solve developing country problems.

Intergovernmental organizations, such as the World Health Organization, support research through provision of funding, as well as technical support. Support is provided through headquarters' programmes, as well as through regional offices. Special research programmes in WHO cover the areas of reproductive health research and tropical disease research.

The not-for-profit private sector includes several foundations, large and small. Examples include the Wellcome Trust in the UK, and the Rockefeller, Ford and Bill and Melinda Gates Foundations in the USA, among others.

In the for-profit private sector, pharmaceutical companies, largely based in industrialized countries, are investing increasingly large sums of money in research and development.

A layer of intermediary support often serves as a bridge between the funders and those conducting research. An increasing number of international research programmes have been active, focusing on particular areas of health research. Examples include the Population Council headquartered in New York and with a number of regional offices, Family Health International (FHI) headquartered in Chapel Hill (North Carolina) and the Program for Appropriate Technology in Health (PATH) with headquarters in Seattle.

6.2.2 Will the project be funded?

Funding organizations receive many more proposals than what they can fund. The selection process is very competitive. The following factors are generally taken into consideration in deciding whether the proposal is to be funded:

- importance and relevance of the research question to the declared interests of the agency; success in obtaining a grant depends on matching the proposal with the interests of a granting agency;
- quality of the research design;
- ability of the investigators to carry out the project;
- capacity of the research facility to carry out the project;
- ability of the institution to handle the administrative and financial procedures;
- satisfactory ethical considerations;
- realistic and justifiable budget, within the limits set by the agency, and normally with no expectations for continued funding after the completion of the project;
- reasonable time-frame for completion of the project;
- understanding of anticipated problems;
- clarity and style of the written proposal.

Writing with enthusiasm is a good idea, but overstatements should be avoided. The applicant should be realistic about the limitations of the study.

6.2.3 How to submit a research proposal

Funding organizations use one or more of the following mechanisms to select and fund research projects: solicit proposals, advertise and invite proposals or have an open door policy.

- Soliciting proposals: In this case, one or more research institutions are approached and are asked for their interest in submitting a research proposal in a certain area of importance to the funding agency. Usually the institution approached is a centre of excellence.
- Advertising an invitation for submitting proposals in certain areas of interest: Here, the process is competitive, and normally there is a time limit for submitting proposals. A part of grantsmanship is to be ready with good ideas, and to be able to speedily compile an attractive proposal. The proposals are then independently reviewed and scored, and a small number is selected for funding. Some funding agencies will ask first for a brief concept outline of the proposal, and then shortlist the applications, and ask for complete proposals from the short list.
- An open door policy for any good proposal: Most funding agencies, however, have areas of interest and areas in which they are not interested. Good advice in this approach is not to send a full proposal. A brief outline of the project with the level of funding requested is enough to get a response about the potential interest of the agency or organization in considering it. If the response is positive, the full proposal can be sent.

Funding organizations have their own websites. Information about their interests and mechanisms can be easily accessed.

6.2.4 Response to comments of reviewers

Research proposals are commonly subjected to peer review. The reviewers may suggest that the proposal can be made more acceptable by revisions. The investigators do not need to make all the changes suggested automatically. They should adopt revisions that will satisfy the reviewer's criticisms wherever possible and justify any decision not to do so. It is good to indicate in a separate page the criticism made and how the revised proposal responded to them.

6.3 Components of a research proposal

A research proposal commonly follows the following format, which includes the components of the research protocol, with some additional information. Some funding agencies have their own formats for standardizing applications and streamlining the review process.

- Title page
- Project summary
- Project description
- Ethical considerations
- Gender issues
- Timetable
- Problems anticipated
- Budget
- References
- Curriculum vitae of the investigator(s)

Title page

This page should provide information on:

- Project title
- Principal investigator(s)
- Institution
- Duration of the project
- Funding requested

Research grants are normally given to institutions not to individuals. The name of the institution should be given in the title page. If the institution is not known to the agency, a brief information about the institution may be given as an annex, or may be requested by the agency. The name of the financial officer who will be in charge of administering the grant should be given, in addition to the names of the investigators.

The duration of the project must be specified. Most agencies will not commit support beyond three years. The funding requested should be specified. In a multi-year project, the amount requested for each year should be outlined.

Project summary

The project summary should be carefully written. It will be the first (and may be the only) part read by the reviewers. It should reveal persuasively the importance and the strengths of the project.

Project description

This should follow the lines of the protocol, as already discussed in the chapter on writing the protocol. The rationale should not only explain why the project is important to do, but should also indicate its relevance to the particular lines of interest of the funding agency. Previous work by the investigators on the research topic will indicate the competence of the investigators to carry out the work. Pilot studies, if already done, are important to demonstrate the feasibility of the research.

Ethical considerations

Approval from the local ethics review committee does not relieve the donor agency from the ethical responsibility for the project. Also approval by a donor agency does not relieve the research institution from ethical responsibility for the project. Ethical issues and concerns should be addressed fully in the research proposal, as outlined in the chapter on writing the research protocol.

Gender issues

Most funding organizations are now increasingly conscious about gender issues. These should be addressed in the proposal, as outlined in the previous chapter on writing the research protocol.

Timetable

The investigators should commit themselves to a timetable. This may include a preparatory phase to train research workers, to procure equipment/supplies, or to complete a pilot phase. The timetable should then estimate the duration for collection of data, final analysis of data and writing up the report. In project proposals of a long duration (more than one year), the timetable should set milestones to be reached. These are taken into consideration when progress reports are reviewed by the funding agency. Funding is often released on the basis of these progress reports.

Problems anticipated

The investigators should demonstrate their awareness of obstacles and difficulties, which may interfere with the successful completion of the project within the timeframe and cost proposed. They should explain how these obstacles and difficulties would be dealt with. An investigator who does not anticipate any problem probably has not thought out the details of the project carefully.

Budget

The budget request should be itemized and each item should be justified.

Budget itemization

The following are examples of categories of expenses:

- Personnel (names, positions, percentage of time spent on the project, salary, fringe benefits)
- Equipment
- Supplies
- Patient care costs
- Travel
- Data processing
- Communications
- Secretarial expenses
- Publication/dissemination of information about the outcome of the project.

Budget justification

All items in the budget need to be justified and are closely scrutinized in the following way:

- Are all personnel needed for the amount of time stated?
- Are critical personnel devoting enough time to the project?
- Major pieces of equipment are difficult to justify in a small project; an exception may be made for a developing country institution as part of research capability strengthening.
- The budget should not include any undue inducement for subject participation.

If the duration of the project is more than one year, a detailed budget is needed for at least the first year. Budget request for the subsequent years should be outlined. Agencies would normally approve the budget for the full duration of the project, but funds will be

released on a yearly basis, subject to the submission of acceptable progress and financial reports.

Agencies normally will allow some flexibility in the use of the budget, provided the total budget is not exceeded. For shifts between budget items, however, it is expected that the agency's approval be sought beforehand.

An unrealistic budget is likely to lead to rejection of the proposal. The budget may be unrealistic in one of two ways. It may ask for more than is needed to undertake the project or it may ask for much less than is realistically needed to undertake the project successfully. The investigators may want to limit the budget to the funding ceiling of the agency, but keep the large project as it is. Instead, they should limit the project objectives to what can realistically be achieved with the requested funds.

References

A number of recent references on the subject should be cited in support of the proposal.

Curriculum vitae (CVs) of the investigator(s)

The ability of the investigators to carry out the project is an important consideration. Biographical sketches of the investigators or CVs should be attached. The track record of the investigators is important. Preliminary studies or other work done by the investigators on the subject should be included.

References and additional sources of information

Cummings SR, Holly EA, Hulley SB. Writing and funding a research proposal. In: Hulley SB, Cummings SR. eds. *Designing clinical research: an epidemiologic approach*, 2nd edition. Philadelphia, Lippincott Williams & Wilkins, 2001: 285–298.

Wingo PA, Higgins JA, Rubin GL, Zahniser SC. *An epidemiologic approach to reproductive health.* Geneva. World Health Organization, 1994: 15–67 (WHO/HRP/EPI/1994).

Health research methodology: A guide for training in research methods. 2nd edition. Manila, World Health Organization Regional office for the Western Pacific, 2001: 147–161.

UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction. *Preparing a research project proposal. Guidelines and forms*. 3rd edition. Geneva, World Health Organization, 2000 (WHO/HRP/PP/2000).

Wyatt KM, Dimmock PW. Applying for a grant. In: O'Brien PMS, Pipkin FB. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 201–209.

Chapter 7

Implementing the research project

7.1 Introduction

How should research be done? The answer to this question can be given in one word: well. Whatever the reason for the research, and whatever the kind of research, it should be done well and should conform to established standards of scientific methodology. It has been said that there is only one type of research: good research. Bad research does not deserve the name of research.

It is not enough that the research question has been well conceived, the appropriate research design selected, and a detailed protocol well thought out and written. All these provide a good anatomy for the research. Physiology matters even more. The research should be implemented with scientific rigour.

7.2 Scientific rigour

The English word "rigour" literally means "strictness". In scientific research, the term rigour is used to imply that:

- the study protocol is being adhered to;
- the research is conducted in accordance with established ethical standards;
- meticulous and detailed records of all observations are maintained;
- methods of measurement are used in an objective way to provide valid and reliable results;
- data are analysed and interpreted using appropriate statistical methods to assess the validity of the results and their generalizability;
- the researchers continue to be well versed with the literature on the subject during the study;
- results are presented in such a way that other investigators can re-analyse the data using the same processes and methods and reach the same conclusions, and other investigators can replicate the study to confirm or refute the findings.

7.3 Pre-testing the protocol

It is always wise to pre-test the protocol after developing it. This is particularly important in large and expensive studies. What appears to be a straightforward and problem-free protocol may prove to have logistic and practical problems in implementation. The pre-test is sometimes called a pilot study. Based on the outcome of the pilot study, the protocol may be modified before the study proper is implemented.

The pilot study can help in determining whether the required number, as well as composition, of study subjects will be recruited within the time frame of the study. The size of the sample may need to be modified, or alternative approaches of recruitment may be explored.

The pilot study can help in testing the methods of measurement. If the study relies on how records have been kept, these records may be checked for accuracy and completion before the study is started. If a questionnaire has been designed, this will need to be pre-tested to check that the questions are clear without any ambiguity and that the answers will be consistent. Modifications may have to be made for the final instrument. If the methodology involves a clinical or laboratory measurement, this has to be tested for inter-rater and intra-rater reliability, i.e. for consistency in results obtained by different workers and by the same worker at different times.

The pilot study can also help in testing the system for data management. Entering and editing the data from the pilot study will show whether the system is working well. This includes designing the forms for recording measurements, choosing a computer, developing programmes for data entry, management and analysis; and planning dummy tabulations to assure that the appropriate variables are collected.

7.4 Monitoring of the study

The study should be monitored. In large clinical trials, a monitor may be appointed with the responsibility of reporting on the progress of the trial and for verification of data. There are two components to monitoring: data management (record keeping and data handling) and data quality (quality assurance and quality control).

Record-keeping and handling of data

All steps involved in data management should be documented in order to allow step-by-step retrospective assessment of the quality of the data and the performance of the clinical trial (“the audit paper trail concept”). A basic aspect of the integrity of data is the safeguarding of “blinding” with regard to assignment of subjects to different treatments. Subject files and other supporting data must be kept for a period of time as required by local regulations.

A common problem in research is the tendency of investigators to collect many data, much more than they can analyse or publish. This can result in an excessively large database and increase the chance of inaccuracy. Limiting the data to be collected to the essential data for the study, and eliminating redundancies, will enhance the accuracy of the study. A general advice to investigators is to be parsimonious (not to expand more than is necessary).

Quality assurance and quality control

A system for quality assurance must be implemented to ensure that the study is performed and the data are generated, recorded and reported in compliance with the protocol, good clinical practice and national regulations. In clinical trials, all sites and all data and documents must be available for verification. All observations and findings should be verifiable in order to ensure the credibility of data and to ensure that the conclusions presented are derived correctly from the raw data. Quality assurance is carried out with the following objectives: to ensure that no data are missing and to ensure that data are precise and accurate.

Missing data will introduce a problem in the analysis of results, whether the data are missing because the measurement was not made or was not recorded. A special type of missing data is loss to follow-up. Loss to follow-up will decrease the number of subjects. Generally, when sample size is estimated, a provision is made for an estimated percentage of loss to follow-up. But this does not completely solve the problem. Loss to follow-up may introduce a bias in the study and discredit its conclusions. The subjects lost to follow-up may be different from the subjects who continued in the study. For example, subjects who develop serious side-effects, complications or even die may be disproportionately represented in the loss to follow-up.

Monitoring during the study can help in reducing the problem of missing data. A computer program can help during data entry to ensure that all data are entered. The computer program will flag missing and out-of-range values.

Inaccurate and imprecise data are a worse problem than missing data, because they may not be discovered after the fact. Only a systematic quality control programme will avoid the problem.

Reliability of measurements is an important component of quality assurance. To test for intra-observer reliability, a common way is to do the measurements twice (test–retest reliability). The results obtained from the first test are then correlated with the second test. To test for inter-observer reliability, a common way is to have the same measurements done by two observers. The results obtained from the first test are then correlated with the second test.

To ensure the quality control of laboratory measurements, blinded duplicates or standard pools can be used In multicentre studies involving laboratory measurements, a common practice is to have a reference laboratory. This reference laboratory will standardize the test to be used and will periodically send the same sample to the different centres and provide them with feedback on how their results compare with each other and with results as determined in the reference laboratory. This mechanism of quality control is essential before a decision is made to pool the results from the different centres together for analysis.

7.5 Periodic tabulations and reports

Periodic tabulation of the data is useful in the monitoring process. Periodic frequency distribution tables will reveal aberrant values. Periodically looking at the data should never mean breaking the code for blinded studies.

7.6 Validation of results in qualitative research

The researcher doing qualitative research may use two or more methods (observation, interviews, focus group discussions) to answer the same question, or may use more than one source for data collection. The objective is to enhance the validity and reliability of the results by comparing the data obtained from different methods or different sources. This process in qualitative research is sometimes referred to as “triangulation”. The idea of triangulation originated from a craft used by land surveyors, who increase the validity of a map by incorporating measures from different angles. Multiple and diverse observations can enrich the description of a phenomenon. The researchers may also cross-check interim research findings with the respondents. This is called “respondent validation”.

7.7 Good clinical practice

Results of clinical trials on novel pharmaceutical products have to be submitted to drug regulatory authorities before the products can be approved for general use. The drug regulatory authority will not only look into the results. It will also consider the process by which these results were obtained, and how the research was carried out. The research should have been conducted according to good clinical practice (GCP). The drug regulatory authority will discard any results of research that did not conform to the guidelines for GCP.

GCP is a standard for clinical studies which encompasses the design, conduct, monitoring, termination, audit, analyses, reporting and documentation of the studies and

which ensures that the studies are scientifically and ethically sound and that the clinical properties of the pharmaceutical product under investigation are properly documented. A World Health Organization technical report provides guidelines for good clinical practice for trials on pharmaceutical products, and is the basis for some of the material in this chapter (WHO, 1995).

Audit is an important component of GCP. An audit is a systematic examination, carried out independently of those directly involved in the clinical trial. Its objective is to determine whether the conduct of a trial complies with the agreed protocol, and whether the data reported are consistent with the records on site. For example, it may check whether data reported or recorded in the case report forms are consonant with those found in hospital files and other original records. The auditor may use statistically controlled sampling to verify data obtained in a trial.

7.8 Research on new pharmaceutical products

Clinical trials of pharmaceutical products should be done in a stepwise fashion. Progress to the next phase should follow the successful completion of the previous phase. The number of subjects in the trial is increased from one phase to the next, as safety and efficacy of the product becomes better established. Animal toxicology studies are usually required, and specific toxicology studies should be completed before moving from one phase to the next.

Clinical trials are generally classified into phases I to IV. It is not possible to draw distinct lines between the phases, and diverging opinions about details and methodology exist. A brief description of the individual phases, based on their purposes as related to clinical development of pharmaceutical products, is given below.

- Phase I clinical trials: These are the first trials of a new active ingredient or new formulation in humans, often carried out in healthy volunteers. Their purpose is to establish a preliminary evaluation of the safety, and the pharmacokinetic and, where possible, pharmacodynamic profile of the active ingredient in humans.
- Phase II clinical trials: These trials are performed in a limited number of subjects and are often of a comparative (e.g. placebo-controlled) design. Their purpose is to demonstrate therapeutic activity and to assess the short-term safety of the active ingredient in patients suffering from a disease or condition for which the active ingredient is intended. This phase also aims at the determination of appropriate dose ranges or regimens and (if possible) clarification of dose–response relationships in order to provide an optimal background for the design of expanded therapeutic trials.

- Phase III clinical trials: Phase III trials include larger (and possibly varied) patient groups, with the purpose of determining the short-term and long-term safety/efficacy balance of formulation(s) of the active ingredient, and of assessing its overall and relative therapeutic value. The pattern and profile of any frequent adverse reactions must be investigated and special features of the product must be explored (e.g. clinically relevant drug interactions, factors leading to differences in effect such as age). These trials should preferably be of a randomized double-blind design, but other designs may be acceptable. Generally, the conditions under which these trials are carried out should be as close as possible to normal conditions of use.
- Phase IV clinical trials: Phase IV trials are studies performed after marketing of the pharmaceutical product. They are carried out on the basis of the product characteristics for which the marketing authorization was granted and are normally in the form of post-marketing surveillance, or assessment of therapeutic value or treatment strategies. Although methods may differ, these studies should use the same scientific and ethical standards as applied in pre-marketing studies. After a product has been placed on the market, clinical trials designed to explore new indications, new methods of administration, new combinations, etc. are normally considered as trials on new pharmaceutical products.

7.9 Termination of the study

A study on a new pharmaceutical product should be closely monitored. The study should stop if: a) unanticipated, potentially serious side-effects are encountered; or b) the comparative study shows clearly, before the study is completed, that one drug is clearly superior to the other.

7.10 Changes in the protocol

The study should be carried out in accordance with the written protocol. Any subsequent change must be agreed upon and documented. For large studies, standard operating procedures in the form of detailed written instructions should be developed and followed.

The protocol should be adhered to. Unauthorized changes in the protocol are termed violations. Violations of the protocol, if discovered late, discredit the study. If discovered before the analysis, the data should be discarded from the analysis. Minor changes can be made in the protocol if this does not affect the characteristics of the data. Otherwise, the data before and after the change cannot be pooled together. If major changes have to be made in the protocol, data before and after the change should be analysed separately.

7.11 Ethical issues in the implementation of the study

7.11.1 Ethical principles

An ethically acceptable design is only as good as the carrying out of the design. The ethical principle of non-maleficence or "do no harm" implies that during implementation of clinical trials, a "cut-off point" should be defined so that if the proposed treatment proves risky or inferior to the alternative, it should be stopped. The ethical principle of respect implies during implementation that patients should be able to withdraw their consent at any time without losing any benefit.

7.11.2 Experimentation on animals

Ethical approval is needed for animal research from the appropriate local and national authorities. Only investigators and personnel who have the appropriate qualifications and experience should carry out research on animals. Experimental work on animals should only be done in qualified and certified facilities. Research animals should be properly cared for as regards housing, environmental conditions, nutrition and veterinary care. Normally the care of animals should be under the supervision of veterinarians having experience in laboratory animal science. The avoidance or minimization of discomfort, distress or pain to the animal is an ethical imperative. Procedures with animals that may cause more than momentary or minimal pain or distress should be performed with appropriate sedation, analgesia, or anaesthesia in accordance with accepted veterinary practice. At the end of, or, when appropriate, during an experiment, animals that would otherwise suffer severe or chronic pain, distress, discomfort, or disablement, that cannot be relieved, should be painlessly killed.

7.11.3 Scientific honesty

Data should be carefully and accurately collected, without any subjective bias on the part of the investigators. As discussed in Chapter 4, a methodology that is relevant in this regard is the double-blind controlled clinical trial, where the investigators are not aware of the type of medicine the subject is given. In a "triple-blinded" design, patients, clinicians and statisticians (or persons measuring the outcome) are unaware of which group is subject to which intervention. Another research methodology is randomization, whereby it is not up to the investigator to assign particular treatments to different subjects. The decision is made by random allocation.

Deliberate scientific fraud is ethically unforgivable. Fraud involves deliberate deception and may take the form of fabricating data, inventing patients, or manipulating data to provide a desired answer. The pressure to "publish or perish" in academic institutions may be a factor, as well as the practice of drug companies of paying a fee to

the investigator for every patient participating in a clinical trial. Local research ethics committees should have the authority to audit the implementation of the research, and to contact research subjects.

7.11.4 Fiscal honesty

Research programmes and projects are commonly supported by government, private or international funds. The research funds have to be used to meet the expenses as agreed upon in the research proposal. Expenditure has to be documented. Accurate periodic and final financial reports are required and should be submitted.

References and additional sources of information

Hulley SB, Cummings SR. Implementing the study: Pre-testing, quality control, and protocol revisions. In: Hulley SB, Cumming SR, eds. *Designing clinical research: an epidemiologic approach*, 2nd edition. Philadelphia, Lippincott Williams & Wilkins, 2001: 259–271.

Guidelines for good clinical practice (GCP) for trials on pharmaceutical products. Geneva, World Health Organization, 1995: 97–137 (WHO Technical Report Series, No. 850).

Chapter 8

Describing and analysing research results

8.1 Introduction

Data accumulated during the research can be voluminous and will need to be summarized, and presented in a clear accurate format. For this, tools of descriptive statistics are used, including tabulation, calculations, figures and correlation.

After the investigators have summarized and described the results, the next step is to analyse the results. In the analysis, they need to question whether the estimates made in the study can be generalized beyond the relatively small number of the sample studied. They need also to question whether the differences or associations found can be explained by chance, and so may not be real. Inferential statistics provide the tools to help in answering these questions.

This chapter emphasizes the underlying concepts in describing and analysing research results. To keep the clarity of the message, technical detail and mathematical calculations are not addressed. For these, the reader can consult the list of references and additional sources for the chapter.

8.2 Descriptive statistics

The results of the study must be clearly summarized to allow their proper analysis and interpretation. Descriptive statistics helps us to make sense of a large volume of data. Its first use is credited to John Graunt, a storekeeper in London in the mid-17th century (Weaver, 2000). Beyond the boring business of the store and bookkeeping, he developed an outside interest in areas of mathematics. He exercised his talent in reviewing a weekly church publication issued by the local parish clerks that listed the numbers of births, christenings and deaths in each parish. These so-called "bills of mortality" also listed the causes of death, thus providing Graunt with a massive but unorganized mass of information about the ongoing drama of birth and death occurring all around him. Graunt made a big effort to organize the data in a way that was probably inspired by his techniques for tracking his shop inventory. He devised tables that could be easily updated. He took great pains to reduce several confused volumes of information into tables and

succinct paragraphs. He was able to compare changes in mortality tables over the years. Graunt published the summary of his work, entitled "Natural and political observations made upon the bills of mortality", in 1662. His book immediately attracted the attention of government leaders and prominent private citizens. The story goes that King Charles II was so impressed with Graunt's work that he proposed his name for membership of the newly created prestigious Royal Society, a forum in which the nation's most brilliant scientists could gather together and exchange ideas. Grant's trade as a shopkeeper provoked objections from the members, but they were over-ruled by the king.

The following tools can be used in describing and summarizing the results: tabulation, calculations, graphs/figures and correlation.

8.3 Tabulation

A first step in summarizing the data is commonly to group them in summary tables. The plan for the tables should be developed in the research design phase. The term "dummy tables" is used to describe tables that are not yet filled with data. During the implementation phase of the research, these dummy tables may be filled with the available data to see how the results are shaping up.

Frequency distribution tables

A frequency distribution table gives the frequency with which a particular value appears in the data. In designing the frequency distribution table for numerical data, suitable class boundaries are needed. The number of classes is important. If the classes are too small, the table will be unwieldy. If they are too large, information may be lost by being too summarized. If in doubt, classes are better initially chosen to be small rather than large. Small ones can be easily amalgamated to form larger ones if needed. Classes must also be mutually exclusive. For example, if we tabulate data about the diastolic blood pressure, we may make the classes as 70 to 79, and 80 to 89, not 70 to 80 and 80 to 90.

Cross-tabulation tables

Frequency distribution tables may describe one variable at a time, for example age distribution. Depending on the objectives of the study, there is often the need to examine the relationship between several of the variables at once, for better description of the data or in order to look for differences or relevant associations.

Cross-tabulation tables may be descriptive or analytical. Descriptive cross-tabulations may be used to describe the sample. An example is a composite table describing the background of subjects, such as age, sex, profession, etc. Analytical

cross-tabulations may be used to determine differences between groups. An example is a table comparing low-birth-weight babies (less than 2500 g) and normal-birth-weight babies in women who received and who did not receive prenatal care. The accepted convention in analytical cross-tabulations is to put the categories of the dependent variable (birth weight) as column headings, and the independent variable (prenatal care) as row headings. The totals are also put for the columns and for the rows. If percentages are used, they should add up to a total of 100%.

Analytical cross-tabulations may focus on exploring associations or relationships between variables. An example is the relationship between the age of the mothers and the duration of breastfeeding. Columns may have three categories for duration of breastfeeding: 0–5 months, 6-11 months, and 12 months or more. Age is put in rows of age groups, for example less than 20, 20–29, 30–39, and 40 or more.

The need for cross-tabulations is dictated by the objectives of the study. Possible conclusions are anticipated during the research design. Dummy cross-tabulations that will allow the conclusions to be made are developed and left empty to be filled when the data are available.

8.4 Calculations

Numerical data can be summarized by calculating their central tendency and variability, by calculating percentage and proportions, and by calculating ratios and rates. Computer software programmes have facilitated these calculations.

Central tendency

The most commonly used measure of central tendency is the arithmetic mean. Less familiar but also useful measurements of central tendency are the median and the mode.

The mean, also called arithmetic mean, is derived by summing up the individual values and dividing by the total number of measurements.

The median of a distribution is a midpoint at which one half of the observations fall below and one half fall above the value.

The mode is the most frequent measurement in a distribution.

If the data fall in a “normal” (evenly spread around the mean) distribution, the mean, median and mode coincide. In “skewed” distributions (data not evenly spread) they vary and may all be meaningful in the presentation of the data.

Variability

In addition to knowing the mean value of a series of measurements it is important to have some idea about their variation around the mean. There are three ways to present the variability of data around the mean: the range, the standard deviation and the percentiles.

The range gives the values at the top and at the bottom, but does not give much indication of the spread of observations around the mean. This spread is provided by the standard deviation.

The standard deviation (SD) is calculated from a formula that sums the squares of differences between the group mean and each individual value. This sum total is termed the variance. The square root of the variance provides the standard deviation. The greater the differences between the values, the more spread the distribution and the larger the standard deviation. Mathematicians have calculated that if the observations follow a "normal" distribution (values evenly spread around the mean), a range covered by one standard deviation above and below the mean will include about 68% of the observations. A range of ± 2 SD will cover about 95% of observations. A range of 3 SD will cover about 99.73% of the observations. Calculating the mean and the standard deviation gives us a good summary of the data.

Percentiles provide another way of looking at variations in distributions. Just as the median is the 50th percentile of a collection of data, the 75th or 95th percentile can be determined and indicates that a particular measurement is larger than 75% or 95% of all the other values. The interquartile range is the distance between the scores representing the 25th and 75th percentile ranks in a distribution. One advantage of percentiles is that they can be applied to data with skewed, not normal distribution, that is data not distributed evenly around the mean.

Percentages, proportions, ratios and rates

A percentage is the number of units with a certain characteristic divided by the total number of units in the sample and multiplied by 100. Usually missing data are not included in the calculation of percentages. Caution should be exercised when describing percentages based on small numbers. In such cases, a small difference may appear as a big difference in percentages.

A proportion is a numerical expression that compares one part of the study units to the whole. A proportion can be expressed as a fraction (for example a proportion of 2/5) or a decimal (for example 0.40)

A ratio is a numerical expression of the relationship between one set of frequencies and another. An example is the ratio of males to females in a sample.

A rate is a numerical expression of the frequency of a condition in a given population measured in a specified period of time. Two rates commonly used in health sciences are incidence rate and prevalence rate. Incidence rate relates the number of new cases of a condition in a population within a time period. Prevalence rate relates the total number of cases with a condition in a population at a given time.

An illustration of the difference between rates and ratios is the measurement of maternal mortality. If we relate the number of women who die because of pregnancy and childbirth to the number of women who have live births, we are calculating a ratio. If we relate them to all women in the childbearing period over a certain time, we are calculating a rate.

8.5 Graphs/figures

Figures improve the readability of the results. A Chinese sage once said that a picture is worth more than a thousand words. Figures include bar charts, pie charts, histograms, line graphs and maps. They are generally useful for the presentation of data. A histogram resembles a bar graph but the bars are drawn to touch each other, reflecting the underlying continuity of the data.

A common first step in looking at and summarizing data is to plot them in a frequency distribution curve. Each variable is plotted against the frequency with which it is found. The shape of this distribution curve tells a lot about the data, and has implications for subsequent analysis.

When the frequency distribution is a bell-shaped curve, it is described as normal or Gaussian distribution. Gauss, a German mathematician who lived in the early nineteenth century, proposed the “law of errors”, which states that repeated measurements made on the same physical object fall in a predictable pattern or distribution. Gauss’s original law was intended for measurement on the same object. It has been later applied to the grouping of measurements made on different objects.

When the frequency distribution curve is asymmetrical, it is described as skewed distribution. In a skewed distribution, the curves are asymmetrical, with one side of the curve extending in an elongated fashion.

Less commonly the frequency distribution curve may show more than one peak.

8.6 Correlation

In the context of correlation, data are classified as either independent or dependent variables. Independent or input variables ordinarily have values that are autonomous of the dependent or outcome variables. Because independent variables precede dependent

variables, they often are called predictors. Dependent (also called output or outcome) variables have responses that are contingent on independent variables. Independent variables are antecedents; dependent variables are consequents. In epidemiology, independent variables are often called risk factors or exposure variables.

Scatter diagram

When an investigator has collected two sets of observations and wants to see whether there is a relation between them, it is best to construct a scatter diagram first. The vertical scale represents one set of measurements and the horizontal scale the other. The dots in a scatter diagram generally lie neither in a single straight line nor equidistant on either side of a central line. They often lie in a roughly elliptical area. The scatter diagram gives an indication whether a correlation may exist and its direction.

Usually, independent variables are graphed on the x-axis (horizontal axis) and dependent variables are graphed on the y-axis (vertical axis).

Correlation coefficient

When the relationship between two variables can be expressed graphically by a straight line, correlation can be expressed as the correlation coefficient. Correlation may be positive or negative. When one variable increases as the other increases, the correlation is positive; when one decreases as the other increases it is negative. The correlation coefficient (r) is measured on a scale that varies from +1 through 0 to –1. Complete correlation between two variables is expressed as 1. It should be clear that correlation means association, but does not necessarily mean causation. This conclusion is left to the interpretation of the results.

Regression equation

Correlation between two variables means that when one of them changes by a certain amount the other changes on the average by a certain amount. The relationship can be described by a simple equation called the regression equation. The regression equation, can be used to construct a regression line on a scatter diagram. As the line must be straight, it will probably pass through few, if any, of the dots. The regression coefficient is the term used to signify the amount by which a change in one variable (independent variable) must be multiplied to give the corresponding average change in another variable (dependent variable). It represents the degree to which the regression line slopes upwards or downwards.

8.7 Inferential statistics

8.7.1 Analysis

After summarizing and describing the results, investigators move to the next step of analysing the results. The investigators should question whether the findings in the study could be generalized beyond the relatively small number of the sample studied. This is referred to as external validity or generalizability.

Statistics helps us in making inferences, and are therefore called inferential statistics. An inference is a generalization made about a population from the study of a subset or sample of that population. In statistics the term population has a different meaning from the general usage of the word. It needs not to refer only to people or animate creatures. Since a population commonly contains too many individuals to study conveniently, an investigation is often restricted to one or more samples drawn from it. It should be emphasized that if the study sample is not representative of the population, the inference we make from the result will be misleading. Analytical statistics will be of no help if the sample is not representative. Analytical statistics cannot correct our mistakes in designing the study.

But even with properly selected samples, results from a single sample are still subject to some degree of uncertainty, or chance. This sampling error cannot be eliminated completely, but its probable magnitude can be calculated.

Statistics is more about common sense than about mathematics. Generalization from the finding from the sample to the population from which the sample was drawn depends mainly on the two factors: the size of the sample and the variability in the results. Naturally, if we have examined 100% of the community, then the result will represent the finding in the whole community. The smaller the sample drawn, the less likely its findings can be generalized. Also, marked variation between subjects in the measurements obtained means that different results are more likely to be obtained from different samples. If the results fall within a wide range, i.e. the variability is high, then a small sample will be less likely to represent the result in the whole population. What analytical statistics does is to translate this common sense into quantitative terms by putting a figure on the probability. This is illustrated by the concept of standard error.

8.7.2 Standard error

The standard error (SE) is a statistical measure about the probability that the finding in the sample will reflect the finding in the population. The standard error depends on two factors: the size of the sample, and the variations of measurements in the sample indicated by the standard deviation. For example, the standard error of a mean is calculated as the standard deviation divided by the square root of the number of observations.

By itself, the standard error may have a limited meaning, but it can be used to produce a confidence interval, which has a useful interpretation. In simple terms, it has been calculated that the sample mean plus or minus 1.96 times its standard error gives the 95% confidence interval, meaning that there is only a 5% chance that this interval does not include the mean of the population. A confidence interval is thus a range of values which includes the population parameter at a specified level of probability.

The standard error (SE) can be calculated not only on a mean, but also on the difference between two means, on a percentage, on a difference between two percentages, and on a correlation coefficient.

The standard error should be clearly differentiated from the standard deviation. The standard deviation is a measure of the variability in the sample studied. The standard error is a measure of the uncertainty in a sample statistic. The standard error, which depends on both the standard deviation and the sample size, is a recognition that a sample is unlikely to determine the population value exactly. In many publications, the ± sign is used to join the SD or SE to an observed mean. This may be confusing as to whether it refers to the SD or SE. The present policy of many scientific journals is to remove the ± signs and to indicate clearly between brackets whether the SD or SE is being quoted, e.g. the mean was 51 kg (SD 8.4 kg).

8.7.3 Testing the research hypothesis

The formulation of the research hypothesis has been discussed in Chapter 4. Researchers may feel strongly that their hypothesis is true. This, however, should not influence the vigour with which the hypothesis should be tested. Sir Peter Medawar, in his book "Advice to a young scientist", said "I cannot give any scientist of any age better advice than this: the intensity of the conviction that a hypothesis is true has no bearing on whether it is true or not." (Medawar, 1979). Scientists should avoid falling in love with their pet hypotheses. The important question is whether the hypothesis can stand up to critical evaluation.

In scientific methodology, the research hypothesis is therefore not tested directly. Instead, we start with an assumption that there is no difference or association between the variables compared. This is called the null hypothesis (H_0). The null hypothesis is thus the contrary to the research hypothesis (also termed alternative hypothesis).

In scientific methodology, even if a difference or an association is found, it should be assumed that it is due to chance, until it is proven, by statistical analysis, that it is unlikely to be explained by chance. The research hypothesis is accepted by exclusion if the statistical test rejects the null hypothesis.

Using the null hypothesis in scientific work has been likened to the judicial process of assuming innocence until guilt is proved (Browner et al., 2001). In some ways, the

investigator's problem is similar to that faced by a judge trying a defendant. The absolute truth about whether the defendant committed the crime cannot be determined. Instead, the judge begins by presuming innocence: the defendant did not commit the crime. The judge must decide whether there is sufficient evidence to reject the presumed innocence of the defendant; the standard in legal language is known as "beyond reasonable doubt". A judge can err, however, by convicting a defendant who is innocent, or by failing to convict one who is actually guilty. Along the same reasoning, two kinds of mistakes can be made in the testing of the null hypothesis in research methodology, and determining statistically whether the finding could be due to chance. The first is when we reject the null hypothesis and it is true. This is similar to the mistake in the judicial process of rejecting the innocence and convicting an innocent defendant. In statistical language, this is called a type I error. Failing to reject the null hypothesis when it is not true is called a type II error. This would be similar in the judicial process of failing to convict a defendant who is actually guilty. The statistical tests used to assess whether the findings can be explained by chance are, in a sense, similar to the judicial process of assessing proof beyond reasonable doubt.

The probability of committing a type I error (rejecting the null hypothesis when it is actually true or proving an association when none exists) is called alpha. Another name for alpha is the level of statistical significance. The probability of making a type II error (failing to reject the null hypothesis when it is actually false or failing to prove an association when it actually exists) is called beta. The quantity (1-beta) is called power. Statistical power of a study is thus the probability of observing an effect (of a specified effect size) if one exists.

8.8 What statistical tests tell us

8.8.1 Probability

Albert Einstein said, "As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality." There is no certainty in science. There are probabilities. What is certain about science is the uncertainty. In scientific methodology, we try to minimize the probability of finding an association when no association actually exists, and to minimize the probability of missing an association when an association actually exists. We cannot eliminate this probability of error, but analytical statistics can give us an estimate of its magnitude. The probability of making an error depends on the size of the sample studied in order to test the null hypothesis. The larger the size of the sample, the less likely will be the probability of making an error. This is why determination of sample size is an essential component of research design.

8.8.2 Statistical significance

A statistical significance test estimates the likelihood that an observed study result, for example a difference between two groups or an association, can be due to chance and therefore no inference can be made from it.

Tests of statistical significance are based on common logic and common sense. That a difference is likely to be real and not due to chance is based largely on three criteria. The first is the magnitude of the difference observed. It is reasonable to expect that the larger the difference, the more likely that it is not due to chance. The second is the degree of variations in the values obtained in the study. If the values fall within too wide a range, differences in means would be more likely to be due to chance variations. The third very important criterion is the size of sample studied. The larger the size of the sample, the more likely that the result drawn from it will reflect the results in the population. What statisticians do is to turn this simple logic, through mathematics, into a quantitative formula, to describe the level of probability.

When the data are analysed, we set an arbitrary value for what we can accept as alpha or level of statistical significance, i.e. the probability of committing a type I error (rejecting the null hypothesis when it is actually true, or proving an association when none exists). The statistical tests then determine the *P* value. *P* is the probability that a difference or an association as large as the one observed could have occurred by chance alone. The null hypothesis is rejected if the *P* value is less than alpha, the predetermined level of statistical significance. Probability or *P* is usually expressed as a percentage. A result is commonly considered to be unlikely to be due to chance, or to be statistically significant, if the *P* value is less than 5% (*P* less than 0.05) and is said to be highly significant if *P* is less than 0.01. There is nothing magical about these levels of probability. They are arbitrary cutoff points, a tradition that began in the 1920s with an influential statistician named Fisher. It is important to keep in mind that the size of *P* or the likelihood that a finding is a chance finding, depends on two values: the magnitude of the difference and the size of the sample studied.

8.8.3 Confidence intervals

Statistical significance of the result, for example a difference, found in a particular study gives us an indication that the difference was unlikely to be explained by chance. But it does not give us an indication of the magnitude of that difference in the population from which the sample was studied. For this, the concept of confidence intervals has been developed. Different from a test of statistical significance, a confidence interval (CI), allows us to estimate whether the strength of the evidence is strong or weak and whether the study is definitive or whether other studies will be needed. If the confidence interval is narrow, the strength of evidence will be strong. Wide CIs indicate greater uncertainty

about the true value of a result. A statistician can calculate CIs on the result of just about any statistical test.

We can take an example, where an investigator found that the haemoglobin (Hb) level appeared to be different in males and females. In males, the mean Hb level was 13.2. In females, the mean Hb level was 11.7. A statistical significance test, based on a *P* value will tell us about how likely this difference is to be real, or to be a chance finding. But the statistical test does not tell us about the range of the difference that can be expected, on the basis of the data, between mean Hb levels of males and females in the whole population, if other samples were taken and studied. The difference between the two means in this particular study is 1.5. But confidence intervals could be, for example, 0.5 to 2.5.

When confidence interval (CI) reporting is used, a point estimate of the result is given together with a range of values that are consistent with the data, and within which one can expect the true value in the population to lie. The CI thus provides a range of possibilities for the population value. This is in contrast to statistical significance which only indicates whether or not the finding can be explained by chance.

As in statistical tests, the investigators must select the degree of confidence or certainty they accept to be associated with a confidence interval, though 95% is the most common choice, just as a 5% level of statistical significance is widely used.

In general, when a 95% CI contains a zero difference, it means that one is unable to reject the null hypothesis at the 5% level. If in the example above, the CI for the difference in Hb level between males and female is –0.4 to +3, we cannot reject the null hypothesis that there is no difference because the confidence interval includes 0. We do not use dash when putting the CI. It may be confusing because the CI may be minus (–). We also do not use ± because the intervals are commonly not equal.

The CI is also useful in analysing correlation. The correlation coefficient (r), as discussed in section 8.6, is measured on a scale that varies from +1 through 0 to –1. Complete correlation between two variables is expressed as 1. A statistical test of significance will tell us the probability that a degree of correlation found in the study is likely or not to be due to chance. But it does not tell us, on the basis of the data, about the range of correlation coefficients that may be expected if a large number of other similar studies is done on the same population. Confidence intervals provide this range. Again, if this range includes 0, we cannot reject the null hypothesis that actually there is no real correlation.

The two extremes of CI are sometimes presented as confidence limits. However, the word “limits” suggests that there is no going beyond and may be misunderstood because, of course, the population value will not always lie within the confidence interval. If we

have accepted a certainty level of 95% , then there is still a 5% chance that the range will go beyond the confidence interval.

8.8.4 Statistical power

A study designed to find a difference or an association may find no such difference or association. Alternatively, it may find such a difference, but application of the statistical test shows that the null hypothesis cannot be rejected. Thus any difference or association found in the study may be due to chance, and no inference can be made from it. We cannot accept this conclusion without questioning whether the study had the statistical power to identify an effect if it was there. Calculation of the statistical power helps us to know how likely a “miss” is to occur at a given effect size.

Power is an important concept in the interpretation of null results. For example, if comparison of two treatments does not show that one is superior to the other, this may be due to lack of power in the study. A possible reason could be a small size of the sample.

As discussed in Chapter 4, section 4.7, the statistical power for a given effect size is defined statistically as 1 minus probability of a miss, i.e. type II error or beta. It is commonly, but arbitrarily set, at 0.8. This means that we accept a 20% chance that a finding or a difference will be missed. The scientific tradition is to accept a lower level of certainty for not missing a finding when it is true than for accepting a finding when it is not true. This can be seen as an analogy to the judicial tradition that convicting an innocent defendant is a worse error than aquitting a guilty defendant, and requires more certainty.

8.9 Selection of statistical test

There are a large number of statistical tests for analysing scientific data. Standard textbooks can be consulted about the type of statistical test and their applications and methodology. The computer has facilitated statistical work to a great degree. A number of software packages are available, commercial and non-commercial. Microsoft Excel is a program commonly included in computer software packages. *Epi-Info* is a software program available free from the Centers for Disease Control and Prevention, Atlanta, USA, (web site http://www.cdc.gov). It was developed in collaboration with the World Health Organization, as a word-processing, database and statistics system for epidemiology to be used on IBM-compatible microcomputers. The commercial statistical software package SPSS provides a good balance of power, flexibility and ease of use. Another commonly used package is SAS. There are also other packages.

One disadvantage of computerization is that it may give investigators a blind trust in statistics as an accurate and precise science. Statistics is based on probabilities and not on certainties. Statistical calculations are based, to a certain extent, on assumptions. A complex statistical test does not necessarily mean a more robust test. A complex test may have to be based on more assumptions, and the resulting estimates may be less rather than more robust.

For large studies, the advice and help of a professional statistician should be sought from the beginning. But it is the investigator who knows the type of data and the questions to be answered, and who must fully grasp the concepts behind statistical calculations and the meaning and limitations of the exercise. Investigators should also familiarize themselves with terms used by statisticians to be able to communicate well with them. They should also understand the factors taken into consideration by statisticians when they decide on the appropriate test to be used, and the common logic behind the tests.

In general, the type of statistical test to be used depends on type of data to be analysed, how the data are distributed, type of sample, and the question to be answered.

Type of data

Statisticians use certain terms in describing the properties of the data to be analysed. The type of data influences the choice of the statistical test to be used.

For the purposes of data description, and statistical analysis, data are looked at as variables. Data are classified as either numerical or categorical. Data are classified as numerical if they are expressed in numbers. Numerical data may be discrete or continuous. Continuous variables are those which are measured on a continuous scale. They are numbers that can be added, subtracted, multiplied and divided,

Categorical variables are ones where each individual is one of a number of mutually exclusive classes. Categorical data may be nominal or ordinal. In nominal data, the categories cannot be ordered one above another. An example of categorical nominal variable is sex (male or female) or marital status (married, not married, divorced). In ordinal data, the variables can be ordered one above another. An example of ordinal categorical data is the grading of pain (mild, moderate, severe), or the staging of tumours (first stage, second stage, third stage, fourth stage).

A continuous variable may be grouped into ordered categorical variables, for example in age groups. In grouping continuous variables care should be taken that groups do not overlap, for example age groups of 1–4 years, 5–9 years, etc.

The type of statistical test applied depends on whether dealing with numerical or categorical data.

Distribution of the data

The distribution of the data is important for the statisticians. Data fall in a normal distribution when they are spread evenly around the mean, and the frequency distribution curve is bell shaped or Gaussian. For such data, which are more common, statisticians apply what they call parametric tests statistics. When the distribution curve is skewed, statisticians use other types of tests, called non-parametric or distribution free statistics.

Type of sample

Tests also differ when the data were obtained from independent subjects or from related samples such as those involving repeated measurements of the same subjects. Tests for analysis of paired and unpaired observations are different. By paired observations, we mean repeated measurements made on the same subject, or observations made on subjects and matched controls. Unpaired observations are made on independent subjects. A different type of test may also be needed if the sample size is small.

Questions to be answered

Statisticians can only look for answers to questions, which the investigators put to them. They may be asked to look at differences between groups or for an association. Selection of the appropriate statistical test for differences between groups will depend on whether investigators are looking for a difference between two groups, or are comparing more than two groups.

If investigators are looking for relationship, association and correlation, selection of the statistical test will depend on whether they are looking for an association between only two variables, or are interested in multiple variables. Univariate analysis is a set of mathematical tools to assess the relationship between one independent variable and one dependent variable. Multivariate analysis assesses the independent contribution of multiple independent variables on a dependent variable, and identifies those independent variables most significant in explaining the variation of the dependent variable. It also permits clinical researchers to adjust for differences in patient characteristics (which may influence the outcome of the study). Logistic regression is a method commonly used by statisticians in multivariate analysis.

If investigators are looking for an effect of one variable on another, they need to decide on whether they are looking to the effect in one expected direction only or without reference to an expected direction. The alternative hypothesis outlining a relationship may be directional or non-directional. For example, a relationship between smoking and cardiovascular disease can only be directional. It is not expected in the hypothesis that it may decrease cardiovascular disease. However, the relationship between oral hormonal

contraceptives and certain disease conditions, for example, can be non-directional. The disease conditions may increase or decrease as a result of oral hormonal contraceptive use. To test a non-directional hypothesis, the statistician will need to use a two-tailed test. Usually a larger sample size is needed for a two-tailed test, compared with a one-tailed test.

8.10 Examples of some common statistical tests

The following two examples illustrate the concepts behind the calculations made in statistical tests and the logic on which they are based.

The t *test*

The *t* test is used for numerical data to determine whether an observed difference between the means of two groups can be considered statistically significant, i.e. unlikely to be due to chance. It is the preferred test when the number of observations is fewer than 60, and certainly when they amount to only 30 or less. An example would be a study of height in two groups of women: one group of 14 women delivered normally and the other group of 15 delivered by Caesarean section. A difference in the average height is found between the two groups, and we want to know whether the difference is significant or is more likely to be due to chance.

The basis of the *t* test is the logic that when the difference between the two means is large, the variability among data is small, and the sample size is reasonably large, the likelihood is increased that the difference is not a chance finding. A *t* value is calculated on the basis of the difference between the two means, and the variability among the data, using a special formula.

A special statistical table has been developed to provide a theoretical *t* value, corresponding, on one side, to the significance level and on the other side, to the size of the sample studied. The significance level (*P* value or the probability of finding the difference by chance, when there is no real difference) is set by the investigator. A *P* value of 0.05 is commonly used. The sample size used by statisticians is called “degrees of freedom”. For the *t* test, the number of degrees of freedom is calculated as the sum of the two sample sizes minus 2. The concept of degrees of freedom is based on the notion that since the total of values in each set of measurements is fixed, then all the measurements minus one are free to have any value. The last measurement, however, can only have one value, the value needed to bring the total to the fixed total value of the sum of all measurements.

The calculated *t* value is then compared with the *t* value as obtained from the table. If the calculated *t* value is larger than the table *t* value, we can reject the null hypothesis at the level of statistical significance that we chose.

The *t* test was developed in 1908 by the British mathematician Gosset who worked, not for any of the prestigious research institutions, but for the Guinness brewery. The brewery employed Gossett to work out statistical sampling techniques that would improve the quality and reproducibility of its beer-making procedures. Gossett published his work under the name of "Student". The test is sometimes referred to as the Student test.

Chi-square test (χ^2)

The Chi-square test is used for categorical data to find out whether observed differences between proportions of events in groups may be considered statistically significant. For example, a study looks at a clinical trial comparing a new drug against a standard drug. In some patients, the drugs resulted in marked improvement. In others, they resulted in some improvement. In a third group, there was no improvement. The performance of the two tested drugs was different. Can this finding be explained by chance? The logic is that if the differences were large, and if the size of the sample was reasonable, the likelihood that the findings are due to chance would be less.

In compliance with the null hypothesis, we assume there is no difference, and calculate the expected frequency for each cell (marked improvement, some improvement and no improvement) if there was no difference among the groups. Then, we calculate how different are the observed results from the expected results if there was no difference. From this, using a special formula, a Chi-square value is then calculated. Because the differences between the observed and expected values can be minus or plus, the differences have to be squared before summing them up (hence the name of the test).

Statisticians have developed a special statistical table, to find the theoretical Chi-square value corresponding to what *P* value is accepted by the investigator (usually taken as 0.05), and to the size of the sample studied.

If the calculated Chi-square value is larger than the hypothetical value obtained from the table, the null hypothesis can be rejected at the specified level of probability.

8.11 Description and analysis of results of qualitative research

Description and analysis of results of qualitative research differs from quantitative data (Pope et al, 2000). Qualitative studies are generally not designed to be representative in terms of statistical generalizability. They do not gain much from a larger sample size.

The term “transferability” or external validity describes the range and limitations for application of the study findings, beyond the context in which the study was done

While quantitative analytical research starts with the development of a research hypothesis and then tests it, in qualitative research hypotheses are often generated from the analysis of data.

Unlike quantitative studies, qualitative studies deal with textual material. During data collection, the investigator may be taking notes, using an already prepared outline or checklist, or using audiotapes. Audiotapes should be transcribed as soon as possible after the interview or discussion group. Transcripts and notes are the raw data of qualitative research. They provide a descriptive record of the research, but they need to be analysed and interpreted, an often time consuming and demanding task. Analysis of qualitative data offers different challenges from quantitative data. The data often consist of a mass of narrative text.

Data immersion

The first step in the analysis of qualitative data is for the investigator to familiarize herself/himself completely with the data, a process commonly described as data immersion. This means that the researcher should read and re-read the notes and transcripts, to be completely familiar with the content. This step does not have to wait till all the data is in. It may progress as the data are being collected. It may even help in re-shaping the ongoing data collection and further refinement of the methodology. Familiarization with the raw data helps the investigators to identify the issues, themes and concepts for which data need to be examined and analysed.

Coding of the data

The next step is coding. In a quantitative questionnaire, coding is done in numbers. In qualitative analysis, words, parts of words, or combination of words are used to flag data, which can later be retrieved and put together. Codes are called labels. Pitfalls in coding should be avoided. Coding too much can conceal important unifying concepts. Coding too little may force the researcher to force new findings into existing codes, into which they do not perfectly fit.

Modern computer software can greatly enhance qualitative analysis, through basic data manipulative procedures. The type of software needed depends on the complexity of the study. For some studies, analysis can be done using a word processor with search, copy, and paste tools, as well as split screen functions. More complex studies need software specifically designed for qualitative data analysis.

For example, instead of typing every code into computer-stored text, the special software can keep a record of codes created, and allows the investigator to select from

already created codes from drop-down menus. Apart from facilitating the coding, this avoids mistakes in typing the code each time, and helps to assemble text segments for further analysis. It may also enable revising automatically a particular coding label across all previously coded text. One change in the master list changes all occurrences of the code.

Another function that can be provided by the special software program is the construction of electronic indexes and cross-indexes. An electronic index is a word list comprised of all substantive words in the text and their locations in terms of specific text, line number, or word position in a line. Once texts have been indexed, it is easy to search and find specific words or combinations of words, and to move to their next occurrence.

The software program may also construct hyperlinks in the text allowing cross-referencing or linking a piece of text in one file with another in the same or different file. Hyperlinks help to capture the conceptual links observed between sections of the data, while preserving the continuity of the narrative. Hyperlinks may also be useful when different focus group discussions have been conducted. Hyperlinks also can relate codes and their related segments to one another.

Different software packages are available. The Centers for Disease Control and Prevention (CDC), Atlanta, USA has developed packages which are free and available online from its web site (http://www.cdc.gov). Commercial software is also available.

Coding sort

The next step after coding, is to conduct a “coding sort”, by collecting similarly coded blocks of text in new data files. Coding sorts can be done manually, using highlighting and cut and paste techniques with simple word- processing software, or can be done with qualitative data analysis software. After extracting and combining all the information on a theme in a coding sort, the investigator will be ready for a close examination of the data.

Putting qualitative data in tables and figures is often called “data reduction”. A table that contains words (not numbers as in quantitative research) is called a “matrix”. A matrix enables the researcher to assemble a lot of related segments of text in one place, to reduce a complicated data set to a manageable size. Some software packages make it easy to develop such matrices. They can also be developed manually. Sometimes qualitative data can be categorized, counted and displayed in tables. Answers to open-ended questions in questionnaires can often be categorized and summarized in a table. For qualitative data, a diagram is often a figure with boxes or circles containing variables and arrows indicating the relationship between the variables. Flow charts are special types of diagrams that express the logical sequence of actions or decisions.

References and additional sources of information

Briscoe MH. *A researcher's guide to scientific and medical illustrations*. New York, Springer-Verlag, 1990.

Browner WS et al. Getting ready to estimate sample size: hypotheses and underlying principles. In: Hulley SB, Cummings SR, eds. *Designing clinical research: an epidemiologic approach*, 2nd edition. Philadelphia, Lippincott Williams & Wilkins, 2001: 51–62.

Gardner MJ, Altman DG. *Statistics with confidence: confidence intervals and statistical guidelines*. London, BMJ Books, 1997.

Gehlbach SH. *Interpreting the medical literature*. 3rd edition. New York, McGraw-Hill Inc., 1993: 138.

Hill AB. *Principles of medical statistics*, 9th edition. New York, Oxford University Press, 1971.

Malterad K. Qualitative research: standards, challenges, and guidelines. *Lancet*, 2001, 358: 483–88.

Medawar PB. *Advice to a young scientist*. New York, Basic Books, 1979: 39.

Pope C, Ziebland S, Mays V. Qualitative research in health care: Analysing qualitative data. *British Medical Journal*, 2000, 320:114–116.

Swinscow TDV, Campbell MJ. *Statistics at square one*. 10th edition. London, BMJ Books, 2002.

Weaver JH. *Conquering statistics: numbers without the crunch*. Cambridge, Massachusetts, Perseus Publishing, 2000: 5–12.

Chapter 9

Interpreting research results

9.1 Introduction

Researchers should describe their results clearly, and in a way that other researchers can compare them with their own results. They should also analyse the results, using appropriate statistical methods to try to determine the probability that they may have been chance findings, and may not be replicable in larger studies. But this is not enough. Results need to be interpreted in an objective and critical way, before assessing their implications and before drawing conclusions. Interpretation of research results is not just a concern for researchers. Health professionals reading or hearing research results should be able themselves to interpret them correctly, and to assess their implications for their work. Policymakers should also be aware of the possible pitfalls in interpreting research results and should be cautious in drawing conclusions for policy decisions.

9.2 Interpreting descriptive statistics

The mean or average is only meaningful if the data fall into a normal distribution curve, that is, they are evenly distributed around the mean. The mean or average, by itself, has a limited value. There is an anecdote about a man having one foot on ice and the other in boiling water; statistically speaking, on average, he is pretty comfortable. The range of the data, and their distribution (expressed in the standard deviation) must be known. It is sometimes more important to know the number or percentage of subjects or values that are abnormal than to know the mean.

Descriptive statistics cannot be used to define disease. The average should not be taken to indicate the "normal". The standard deviation should not be used as a definition of "normal" range. To allow a cut-off point in a statistical distribution to define a disease is wrong. This is particularly important in laboratory data, where the range of normal is often based on measurements in a large number of healthy people. The standard deviation is based on the values in 95% of the apparently normal healthy people. Outlying values are considered abnormal though they do not indicate disease. With the modern tendency of using a large battery of laboratory tests for each patient, the likelihood of so-called abnormal values becomes high. For example, when 5% of each of 20 biochemical determinations in healthy people are routinely classified as deviant, the likelihood that

any non-diseased individual will have all 20 determinations reported as normal will be only 36% (Gehlbach, 1993). Graphs may distort the visual impression of relationships, if the scale on the *x* and *y* axes is put in different ways. An association or correlation does not mean causation. An association or correlation needs explanation. Because of the importance of this question, it will be dealt with in detail in another section in this chapter.

9.3 Interpreting "statistical significance"

Albert Einstein said, "Not everything that can be counted counts, and not everything that counts can be counted." A statistically significant finding simply means that it is probably caused by something other than chance. Significant does not mean important.

To allow proper interpretation, exact *P* values should be provided, as well as the statistical test used. The term "orphaned" *P* values is used to describe *P* values presented without indication of the statistical test used.

Statistical tests need to be kept in proper perspective. The size of the *P* value should not be taken as an indication of the importance of the result. The importance of the result depends on the result itself and its implication. Results may be statistically significant but of little or no importance. Attaching a fancy *P* value to trivial observations does little to enhance their importance. A statistically significant or even a highly significant difference does not necessarily mean a clinically important finding. A difference is a difference only if it makes a difference.

Differences may not be statistically significant but may still be important. The differences may be real but, because of the small size of the sample, they are not statistically significant. A *P* value in the non-significant range tells you that either there is no difference or that the number of subjects is not large enough to show the difference. As discussed in Chapter 8, the study may not have had the power to show an effect of that size.

9.4 Bias

All studies are potentially subject to bias (literally defined as systematic deviation from the truth). Bias is a systematic error (in contrast to a random error due to chance). The effect of bias is called "like is no longer compared with like". Bias has a direction. It either increases or decreases the estimate, but cannot do both. This is in contrast to chance findings that can have any effect on the estimate.

If the study sample is not representative of the population, the inference we make from the result may be misleading. Analytical statistics will be of no help if the sample

is not representative. Analytical statistics cannot correct our mistakes in designing the study. Every attempt should be made in the design of the study to ensure that the sample is representative. Bias cannot be addressed or corrected by statistics. The main protection is to think of the possibility of the bias and design it out. Using sophisticated computer programs does not guarantee the validity of the study. In computer jargon, they say "garbage in, garbage out". If you feed the computer with the wrong information, you will get a wrong outcome. If the possibility of bias cannot be avoided completely in the planning of the study, the investigators must point this out when they present the findings of the study. Bias can occur when groups being compared differ systematically in a way that is related to the outcome. Main types of bias can occur at two levels: at the level of selection of subjects (selection bias) and at the level of collecting the information (information or measurement bias).

Selection bias is a systematic difference between subjects selected for a study and those who are not selected. Loss to follow-up can cause a selection bias. Attrition is the term used for reduction in the number of subjects who remain in a study. Attrition bias occurs when the subjects who drop out of a study are systematically different from those who complete the study. For example, those who develop complications or side-effects may be more likely to drop out of the study. Response bias is a specific type of selection bias in which respondents differ systematically from non-respondents to a questionnaire.

Measurement or information bias occurs when the methods of measurement or obtaining information are consistently dissimilar in different groups of patients. One type is recall bias. It is encountered, for example, when people with a certain condition are more likely to remember exposure to the variable under study than people without the condition. An example is if a study tries to compare the frequency of past oral contraceptive use among women admitted to hospital because of thrombophlebitis and a group of women admitted for other reasons. It is entirely possible that women with thrombophlebitis, if aware of the reported association between oestrogens and thrombotic events, might report use of oral contraceptives more completely than women without thrombophlebitis. A special type of information bias is surveillance or diagnostic suspicion bias encountered when patients with a risk factor may be tested for the outcome more frequently and carefully than those without the risk factor. For example, women using hormonal contraceptives may be screened more frequently and more carefully for neoplasia of the uterine cervix, than other women, leading to the diagnosis of more cases.

9.5 Confounding

In simple terms, confounders are all of the "other things" that could explain the result of the research. A careful investigator should look for all possible explanations of the results, before making a conclusion. In good scientific thinking, one should not try to assume one interpretation of the results, when other interpretations are also possible.

For example, a study may find that the risk of lung cancer is more in manual workers. A good investigator will not assume that manual work predisposes to lung cancer, before looking for other possible explanations. The result may, for example, be due to a fact that manual workers are more likely to smoke and it is smoking, not manual work, which is associated with lung cancer.

Another example is when an association is reported between herpesvirus infection and cervical cancer. Both herpesvirus and a number of other infectious agents, which may possibly cause cervical cancer, are transmitted by sexual contact. There is strong evidence that human papilloma virus infection leads to cervical cancer. It could be that the higher prevalence of herpesvirus infection in women with cervical cancer is only a consequence of greater sexual activity and so is indirectly related to a true cause, which is also transmitted sexually.

There are three ways to deal with confounding: to think of it in planning and designing the study; to measure/record the presence of the confounder during implementation of the study; and to allow for it in the analysis.

The case mix or patient mix, which refers to baseline differences among research subjects, can be a confounding factor. Matching is an important technique for creating a control group by pairing subjects, based on one or more confounding factors. An alternative is to use the control-table method, in which stratification is done afterwards. Rather than arranging subjects by groups as the study is designed, results are calculated within specified subdivisions. When more than a few confounding variables are present, the statistical technique of multivariate analysis is used.

As an example of analysis for confounding factors we may look at a study of the relationship between the working status of mothers and the duration of breastfeading. The study may show that women who are employed full-time are less likely to breastfeed for a long duration than women who are employed part-time and women who are not employed. However, the level of education of the mother may be a confounding variable, since it can affect the outcome (duration of breastfeeding) and it may correlate with the working status. Before blaming work for the shorter duration of breastfeeding, there is a need to consider the confounding factor of education. Stratification may be used. A cross-tabulation table may be constructed for mothers at different educational levels, for example those who had no schooling, less than 5 years of schooling, 5–9 years and 10 years or more. For each table, we look at duration of breastfeeding in mothers who

are employed full-time, employed part-time and not employed. An alternative way of considering this confounding factor is matching at the design and implementation phase. For each employed mother with less than 5 years of schooling, we would choose a non-employed mother with a similar educational level.

Crude rates are the terms used when results have not been adjusted for confounding factors. Adjusted rates are the terms used when results have undergone statistical transformation to permit fair comparison between groups differing in some characteristic that may affect risk of disease.

9.6 Making the case for causation

The association of two variables does not necessarily mean causation and should not be interpreted as a causal relationship. Historically, scientists had to struggle with this issue, in the early days when microbiologists began to discover and report on the association of certain microorganisms with some disease conditions. In 1882, Koch stipulated that for an infectious agent to be considered the cause of a disease, the following criteria must be established:

- the organism must be present in every case of the disease;
- the organism must be isolated and grown in pure culture;
- the organism must cause a specific disease when inoculated into an animal; and
- the organism must then be recovered from the animal and identified.

A false appearance of association can occur through three mechanisms: chance, bias, or confounding. But, even after excluding, to the best of our effort, the possibilities of chance, bias and confounding variables, other criteria are needed to turn an association into causation. Sir Austin Bradford Hill proposed a set of features that should be sought when deciding whether a relationship is causal or just an association (Hill, 1965). They are still valid and are referred to as the Bradford Hill criteria:

- strength of the association
- consistency of the observed evidence
- specificity of the relationship
- temporality of the relationship
- biological gradient of the dose–response
- biological plausibility
- coherence of the evidence
- experimental confirmation
- reasoning by analogy.

- Strength of the association: A strong association between a purported cause and effect, as expressed for example by a large relative or absolute risk, is better evidence for a causal relationship than a weak association.
- Consistency of the observed evidence: When several studies conducted at different times in different settings and with different kinds of patients all come to the same conclusion, evidence for a causal relationship is strengthened.
- Specificity of the relationship: Specificity (one cause, one effect) is more often found for acute diseases. But for other diseases, there are often many causes for the same effect, and many effects may arise from the same cause. An example is the association of smoking and lung cancer. Smoking causes other diseases and lung cancer has other causes. Strong specificity is evidence for cause, but the absence of specificity is only weak evidence against a cause-and-effect relationship.
- Temporality of the relationship: Causes should obviously precede effects. This self-evident principle may, however, be overlooked when interpreting cross-sectional or case-control studies, in which both the cause and effect are measured at the same point in time.
- Biological gradient of the dose–response: A dose–response relationship is present when varying amounts of the purported cause are related to varying amounts of the effect. An example of a dose–response curve is when lung cancer rates are plotted against number of cigarettes smoked.
- Biological plausibility: When the assertion of cause and effect is consistent with our knowledge of the mechanisms of disease, as they are currently understood, plausibility is often given considerable weight when assessing causation. It is important, however, to remember that what is considered biologically plausible depends on the state of medical knowledge at the time.
- Coherence of the evidence: A factor is also more likely to be a cause of disease if its removal results in a decreased risk of disease. For example, if people give up smoking does this decrease the likelihood of lung cancer?
- Experimental confirmation: A causal association is more likely if supported by experimentation in animals.
- Reasoning by analogy: The argument of analogy for a cause-and-effect relationship is strengthened if there are examples of well established causes and effects that are analogous to the ones in question.

9.7 Interpreting end points to measure the outcome

The use of one end point may ignore the possible effect on other variables that may have a clinical impact. For example, one study reported a 44% reduction in heart attack for physicians who took low-dose aspirin. But there was a trend toward increased haemorrhagic stroke among treated subjects and an undiminished overall death rate from cardiovascular causes. The aspirin appears to be performing the role of platelet inhibition quite well, but selecting only one end point would have masked its possible other effects (Steering Committee of the Physicians' Health Study Research Group, 1989).

It is also important to consider when an end point or outcome occurs in relation to the intervention. Outcomes that occur long after subjects stop taking a drug, or before benefits of an intervention can logically be expected, can complicate the interpretation.

Surrogate end points are sometimes used to determine an outcome, and caution should be exercised in extrapolating from the result. A surrogate end point can be defined as a variable that is relatively easily measured and that predicts a rare or distant outcome, but which is not itself a direct measure of either harm or clinical benefit. The two main advantages of the use of end points are that they can considerably reduce the sample size, duration, and therefore cost of studies. They can also allow treatments to be assessed in situations when the use of primary outcomes would be excessively invasive or unethical. For a surrogate end point to be a good measure of the outcome it must have a good positive predictor value, and a good negative predictor value, i.e. it should be both sensitive and specific. It is not enough for the end point to be biologically plausible to draw clinical conclusions. The end point should also be amenable to quality control monitoring. Examples of the use of surrogate end points include the use of lipid profile as a surrogate for the development of cardiovascular disease, and the use of the CD4 cell count to predict survival in HIV infection.

9.8 Interpreting studies of risk factors

Studies of risk factors are very important in the prevention of disease. They often attract public and media attention. Unless interpreted properly, they can lead to misinformation. Studies of risk factors cannot be interpreted without proper understanding of the following concepts: basic risk, relative risk, confidence intervals, attributable risk, and balancing risks and benefits.

Basic risk

Basic risk statements express the likelihood that a particular event will occur within a particular population. For example, the US Women's Health Initiative (WHI) Study reported in July 2002, that postmenopausal women using hormone replacement therapy

(HRT) had an incidence of colorectal cancer of 10 per 100 000 women per year. This basic risk, however, does not mean much unless we know how many women would have developed the same disease condition without having used HRT. Without this information, we cannot interpret the finding. It may indicate a risk, but it may also indicate a protective effect. In the placebo-controlled group followed up in the same study, the incidence was 16. This means that HRT was actually protective against colorectal cancer (Writing group for the women's health initiative investigators, 2002).

Relative risk

Relative risk is the ratio of the incidence of outcome in the exposed group to the incidence of the outcome in the unexposed group.

The odds ratio, a term used in case-control studies, measures the odds of having the risk factor among people with the disease divided by the odds of having the risk factor among people without the disease.

Confidence interval

The statistical concept of confidence interval has been discussed in Chapter 8. A study that reports the relative risk or the odds ratio without reporting the confidence interval cannot be adequately interpreted. Confidence intervals should always be presented for the relative risk and odds ratio. The important feature to look for in assessing the confidence interval is whether the boundaries include unity. A relative risk or odds ratio of 1 means there is no association between the risk factor and the disease. A relative risk may be much higher than 1.0, but if the 95% confidence interval overlaps 1.0, it can be concluded that the increase in risk is not statistically significant, and could have been a chance finding.

Attributable risk

The importance of a risk factor cannot be interpreted on the basis of the magnitude of the relative risk, without relating it to the prevalence of the particular disease condition. The term attributable risk is an estimate to quantify the contribution, which the particular risk factor makes in producing the disease within a population. The following two examples illustrate the importance of calculating the attributable risk.

The relative risk of lung cancer due to smoking is much greater than is the relative risk of myocardial infarction among smokers. However, heart disease is much more common than lung cancer. So, even though the risk associated with heart disease and smoking is small, its importance to the general health is magnified by its relatively higher incidence.

Women taking oral contraceptive pills are more likely to have a fatal heart attack than women not taking the pill. However, women of reproductive age (pill users) have a low incidence of myocardial disease. This increased relative risk translates into an attributable risk of only very few deaths per 100 000 users per year.

Balancing risks and benefits

Decisions cannot be made on risks alone, if there are benefits as well. This applies, for example, to the case of oral contraceptives, which have health risks, but also health benefits which outweigh the potential risks.

9.9 Interpreting studies of diagnostic tests

The term diagnostic test is used broadly to describe the value of a symptom (or collection of symptoms), signs or special investigation in the diagnosis of a clinical condition or health situation. In order to assess the diagnostic worth of tests, they must be compared to the gold standard, i.e. the best test currently available. Otherwise, there is no way to make sure that the test is diagnosing the condition in question. Proclamations of highly statistically significant associations between the test and the diagnosis are not sufficient. It is essential to provide information on the extent to which diagnostic tests misclassify subjects, i.e. make a diagnosis of a disease when it is not present, or miss the diagnosis when a disease is present. For this, the concepts of sensitivity, specificity, predictive value and efficiency are used.

Sensitivity

Sensitivity is the ability of a test to single out people who have the disease. Low sensitivity will mean that there will be many false negatives.

Specificity

Specificity is the ability of a test to label people who do not have the disease as negative. Low specificity means there will be many false positives.

Predictive value

The predictive value of a test gives the frequency with which a positive test actually signifies disease. It is more appropriately labelled positive predictive value. (The negative predictive value is less often used).

Efficiency

Efficiency is an overall estimate of a test's ability to classify patients correctly. It is estimated by adding the numbers of the two correct classifications (true positive and true negative) and dividing by the total number of patients assessed.

Balancing sensitivity and specificity

Diagnostic tests cannot be expected to be perfect. Increasing the sensitivity of a test often results in decreased specificity and vice versa. The value of a test cannot be made on the basis of sensitivity and specificity alone, or on the basis of overall efficiency. Much depends on the prevalence of the disease condition. The predictive value is at the mercy of prevalence. Even a small percentage of false positives can become magnified when a disease is rare.

Choices between sensitivity and specificity must be made. The decision is not statistical; it is clinical and economical. When the consequences of missing a disease are crucial, sensitivity is paramount. But if the burden of creating false positives outweighs the advantages of capturing all cases of a disease, increasing specificity should be the goal.

With the economic aspects of health care drawing increasing attention, costs are becoming a concern in the evaluation of screening procedures. The following example of mandatory premarital serological testing of HIV infection illustrates the point. A test used may have a sensitivity of 98%, and a specificity of 99%. This sounds like an impressive performance. But in a community where HIV prevalence is low, even the low rate of false positives will mean that a large number of people will be unnecessarily alarmed. It also means that a very large number of people will need to be tested in order to detect one single true positive case. An estimate will need to be made of the cost of diagnosing one single case.

Many diagnostic tests yield continuous results, for example serum levels of prostate specific antigen (PSA) as a screening test for prostate cancer. With such tests, a decision must be made as to what value will constitute a positive test, a value called the "cut-off point". This decision requires trading an increase in specificity for a decrease in sensitivity, or vice versa. Receiver operator characteristic (ROC) curves are useful for visualizing and selecting the most appropriate cut-off point for screening tests. The terminology comes from its first use in the field of electronics. ROC curves are a graphic way of portraying the trade-offs involved between improving either a test's sensitivity or its specificity when different cutoff values are selected. For each cut-off point, statisticians plot the sensitivity on the vertical axis, and the value that is 1 minus specificity (false positives) on the horizontal axis (Newman et al. 2001).

9.10 Interpreting studies that report the results of interventions

Results based on uncontrolled studies do not mean much. In many cases, it is possible that no treatment could have achieved comparable results.

Results of controlled studies based on comparison of a treatment with a placebo do not mean much, if there are other available treatments. Results should be based on comparator drugs currently available.

The preference of one drug over another should not be based on one aspect only of its performance. It should give equal consideration to the four elements of the acronym "STEP": safety, tolerability, efficacy, price.

A study may indicate that a new drug has resulted in more improvement than the currently available and used drug. A statistical test shows that this difference between the two drugs is statistically significant. What the statistical test says is that the difference in the result is unlikely to have happened by chance, and that the probability of its being a chance finding is less than 5%. The *P* value can tell us how remote the possibility that the difference can be found by chance. This, however, does not mean that the result is clinically significant. The *P* value does not tell us anything about the magnitude of the difference between the two treatments, and how the point of estimate of the difference will be changed if other samples from the same population are studied. For this we need to know the confidence intervals for the difference between the performance of the two drugs. Confidence intervals will show the likely range of the magnitude of the difference in the performance of the two drugs.

Even if a drug is proven to be superior to another, the question still remains about what this means to the individual patient or individual clinician. One has to calculate the number needed to treat in order to achieve the therapeutic advantage of the new intervention. If for example, one intervention reduces the absolute risk of dying by say 4%, it means that the number needed to treat with the new intervention to avoid one death will be 25. The "number needed to treat" has important cost implications when the result of the intervention is interpreted.

9.11 Interpreting results of qualitative research

Qualitative research methods involve the interpretation of textual material derived from talk or observation.

In interpreting qualitative findings, the investigators should carefully look into their credibility, dependability, confirmability and transferability.

Credibility means interpreting the qualitative data in a way that offers explanations that are consistent with the data collected. Negative findings should be adequately presented and addressed, and alternative explanations considered. As in quantitative research, the investigators should look for confounding variables. For example, a study may reveal that homes sprayed for malaria control had a higher incidence of malaria when sprayed in the afternoon. It could be that sprayers used most of the spray in the morning so that the load to carry in the afternoon would be lighter. To ensure credibility of the interpretation, the investigator should act as the "devil's advocate", considering all potentially competing explanations of the results.

Possible sources for bias should be checked, for example observer bias or the influence of the researcher on the research situation. A researcher's background and position will affect the process of qualitative research. The investigator always enters a field of research with certain opinions about what it is all about. In qualitative research, this potential bias cannot be eliminated, but it should be exposed in a process termed "reflexivity". Reflexivity starts by identifying preconceptions brought into the project by the researcher, representing previous personal and professional experiences, pre-study beliefs and qualifications for exploration of the field. During all steps of the research process, the effect of the researcher should be assessed, and, later on, shared. Adequate accounts of these effects should be considered in the limitations and strengths of the study, and transferability of findings.

Dependability means that data can be replicated. The replication is not necessarily of the results, but of the process used to obtain the results. Other investigators should be able to replicate the study.

Confirmability means that other researchers can have access to the data and can do their own analysis. The concept of "audit trail" enables others, on the basis of the collected data, to review the analysis decisions and verify the interpretations.

Transferability means the use of the findings to make inferences to other populations. This may not be possible because qualitative research is often context-specific. Qualitative research emphasizes depth more than breadth, and insight rather than generalization. In such cases, however, there are lessons learnt that may help in understanding the situation in other populations.

References and additional sources of information

Altman DG, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*, 2001, 134: 663–694.

Bossuyut PM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *British Medical Journal*, 2003, 326: 41–44.

Newman B, Browner WS, Cummings SR. Designing studies of medical tests. In: Hulley SB, Cumming SR, eds. *Designing clinical research: an eidemiologic approach*, 2nd edition. Philadelphia, Lippincott Williams & Wilkins, 2001: 175–192.

Byrne DW. *Publishing your medical research paper*. Baltimore, Lippincott Williams & Wilkins, 1998.

Devers KJ, Sofaer S, Rundall TG, eds. Qualitative methods in health services research: A special supplement to HSR. *Health Services Research* 1999, 34 (5) Part II: 1083–1263.

Doll R, Hill AB, Mortality in relation to smoking: ten years' observation of British doctors. *British Medical Journal*, 1964, 1:1399–1460.

Einstein A. http://www.memorable quotations.com/einstein.htm (accessed 30/04/2004).

Gehlbach SH. *Interpreting the medical literature*, 3rd edition. New York, McGraw-Hill Inc., 1993.

Greenhalgh T. *How to read a paper: the basics of evidence-based medicine*. London, BMJ Books, 1997.

Hill BA. The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, 1965, 58: 295–300.

Huff D. *How to lie with statistics*. London, Penguin Books, 1991.

Malterad K. Qualitative research: standards, challenges, and guidelines. *Lancet*, 2001, 358: 483–88.

Schulz KF, Grimes DA. Uses and abuses of screening tests. *Lancet*, 2002, 359: 881–884.

Steering Committee of the Physicians' Health Study Research Group: Final report on the aspirin component of the ongoing physicians' health study. *New England Journal of Medicine*, 1989, 321:129.

Writing group for the Women's Health Initiative investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *Journal of the American Medical Association*, 2002, 288: 312–333.

Chapter 10

Communicating research

> If you have an apple and I have an apple and if we exchange these apples then you and I will still each have one apple. But if you have an idea and I have an idea and we exchange these ideas, then each of us will have two ideas.
>
> George Bernard Shaw

10.1 Introduction

Research is not complete until it is written up and its results shared, not only with other scientists who may build upon it to further advance the science, but also with those who may benefit from it, who may use it, and who have a stake in it. Etymologically, the Latin "communio" relates to participation and sharing. Communication and sharing are two words sharing the same concept.

It is an ethical duty to communicate research results. Editors of scientific journals should consider seriously for publication any carefully done study of an important question relevant to their readers, whether the results are negative or positive. Failure to submit or publish studies with negative findings contributes to publication bias. Pharmaceutical companies have occasionally been held guilty of suppressing research results that show that their products may not be as safe as they claim.

Researchers normally communicate their results to other scientists, by publishing in peer-reviewed journals and presentation in scientific meetings. The internet is revolutionizing the dissemination of scientific information in ways never thought possible before. If the research was funded, researchers have an obligation to submit periodic reports to the funding agency.

The primary aim of health research is to improve health. To achieve this aim, results of research should not be communicated to other scientists only. The information has to reach the health professionals. Research with practical implications should be scrutinized, synthesized and presented in the form of evidence-based reviews and guidelines about best practices. There is a growing awareness of a gap between clinical practice and the findings of research.

If research is to inform public policy, it should be properly communicated to policy makers. Sending a report is not enough. The research should be presented and discussed.

There is a need to communicate scientific information to patients. Patients need to participate in making informed decisions and choices about their treatment options. Health professionals should always keep in mind that they do not treat diseases; they treat patients who have their preferences, values and rights. An informed patient is also more likely to follow prescribed treatment, which is often ambulatory and self-administered. Packages of prescription drugs normally include an insert for patient information. The material in this insert is closely scrutinized for accuracy by the drug regulatory agency. Health lifestyle behaviour is a powerful determinant of health. For certain health conditions, it can be more important than the provision of health care. Empowering people with valid scientific information is more likely to induce a healthy lifestyle. Educating patients about the effectiveness of interventions is sometimes advocated as a way of changing the behaviour of health professionals who may be reluctant to change their traditional ways of treatment and to adopt more novel approaches. Pharmaceutical companies are now exploiting this patient-centred approach by targeting patients for their messages in public media.

Communities that have participated in research are entitled to know about the outcome of the research and its implications for them. Health researchers need to engage the public in what they are doing and what they hope to achieve. For one thing, science needs a favourable public environment. For another thing, there is a growing need to ensure and maintain public trust in science. Research can only thrive in a favourable scientific environment. Chairman Mao once said, when talking of revolutionaries, “the fish need a sea to swim in”. Science also needs a sea to swim in. It can only thrive if a culture for research is present in the society. Without a strong appreciation of science in society, the introduction of technology-driven solutions to everyday problems will be more difficult than it should be. Science should become more comfortably enmeshed in society’s collective consciousness. In this favourable environment, people volunteer as research subjects when they know that the scientific benefit will accrue to others, not to themselves. A distinction should be drawn between the public understanding of science and the public appreciation of science. It does not actually matter whether the public can distinguish a proton from a protein, in order to appreciate science.

This chapter provides general guidance to researchers on communicating their results to other scientists, to the funding agencies, to health professionals, to policy-makers, to patients and to the public at large.

10.2 Communicating to scientists

10.2.1 Publication in scientific journals

Scientists always aim to publish their research findings in scientific journals that are peer-reviewed, that are indexed and that have a high impact factor. Peer-reviewed

journals are journals in which the articles are vetted by independent referees for quality and interest, and are therefore more highly regarded by researchers. Articles published in journals that are indexed by indexing services, such as the Index Medicus, are retrievable and accessible to other researchers, ensuring wider dissemination to the scientific community. Journals are ranked by their impact factor, a term used to indicate how many times, on average, journals papers are cited. This concept and its shortcomings are discussed in more detail in Chapter 14 on assessment and evaluation of research.

It should be realized that much important research is conducted that does not make it into major international journals. Journals can only publish a fraction of all papers submitted to them. There may also be a bias towards publications from institutions in industrialized countries. The tools of the information age hold considerable promise for developing country researchers, enabling them to disseminate the results of their research more widely.

There is growing understanding of intellectual property rights by academics and their institutions, and an increasing knowledge of how to do licensing deals. A distinction is drawn between what is patentable and what is not. Publication of scientific findings puts the findings in the public domain, and jeopardizes any patent application. If there is no patent protection, industry will not be interested in the discovery. Publications from major university centres are now screened for patentable discoveries before proceeding with publication. Many universities now employ patent lawyers.

Chapter 11 provides detailed guidelines on how to write a scientific paper, and Chapter 12 on how to get it published.

10.2.2 Presentations in scientific meetings

Presentation of papers in scientific meetings is another important venue for scientific communication. For many years, it was the major venue of communication among scientists. Researchers should train themselves in the art of scientific presentation. There are both advantages and disadvantage to presenting papers in scientific meetings compared with publishing. Among the advantages are that the information presented is up to date (there is usually a long time lag before a paper is published in a reputable journal), and that presentations allow discussion and questions to the authors, provide an opportunity for meeting other researchers interested in the same topic and promote networking in research. Among the disadvantages are that scientific presentations are not subjected to the same level of peer review and are not retrievable in the literature. A paper presented at a scientific meeting can be submitted subsequently for publication provided that the conference papers as a whole have not already been published in a journal. However, papers presented at scientific meetings usually need substantial reworking before full publication.

10.2.3 The age of paperless papers

The dream that the results of the world's biomedical research can be disseminated freely and widely to all may not be far away. The World Wide Web allows the distribution of information at only a fraction of the cost of distribution on paper. The internet was originally created as a place for scientists to do science. Until just a few years ago, researchers were the main inhabitants of cyberspace. The internet is now changing the process of research publication, ushering in a new age of paperless papers.

The time between the day an article is submitted to a traditional journal to the day it reaches a subscriber's hands can amount to months of peer-review, editing, proof approval and simple queuing for space. To bypass paper altogether, some journals have already adopted the online system for manuscript submission, tracking and peer review. Other journals will soon follow. Researchers submit their articles by e-mail (instantly verifying that it has been received), and editors send them by e-mail for peer review without the delays involved in mailing. Reviewer's comments are sent by e-mail to the authors, and requested revisions are sent back by e-mail. Through a tracking system, using the internet, authors can check on the paper's status. Turnaround times can be much shorter as a result (weeks rather than months).

Some journals post early online papers on their websites ahead of print publication. The version of an article published early online is the definitive version, which will be identical in content to that published in the print journal. When the final article is assigned to an issue of the journal, the early online version is removed.

Electronic journals are supplementing and, in some cases, replacing paper journals. Subscription prices for electronic journals are a fraction of those for paper publications, and are sometimes free. As of the end of 1996, there were 306 electronic journals, 70% more than in the year before, and including the fields of maths, physics, chemistry, biology, medicine and the social sciences. The number has since been increasing.

It is too early to predict the death of the biomedical journal as we know it. But we are certainly experiencing a dramatic metamorphosis of the tools of scientific communication. The World Wide Web makes it inevitable that new systems for disseminating research will partly replace or supplement traditional journals.

Publishing peer-reviewed original research has additional costs, even on the internet. Currently, subscribers meet the costs. A new model is now being experimented with whereby authors (or their institutions or funders) pay the costs of peer-review and electronic dissemination of their articles. Experiments with the "author pay" model are already under way. BioMed Central (http://biomedcentral.com/) is an independent commercial publisher, committed to providing free and immediate online access to the full text of peer-reviewed biomedical research. Authors retain copyright. BioMed Central has more than 90 peer-reviewed journals spanning the fields of biology and medicine,

and provides free technical support and hosting for groups of researchers wanting to run online, open access, peer-reviewed journals under their own editorial control. The company receives no support from governments or from scientific societies. Instead of charging users, BioMed Central covers the costs of peer review and publication by charging authors for processing manuscripts. The charge, US$ 500 per published article in 2003, can be paid directly by authors, usually from their research funds, or via their institutes through BioMed Central's membership scheme. In 2003, BioMed Central had 291 institutional members from 29 countries. The charge is waived for authors from developing countries and others who are unable to pay. Widespread adoption of a US$ 500 charge per published article would represent a ten-fold saving for science and society. It has been estimated that the scientific community currently pays about US$ 5000 per published article (based on publishers' gross revenues from journal subscriptions). It has been estimated that between 1999 and 2002, the global medical publishing sector grew by an estimated 20%, taking its revenue to US$ 2.69 billion.

The Public Library of Science PLoS (http:\\www.plos.org/), a non-profit organization of scientists and physicians, is another initiative committed to making the world's scientific and medical literature a freely available public resource. It is being funded during its first four years by a US$ 9 million grant from the Gordon and Betty Moore Foundation. The Internet and electronic publishing enable the creation of public libraries of science containing the full text and data of any published research article, available free of charge to anyone, anywhere in the world. To realize this potential, a new business model for scientific publishing is required that treats the costs of publication as the final integral step of the funding of a research project. To demonstrate that this publishing model will be successful for the publication of the very best research, PLoS plans to publish its own peer-reviewed journals. PLoS Biology launched its first issue on October 13, 2003, in print and online. PLoS Medicine will follow in 2004. PLoS Biology plans to meet its costs by charging authors $1500 for each published paper. If accepted for publication, the article will be made immediately and freely available online.

10.3 Communicating to funding agencies

Researchers need to report regularly to the funding agency on the progress of their research. Most agencies require a yearly progress report. A few require six-monthly reports. Normally in multi-year funding of a project, funding for the next period is contingent on the receipt of a satisfactory progress report, as well as a financial report on the expenditure during the period covered by the report.

The progress report should provide information to satisfy the agency about the progress of the project. Any problems encountered should be presented. The plan for the next period should be clearly outlined. Any papers submitted, accepted or published

should be mentioned. The financial report should be itemized. If the research did not go on schedule, the investigators may request a no-cost extension of the grant. At the end of the grant, a more detailed final report is expected. A final financial report is needed to close the books on the grant.

10.4 Communicating to health professionals

Researchers have a collective responsibility to ensure that health care providers have access to scientific evidence tailored to their needs. Reliance on passive diffusion of information to keep the knowledge of health professionals up to date is not enough. Although the skills for searching for evidence and critically appraising it need to be mastered, most health professionals cannot keep up with the strides of scientific knowledge. About two million articles on medical issues are published every year. An editorial in the British Medical Journal calculated in 1995 that, for doctors to keep up to date with the explosion of scientific information in their specialty, they need to read about 17 articles a day every day of the year. Most results from research appear first in peer reviewed journals. The small number of studies with practical implications for health professionals is spread thinly through a vast number of publications. The evidence from these studies needs to be synthesized.

Researchers can and should help in communicating new information to health professionals in a manner that is tailored to their needs. The development and publication of evidence-based reviews and clinical practice guidelines are examples of how this communication can be achieved.

Evidence-based reviews

There are now an increasing number of journals and abstracting services that review important papers rigorously and present the results in a way that busy health professionals can easily grasp. An example is *Evidence-based medicine*, published through a collaboration between the American College of Physicians and the BMJ (British Medical Journal) Publishing Group. It includes abstracts and commentaries from most specialties, with preference given to studies that cover conditions that are commonly encountered in practice. It also publishes systematic reviews and editorials of general interest. As the editors put it, the journal will publish the gold that intellectually intense processes will mine from the ore of about 100 of the world's top journals. The journal is available online (http://ebm.bmjjournals.com/), with free access for professionals from low-income and low-middle-income countries.

Systematic reviews of research, such as the work done by the Cochrane Collaboration have also become a useful resource, as described in Chapter 14.

Clinical practice guidelines

The medical literature can be biased towards innovations. But innovations need to be critically assessed. The challenge is to promote the adoption of those innovations that have been proven to be beneficial, to delay the spread of innovations not yet shown to be effective, and to prevent the uptake of ineffective or potentially harmful innovations. There are dangers in uncritical acceptance of medical innovations by health professionals.

Systematic reviews of evidence will not always lead to clear and unambiguous recommendations. Rigorously developed guidelines can translate complicated research findings into actionable recommendations for clinical practice. Evidence-based clinical practice guidelines can decrease the use of inappropriate health care and can promote the introduction of new knowledge about best practices. A growing number of guidelines are being developed after exhaustive reviews of evidence, by a multitude of professional organizations.

To be useful, guidelines should balance the strengths and limitations of all relevant research evidence with the practical realities of the health care and clinical settings. They should also acknowledge the uncertainty. Authoritative medicine is giving way to evidence-based medicine. Uncertainty makes it difficult to make definite recommendations, based on evidence, in all situations. Based on the available level of evidence, recommendations on management or interventions are now commonly graded according to the following categories (ACOG, 1998):

A. There is good evidence to support the recommendation.
B. There is fair evidence to support the recommendation.
C. There is insufficient evidence to support the recommendation; however, the recommendation may be made on other grounds.
D. There is fair evidence against the recommendation
E. There is good evidence against the recommendation.

10.5 Communicating to policy-makers

Health policy-makers need adequate and scientifically validated information to make evidence-based policy. Where the research has policy implications, researchers have the responsibility to communicate the results to the concerned policy-makers. Merely publishing the study or sending a copy of the report of the study is not enough. It is much better, where possible, to have a face-to-face presentation with ample time for discussion. Grant-making bodies usually approve an allocation in the budget for dissemination of the research results. This may include, where appropriate, a meeting with health managers and policy-makers to inform and discuss the results with them.

For research with policy implications, communication should not be left until the completion of the research. It should ideally start during the stage of planning the research to ensure that research questions are framed appropriately and tested in relevant contexts using interventions that can be replicated in practice Where possible, those who are most likely to use the results of research should also be involved in the implementation of the research project.

The following are some guidelines for investigators when making a presentation of their results to policy-makers.

- Know your audience and tailor the presentation to the particular audience. The audience may be physicians only, or may include nurses, community leaders, and donor agencies. If necessary, more than one presentation should be made.
- Avoid technical jargon. It will not impress. It will simply confuse and distract.
- Do not overload the presentation with statistical data. Include only the data that justify and explain the conclusions and recommendations.
- Follow the same steps as in a scientific presentation, with emphasis on the conclusions and recommendations. Recommendations are more likely to be implemented if they are directed to those who should and can implement them and if attention is given to the feasibility of their implementation. Specific recommendations are better than general recommendations. A plan of action is even better. Policy-makers often prefer to be given options about what can be done, with an outline of the advantages and disadvantages of each option. They prefer not to be told what to do, but to be given the information upon which they can make appropriate decisions.
- Visual aids, properly selected and designed, are useful in highlighting the important points in the presentation, including the main conclusions and recommendations.
- Allow adequate time for discussion.
- Be prepared to accept comments, criticisms and suggestions. But be also prepared to defend your results.
- Have an informative executive summary of the study, and make it available for distribution in the meeting. A copy of the visual aids may be given at the end or at the beginning of the presentation. It is always better that participants in the meeting take something with them. The full report is less likely to be read than an executive summary.
- Have a record of the meeting, and a note of any agreements made. This should be prepared shortly after the meeting, before the meeting is forgotten. The record or minutes should be circulated to those who attended the meeting and also to those who were expected to attend but could not.

10.6 Communicating to patients

The health research community has an obligation to ensure that patients have access to appropriate scientific information. There has been an explosion of health information on the internet. More than 100 000 medical websites (of varying quality) exist, and their number is growing rapidly (Kiley and Graham, 2002). The internet revolution in health care is largely driven by a massive consumer demand for online health resources. There is a growing body of health information, directed at patients, which is both scientifically sound and intelligible. In 1998, the US National Library of Medicine (http://www.nlm.nih.gov) launched a consumer health page called MEDLINEplus, designed to direct consumers to resources containing information that will assist in researching their health questions. The pages are designed for educational use only and are not intended to replace advice from a health professional. The pages provide a carefully selected list of resources, not a comprehensive catalogue.

Care should be taken in communicating research findings to patients. People need to be empowered with scientifically valid and intelligible information. Information, particularly on health risks or benefits of different interventions can be confusing if not adequately presented and explained. Confusion can lead to patients making wrong decisions. For example, women aged over 50 years old may be told that mammography screening reduces their risk of dying from breast cancer by 25%. Few patients would understand that this impressive figure means an absolute risk reduction of only one in 1000: of 1000 women who do not undergo mammography, about four will die from breast cancer within ten years, whereas out of 1000 women who do three will die (Gigerenzer and Edwards, 2003).

10.7 Communicating to the community

If the research was a community-based study, the community has a right to know the outcome of the study. It is the duty and responsibility of the investigators to do this and to select the appropriate form and way of doing it. It is advisable to share the information with the community before putting it into the public domain. It would also be useful to check whether they agree with the findings and conclusions, and whether there are additional questions that needed to be addressed. The feasibility of any actionable recommendations can also be discussed.

10.8 Communicating to the public

The public is entitled to accurate scientific information on issues that can influence individual behaviour or public policy. Communicating scientific health information to

the public can be done by popular scientific publications targeted at a lay audience, by using the channels of public media, and, increasingly now, by using the internet.

Scientists need to engage the public in what they are doing. This involves more than just making scientific information freely available. The role of scientists is no longer to preach enlightenment to the ignorant masses. On certain issues, the role of scientists is to present the case objectively to an enlightened citizen jury to allow them to make an informed judgement. Scientists must accept that they are no more qualified than the general public to make value judgements as to the uses to which science shall be put. The uncertainty, inherent in the scientific process, must be adequately exposed. The arrogance of science must give way to a scientific culture of social responsibility.

Scientists, in arguing the case for academic freedom, state that scientific progress should not be stopped because of the possibility of abuse. The public, however, has a right to be concerned. One should never underestimate the ability of human beings for irrational behaviour. The atomic bomb and other scientifically developed weapons of mass destruction are still in the memory of the public. The objective of science is to work for a better world. Science should not be used for purposes intended to harm human beings or the environment. Scientist should consider the ethical implications of their work.

The task of educating the public is achievable, as evidenced by a recent referendum in Switzerland on genetic engineering. Voters were divided on the issue. After Switzerland's scientists opened their laboratories and communicated with the public, the result was a two-third majority against a total ban on genetic modification of plants and animals, and their release into the environment.

10.9 Communicating to the public media

Scientists should be careful in communicating scientific data to the public media. The media, in its presentation of science, aims first to engage and entertain, and only second to inform. Scientists should resist the temptation to communicate just for the sake of publicity. Scientists should help the public media to prepare and present accurate reports of scientific data of interest to the public.

There are ethical considerations in communicating the results of scientific research to the public media. Media reports of scientific research before the work has been peer-reviewed and fully published may lead to the dissemination of inaccurate or premature conclusions. Very little medical research has such urgent implications for public health that it should be released before full publication in a scientific journal. In such a situation, the decision should be made, not by the researchers, but by the appropriate public health authority. Improperly communicated scientific information can result in unjustified public alarm. Researchers who present their work at a scientific meeting may discuss

their presentations with media reporters, but they should not go beyond what they have presented.

References and additional sources of information

Reading the medical literature. Applying evidence to practice. Washington, DC, American College of Obstetricians and Gynaecologists, 1998.

Byrne DW. *Publishing your medical research paper*. Baltimore, Williams & Wilkins, 1998: 210–217.

Delamothe T, Smith R. PubMed Central: creating an Aladdin's cave of ideas (editorial). *British Medical Journal*, 2001, 322:1–2.

Delamothe T, Godlee F, Smith R. Scientific literature's open sesame? *British Medical Journal*, 2003, 326: 945–946.

Davidoff F, Haynes B, Sackett D, Smith R.. Evidence based medicine (editorial). *British Medical Journal*, 1995, 310: 1085–1086.

Eysenbach G, Sa ER, Diepgen TL. Shopping around the Internet today and tomorrow: towards the millenium of cybermedicine. *British Medical Journal*, 1999, 319: 1294.

Gigerenzen G, Edward A. Simple tools for understanding risks: from innumeracy to insight. *British Medical Journal*, 2003, 327: 741–744.

Grimes DA. Communicating research: working with the media. In: O'Brien PMS, Pipkin FB, eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 210–217.

Kiley R, Graham E. *The patient's Internet handbook*. London, Royal Society of Medicine Press Ltd., 2002.

Lock S. Foreword. In: Hawkins C, Sorgi M, eds. *Research: How to plan, speak and write about it*. Berlin, Springer-Verlag, 1985: vii.

Long M. The future of electronic publishing. In: Hall GM, ed. *How to write a paper*, 2nd edition. London, BMJ Books, 1998: 132–138.

Shiffman RN et al. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Annals of Internal Medicine*, 2003, 139: 493–498.

Smith R. Closing the digital divide (editorial). *British Medical Journal*, 2003, 326: 328.

Smith R. Do patients need to read research? *British Medical Journal*, 2003, 326: 1304.

Tamber PS, Fiona F, Newmark P. Open access to peer-reviewed research: making it happen. *Lancet*, 2003, 362: 1575–77.

Chapter 11

Writing a scientific paper

11.1 Introduction

Writing a scientific paper is the most common way of communicating the results of research to other scientists and to health professionals. It goes without saying that authors should at all times have in mind objectivity, clarity and honesty in reporting their research. The format for writing a scientific paper for publication in biomedical journals has been standardized to provide a systematic and organized way to present the data. The text of observational and experimental articles is usually (but not necessarily) divided into sections with the headings: Introduction, Methods, Results, and Discussion. Long articles may need subheadings in some sections (especially the Results and Discussion sections) to clarify their contents. Journals generally provide in each issue, and on their web sites, detailed instructions to the authors on the required format for submitting papers.

The process of writing up the research should begin during the research planning, and continue while the research is being implemented. When the results of the research are analysed, a first draft of the written paper can be produced. Revision of this draft is an important part of the process. It should include revision for the content and revision of the style.

Not all scientific communications fit into the classical format for presentation of research. Two such examples are a case report and a scientific review. There are special considerations in writing a paper describing the results of qualitative research, and also in writing a thesis or dissertation.

The International Committee of Medical Journal Editors issues a set of uniform requirements for submitting manuscripts to biomedical journals. These requirements are revised periodically, the latest version being dated November 2003. The requirements were taken into consideration in developing the guidelines in this chapter and also in Chapter 12 on publishing a scientific paper. For more details, the reader may consult the references and additional sources provided for the chapter.

11.2 Selecting a title for the paper

A good title should adequately describe the contents of the paper in the fewest possible words. It should not be too long or too short; generally, it should consist of 10–12 words. Some journals, but not all, allow sub-titles. The title should not include any unnecessary words, nor waste space with phrases such as “Observations on” or “A study of”. It should not contain abbreviations.

Many journals require a running title (short title) to be printed at the top or bottom of every page of the article when it is published. Usually, this is between 30 and 50 characters.

11.3 Writing the abstract and key words

An abstract should be included at the beginning of the paper. The abstract can persuade or put off readers. The abstract is the part of the paper that will be included in most electronic databases, available for retrieval. The abstract should state the purposes of the study or investigation, basic procedures (selection of study subjects or laboratory animals; observational and analytical methods), main findings (giving specific data and their statistical significance, if possible) and the principal conclusions. It should emphasize the new and important aspects of the study or observations.

A good abstract should be a miniature version of the paper, provide a brief summary of each of the main sections of the paper and follow the structure of the paper. Many journals require a structured abstract, which includes subtitles such as objective, type of design, setting, material or subjects, methods, results, and conclusions. The number of words in an abstract should generally be less than 150 for unstructured abstracts, and less than 250 for structured abstracts. Some electronic databases are programmed to accept only up to this limited number of words. Abstracts are generally written in the past tense. The abstract should be self-contained and able to stand alone without need to consult the full text. As such it should not include references to literature or to figures and tables in the body of paper, should not include information that is not in the paper, and should not contain abbreviations or acronyms unless standard or very well known.

Most scientific journals require authors to provide 3 to 10 key words or short phrases that will assist indexers in cross-indexing the article. Key words are usually placed beneath the abstract. Terms from the Medical Subject Headings (MeSH) list of PubMed (US National Library of Medicine) should be used wherever possible, to facilitate indexing and retrieval (see Annex 3).

11.4 Article structure

A scientific article generally consists of four sections, with the acronym IMRAD: Introduction, Methods, Results, and Discussion. These sections are described by the following questions, called the Bradford Hill questions, after the author (Hill, 1965):

Introduction:	Why did the authors start?
Methods:	What did they do?
Results:	What did they find?
Discussion:	What do the results mean?

Reasoning in the paper should follow a straight line. The flow should not stray from the objective or research question. It cannot be written in the style of a story or novel, where the author can move between the characters and can jump between different time episodes.

11.5 Writing the Introduction

The introduction should:

- tell the reader why the research was started, and make clear what question the research was designed to answer. Research is not a fishing expedition. It is designed with a specific question in mind.
- raise the interest of the reader. The first few lines in the paper may attract or put off the reader. Investigators are advised to convey their enthusiasm but not to exaggerate.

The introduction should not:

- explain what can be found in any textbook in the field
- be over-referenced; it should give only strictly pertinent references
- include data or conclusions from the work being reported.

11.6 Writing the Methods section

Principles

Replicability of results is the heart of science. The methods section should provide a detailed exposition of the research design. A reader of the methods section should be able to repeat the study and to validate the findings. A methods section less than two double-spaced pages is probably inadequate.

The methods section should be organized under meaningful subheadings and describe techniques used in sufficient detail to allow others to replicate the study. Established methods should be referenced but no description is necessary. For published but not well known methods, a reference as well as a brief description should be given. New or substantially modified methods should be clearly described, with reasons given for using them and with their limitations outlined.

The methods section should not:

- refer to patients and animals as material; patients and animals are living things; not inanimate "material". The term "material" should be used only if inanimate specimens have been used.
- use proprietary names of drugs; generic names should be used.

Ethics

When reporting experiments on human subjects, authors should indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration.

Patients' names, initials, or hospital numbers should not be used. Particular care should be taken that these do not appear in illustrative material.

When reporting experiments on animals, authors should indicate whether the institutional or national guidelines or laws on the care and use of laboratory animals were followed.

Statistics

Statistical methods should be described in sufficient detail to enable a knowledgeable reader with access to the original data to verify the reported results. References for statistical methods should be to standard works when possible. Any computer programs used should be identified. Statistical terms, abbreviations, and symbols should be defined.

Details about randomization, if used, should be given, as well as concealment of allocation to treatment groups, and the method of masking (blinding). Losses to observation (such as dropouts from a clinical trial) should be reported.

It is recommended to include the word "considered" in descriptions of statistical significance such as "a P value of less than 0.05 was considered statistically significant", since the choice of this cut-off point is arbitrary.

It is better to avoid non-technical uses of technical statistical terms, such as “random” “significant”, “correlation” and “sample” in non-statistical contexts.

11.7 Writing the Results

Principles

The objective of the research should be kept in mind. Results that do not relate to the research objective should not be mentioned. Sufficient detail should be given to allow other scientists to assess the validity and accuracy of the results. Statistics should not take over the paper, but statistical analysis of the results should be adequately described. Results should be presented in a logical sequence in the text, tables, and illustrations. Tables and graphs are often extremely helpful in summarizing large amounts of data. Authors should not repeat in the text the numerical data contained in figures and tables.

The number of tables and figures should be restricted to those needed to explain the argument of the paper and to support its findings. A good rule about whether to include figures or not is: When in doubt, leave it out.

Tables

Tables should be used to show the exact values of more data than can be summarized in a few sentences of text; or when the objective of presenting data is to present specific inter-relationships. Tables should not be used when the data can be easily presented in the text (tables are more expensive to typeset than text); or when there is no relation between the data or to a time sequence.

A table should be readily understood without reference to the text. After reading the title and abstract, many readers often glance through the tables and illustrations before deciding whether or not to read the text. A table should be cited in the text, be numbered, and have a title which exactly describes the content of the table. It should have short or abbreviated headings for columns and rows and, if necessary, a footnote for explanation of non-standard abbreviations that are used, and for identification of statistical measures of variations, such as standard deviation and standard error of the mean. Tables should have a logical structure. Columns should be arranged from left to right in a logical sequence, e.g. to reflect the sequence in which data were collected or changes over time. Rows should be arranged from top to bottom in a logical order, e.g. by ascending order of age.

A table should not include in its title any unnecessary words, nor a repetition of column and row headings. There should be no ambiguity about the purpose of the

columns and rows. When column headings are grouped, a straddle-line should be used to eliminate any uncertainty about which column headings are included under the grouped column headings. Items in row headings may be indented to indicate groupings.

For purposes of publication:

- The table should not exceed the width of the journal columns. A single-column table, in a journal with a double-column page, should not include more than 60 characters (and equivalent spaces) in a row (with its row heading). A table running the full width of a page should not include more than 120 characters in a row.
- Each table should be typed or printed with double-spacing on a separate sheet of paper. Tables should not be submitted as photographs or images.
- Tables should not have internal horizontal and vertical rules.
- Tables should be numbered consecutively in the order of their first citation in the text. Each table should be cited in the text.
- If data are used from another published or unpublished source, permission is needed and should be acknowledge fully.
- The use of too many tables in relation to the length of the text may produce difficulties in the layout of pages. Issues of the journal to which the paper will be submitted can be checked to estimate how many tables can be used per 1000 words of text. A general rule is no more than one table (or illustration) per 1000 words of text (4 pages of manuscript).
- The editor, on accepting a paper, may recommend that additional tables containing important backup data, too expensive to publish, be deposited with an archival service, such as the National Auxiliary Publication Service in the United States, or made available by the author on request. In that event an appropriate statement will be added to the text. Such tables should be submitted for consideration with the paper.

Illustrations

Illustrations should be used only for a specific purpose. An illustration may be used as evidence to support the argument, since “seeing is believing”. Illustrations may be used as a more efficient way in presenting data. A flow chart is such an example. The use of illustrations for emphasis, just to stress a point, is not a good purpose. It may be more appropriate for a presentation than a written paper.

Graphs are used to illustrate relationships. If exact values are important, a table is preferable to a graph; when trends and relationships are more important than exact

values, a graph is more efficient. A graph is a better alternative than a table with many entries. The same data should not be repeated in figures and tables.

For purposes of publication:

- Figures should be professionally drawn and photographed; freehand or typewritten lettering is unacceptable.
- Instead of original drawings, X-ray films, and other material, authors should submit sharp, glossy, black-and-white photographic prints, usually 127 × 173mm (5 × 7 inches) but not larger than 203 × 254mm (8 × 10 inches). Letters, numbers, and symbols should be clear and even throughout, and of sufficient size that when reduced for publication each item will still be legible.
- Titles and detailed explanations belong in the legends for illustrations not on the illustrations themselves.
- Each figure should have a label pasted on its back, indicating the number of the figure, author's name, and top of the figure. Do not write on the back of the figures or scratch or mar them by using paper clips. Do not bend figures or mount them on cardboard.
- Photomicrographs should have internal scale markers. Symbols, arrows or letters used in microphotographs should contrast with the background.
- If photographs of people are used, either the subjects must not be identifiable or their pictures must be accompanied by written permission to use the photographs.
- Figures should be numbered consecutively according to the order in which they have been first cited in the text.
- If a figure has been published, the original source has to be acknowledged and a written permission from the copyright holder to reproduce the material should be submitted.
- Permission is required irrespective of authorship or publisher except for documents in the public domain.
- For illustrations in colour, it is important to ascertain whether the journal requires colour negatives, positive transparencies or colour prints. Some journals publish illustrations in colour only if the author pays for the extra cost.
- Legends for illustrations should be typed or printed using double spacing, starting on a separate page, with Arabic numerals corresponding to the illustrations.
- When symbols, arrows, numbers, or letters are used to identify parts of the illustrations, each one should be explained clearly in the legend. The internal scale, and the method of staining in microphotographs, should be stated.

11.8 Writing the Discussion and Conclusions

This section of the paper should emphasize the new and important aspects of the study and the conclusions that follow from them. It should not repeat in detail data or other material given in the Introduction or Results sections.

Good papers have a targeted discussion, to keep it focused. The discussion should preferably be structured to include the following six components (Docherty and Smith, 1999):

- statement of principal findings
- strengths and weaknesses of the study
- strengths and weaknesses in relation to other studies
- meaning of the study, possible mechanisms and implications for clinicians and policymakers
- unanswered questions and future research
- conclusion.

- Statement of principal findings: The opening of the discussion usually gives the answer to the research question, or a restatement of the principal findings. This should not normally be more than a few sentences. It is advisable that the discussion start with a sentence that clearly shows that the paper includes new information. Reviewers often start with a "null hypothesis" that the paper does not add anything new.
- Strengths and weaknesses of the study: Equal emphasis should be given to both strengths and weaknesses. Reviewers are more interested in seeing that the author is aware of the weaknesses. If the reader discovers in the paper weaknesses that are not mentioned by the author, the trust in the paper will be shaken. A subheading such as "limitations of the study" or data is useful. Findings that have not been described in the results section should not be discussed.
- Strengths and weaknesses in relation to other studies: All evidence bearing on the argument, with or against, should be considered. Authors should discuss the opposing point of view, taking a "devil's advocate" position. Full credit should be given for supporting evidence. Authors should avoid burying the citation of a previously published paper on the same question, which arrived at the same answer in the discussion. Such a citation is better highlighted in the introduction. It is not enough to simply summarize published papers. The authors should critically evaluate their methodology, findings and conclusions. In particular, any differences in results should be discussed and possible explanations offered. If the authors do not know why their results are different from other studies, they should say so, but not imply that their results are better.

- Meaning of the study, possible mechanisms and implications for clinicians and policymakers: This section should be written carefully. Authors should not move beyond the limited evidence provided by the study. Restraint in stating implications is a virtue appreciated by reviewers and readers. It may also be relevant to emphasize, not only what the results mean, but also what the results do not mean. This will keep readers from making unjustified conclusions.
- Unanswered questions and future research: New research may be proposed to provide the answer to questions that are still not answered. A good study should generate new ideas for further research. A simple statement that further research is needed is less helpful than providing new specific research questions or suggesting particular studies.
- Conclusion: A good paper ends with strong clear conclusions. It has been said that the body of a good paper is a "thunderbolt in reverse": it begins with thunder (introduction) and ends with lightning (conclusions) (Byne, 1998). Conclusions should be linked with the goals of the study, and should be limited to the boundaries of the study. Authors should avoid unqualified statements and conclusions not completely supported by the data. For example, they should not make statements on economic benefits and costs unless their manuscript includes economic data and analysis. Authors should refrain from claiming unjustified priority about the findings. It should be noted that a negative finding could be as important as a positive finding.

11.9 Acknowledgements

At an appropriate place in the article (the title page, footnote or an appendix to the text; depending on the journal requirements), one or more statements should specify: contributions that need acknowledging but do not justify authorship, such as general support by a department chair; acknowledgement of technical help; acknowledgements of financial or material support, which should specify the nature of the support; and relationships that may pose a conflict of interest.

Persons who have contributed intellectually to the paper but whose contributions do not justify authorship may be named and their function or contribution described, for example "scientific adviser", "critical review of study proposal", "data collection", or "participation in clinical trial". Such persons must have given their permission to be named. Authors are responsible for obtaining written permission from persons acknowledged by name, because readers may infer their endorsement of the data and conclusions. Technical help is better acknowledged in a paragraph separate from that acknowledging other contributions.

11.10 Citation of references

The reference section is an important part of a scientific paper. The number of references should be restricted to those that have a direct bearing on the work described. Except for review articles, it is rarely necessary to have more than 40 references in the longest paper (Halsey, 1998).

References should be carefully checked. They should be verified against original documents. One study has shown that in a random check of references in published papers, 20% were misquoted, with half of the misquotations being seriously misleading (DeLacey et al. 1985). Useful advice for the author is to photocopy the first page of every reference cited. This page normally includes all the information needed for correctly citing the reference.

Different standard formats for citing references are used in different scientific disciplines. These formats include: MLA Style established by the Modern Language Association; APA Style, governed by the Publication Manual of the American Psychological Association; CMS Footnote Style, conforming to the Chicago Manual of Style; and CBE Number Style established by the Council of Biology Editors.

In biomedical sciences, there are two major styles for citing the references: the Harvard system and the Vancouver system.

In the Harvard system, the order of references at the end of the paper is strictly alphabetical, regardless of the chronology. In the text of the paper, references are cited by giving in parentheses the name of the author and the year of publication. When the author's name is part of a sentence, only the year is put in parentheses. When several references are given together, they should be listed in chronological order and separated by a semicolon. When a paper written by two authors is quoted, both names are given. If there are more than two authors, all the names may be given the first time the reference is cited. Otherwise, it is sufficient to give the name of the first author only, adding "et al". The term "et al" means "and others". It is an abbreviation for two Latin terms: "et alii" (masculine) and "et aliae" (feminine). When two citations have the same author and the same year of publication, alphabetical annotation is used, for example "2004a". The order of these alphabetically annotated citations ideally should be chronological within the year.

The Vancouver system has been adopted in the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" by the International Committee of Medical Journal Editors (who held their first meeting in Vancouver). Most biomedical journals follow this system. It is based largely on a standard style adapted by the US National Library of Medicine (NLM) for its databases. According to the Vancouver style, references should be numbered consecutively in the order in which they are first mentioned in the text. References in text, tables and legends should be identified by

Arabic numerals in parentheses. References cited only in tables or figure legends should be numbered in accordance with the sequence established by the first identification in the text of the particular table or figure.

In writing the early drafts of the paper, it is advisable to use the Harvard style. If numbers are assigned to references at this early stage, those numbers will very likely have to be changed in subsequent drafts. In the final draft, the authors can switch to the Vancouver style. To track the references in the early drafts using a word-processing program, one can place at the beginning of each citation a character not used elsewhere in the text, for example an asterisk (*).

If journal titles are abbreviated, as is the practice in most but not all journals, this should be in line with the abbreviations in the Index Medicus (which are based on an international standard). The list of journals is published annually in the January issue. The list can also be accessed through the web site of the US Library of Medicine (http://www.nlm.nih.gov).

Unpublished observations are generally not to be used as references; papers accepted for publication but not yet published and given as references are identified as "in press" or "forthcoming"; research papers submitted to a journal but not yet accepted are to be treated as unpublished observations.

Authors should avoid citing a "personal communication" unless it provides essential information not available from a public source, in which case the name of the person and date of communication should be cited in parentheses in the text. Authors should obtain permission and confirmation of accuracy from the source of a personal communication.

Annex 4 provides examples on how different types of references should be cited. Additional information may be obtained from the web site: http://www.nlm.gov/bsd/uniform_requirements.html.

11.11 Steps in the process of writing a paper

The process of writing a scientific paper should start before doing the research, continue during the research, and be completed after the research results have been described, analysed and interpreted. After writing the paper, it should be carefully revised, first for content and then for style.

Before the research

- Search the literature and keep a record of the references.
- Prepare dummy tables for results.

During the research

- Record the results.
- Update the literature.

After completion of the research

- Use a systematic approach, building the paper step by step. Do not try to do the whole thing at once.
- Start with an outline, which will serve as framework.
- The discussion is the part that requires most careful thought and interpretation.
- Begin with the easiest section. Deal with individual sections one at a time.
- Decide on the journal to which the article will be submitted and study its format requirements.
- Write the rough draft: Once you start, write as fast as you can. Do not worry about style.
- Put the paper aside for several days or weeks and then re-read it.
- Give a version of the paper to a colleague or colleagues to review it.
- Date all drafts.

11.12 Revision of the manuscript for scientific content

For creative writing, the word processor is the best invention since the quill pen. The days of retyping are over. Most journals require an electronic copy of the paper.

Revision checklist

- Is the title accurate, succinct and effective?
- Are keywords indexable? It is better to use keywords from the Medical Subject Headings (MeSH vocabulary) of MEDLINE (Annex 3).
- Does the abstract represent the content of all the main sections of the paper, within the length allowed by the journal? Do data in the abstract agree with data in the paper?
- Does the introduction set the stage adequately but concisely for the main question considered, or for the hypothesis tested, in the paper? Is that question or hypothesis made clear by the end of the introduction?
- Are the methods described in enough detail to allow replication of research? Are statistical methods described?

- Are the results presented in a way that allows other investigators to check and to compare? Can any of the tables or illustrations be omitted? Can any of the tables be replaced by a graph? Do data in the text agree with data in the tables? Are all tables and figures cited in the text? Are all tables and figures mentioned in the text included? Are legends of figures correct?
- Does the discussion properly interpret the significance of the data? Does the discussion reflect up-to-date awareness of the literature? Are conclusions justified by the results?
- Are all references cited mentioned in the text? Are all references mentioned in the text cited? Have any necessary references been omitted?
- Is the length of the paper appropriate? Does any of the text repeat information found elsewhere in the paper? Are there paragraphs or sentences that can be omitted? Where possible, it is good to plan to submit an article that is shorter than the average article published in the journal to which the paper will be sent. The best papers are concise. Generally, a manuscript should, on average, be about 10 double-spaced pages, or 3 published pages, with 25 references. (Each printed page is about 3–4 double-spaced typed pages). The sections of a manuscript that are often too long are the introduction and discussion. The sections that are often too short are the methods and results. A good rule is to shorten the introduction and discussion and to expand the methods and results sections.
- Are all pages numbered?

11.13 Revision of the manuscript for style

The acronym "KISS", "keep it simple and short", is the key to good scientific writing. Authors should always choose the simplest and shortest way of saying something. It takes more time to write a good concise paper, than a lengthy one. Pascal once wrote to a friend: "I am sorry this letter is so long but I had no time to write a short one." Most authors do not spend enough time planning. Good planning will shorten the time spent in writing.

In editing oneself, consideration should be given to paragraphs, sentences and words. The following sections provide a few useful hints, particularly for non-English speakers. For additional information, sources such as Strunk (2000) can be consulted.

Paragraphs

Well structured paragraphs are the key to good writing, and should consist of: a topic or lead sentence to introduce the subject of the paragraph; body sentences which expand upon the theme and present a logical argument; and either a transitional sentence,

which leads into the next paragraph, or a concluding sentence. There is no firm rule on paragraph length: more than 25 typed lines would be too long; fewer than 5 or 6 lines represent what is really a fragment of either adjacent paragraph. A new paragraph must either link to that preceding it and/or following it, or should clearly introduce a new subject. In a long discussion, subheadings are a good idea.

Sentences

The following hints may be helpful to authors in revising the style of their paper.

- Long sentences (more than two typewritten lines) are better avoided if possible.
- The active is preferable to the passive because it is much clearer and easier to understand, in general. For example, replace "It was found by *x*" by "*x* found that"). The passive voice has traditionally been used in scientific writing to refer to the thoughts or actions of the author. This tendency is slowly changing, and many editors now encourage authors to use "I" or "we" in their writing.
- Avoid ambiguity in the use of adjectival and adverbial clauses and phrases. It is often better to simplify sentences by splitting the subordinate phrases and clauses and making them sentences on their own.
- Avoid verbosity (to say a thing in a complicated way, to make it sound important) or pompous verbiage.
- Each sentence must have a verb, and the verb should agree with the noun.
- Economy is a virtue. Strike out unneeded words and phrases.
- "Do not use a preposition to end a sentence with"—is a good rule which itself breaks the rule.
- It is a useful convention to put anything that was done in the past tense and to put general statements in the present tense. In general, the introduction and discussion sections are written in the present tense, and the methods and results sections are written in the simple past tense.

Words

It is advisable to look for and try, where possible, to replace the following six groups of words.

- Abstract nouns (nouns formed from verbs and ending in: tion, sion, ance, ment, ness, cy). These nouns are better replaced with verbs. For example, change "Measurements were performed on the variation" to "The variation was measured" or "we measured

the variation"; change "The interpretation of the data was made" to "Data were interpreted" or "we interpreted the data".

- Compound nouns (noun clusters) e.g. patient liver enzyme status (the status of liver enzymes in patients); research result dissemination methods (methods of disseminating research results).
- Abbreviations, unless they are standard and unless they are used at least ten times in the paper. Avoid abbreviations in the title and abstract. The complete term for which an abbreviation stands should precede its first use unless it is a standard unit of measurement.
- Sexist words: Do not use the pronoun "he" or "his" when she or her would be equally appropriate. Use the plural form instead. Try to replace words such as: man (unless referring to a man), mankind, manpower, policeman, foreman.
- Dehumanizing words: e.g. referring to people as cases or subjects (use patients or volunteers for example); using syndromic tags for patients; male/female are more appropriate for animals; men and women are better for human subjects.
- Slang and jargon (words that have an arbitrary meaning).

Do not confuse American and British Spelling. Follow the style prescribed by the journal. If in doubt, use a good dictionary (do not depend on the spell-checker in the computer which is only as good as its content).

Unless otherwise requested in the journal instructions to authors:

- Measurements of length, height, weight, and volume should be reported in metric units (metre, kilogram, or litre) or their decimal multiples, and temperatures should be given in degrees Celsius. Blood pressure should be given in millimetres of mercury.
- All haematological and clinical chemistry measurements should be reported in the metric system in terms of the International System of Units (SI). Editors may request that alternative or non-SI units be added by the authors before publication.

11.14 Writing a case report

Reports of single cases have become less and less acceptable for publication in major journals, mainly because of their tendency to carry relatively little important new information. The following kinds of case reports still merit publication:

- The unique or nearly unique case that appears to represent a previously undescribed syndrome or disease.

- The case with an unexpected association of two or more diseases or disorders that may represent a previously unsuspected causal relation.
- The case representing a new and important variation from an expected pattern: the "outlier" case.
- The case with an unexpected evolution that suggests a therapeutic or adverse drug effect.

A good example of an important case report is the report by Hymes et al. in 1981 of eight cases of the rare skin tumour, Kaposi's sarcoma in New York. Usually a slowly growing tumour, the course in these cases was aggressive. Usually a disease of old people, these cases occurred in young men. The patients were all homosexual men. This report first alerted the world to the AIDS epidemic.

11.15 Writing a secondary scientific paper

A secondary scientific paper is a review paper which summarizes other papers. There are two types of reviews: a narrative review and a systematic review. The distinction between the two types of review should be clear. Meta-analysis is a special type of systematic review.

Narrative review

In the narrative review, the studies reviewed have not been identified or analysed in a systematic, standardized and objective way. Experts, to provide an update on a certain subject, usually write the review.

Systematic review

The systematic review contains an explicit statement on objectives with a spelt out research question. The data sources for the papers (including grey literature) are stated as well as the method of selection. The review is conducted according to an explicit and reproducible methodology. Different from the narrative review generally written by experts, a systematic review may be better done by non-experts on the subject, who are experts on writing systematic reviews.

A systematic review generally includes the following parts:

- Abstract
- Introduction: A well-conceived systematic review answers a question or closely related questions, which should be made clear at the beginning of the review.

- Methods: The methods section in a systematic review should fully describe the methods used for locating, selecting, extracting and synthesizing the data. It should outline the literature search, including the bibliographic indexes and databases searched, limits on years and languages, as well as search terms used.
- Body of the review: Topics in the body of the review depend on subject. The sequence should have a logical basis. Sequence should be made clear by subheadings. The argument should be critical.

Assessment of the quality of systematic reviews is discussed in Chapter 14.

Meta-analysis/pooling

Meta-analyses critically review research studies and statistically combine their data to help answer questions that are beyond the power of single papers. "Power" is the term to describe the value of this technique. Combining data from a number of studies increases the sample size. The technique of meta-analysis has great potential for synthesizing research results and adding precision and power to our estimates of effect.

The results of these meta-analyses now tend to be presented in a standard format, because they mostly use a common computer software known as MetaView to do the calculation and express the results in a graphic form. This format is colloquially known as a "forest plot" or "blobbogram". It shows a number of horizontal lines, each representing one study. The blob in the middle of each line is the point estimate, and the width of the line represents the 95% confidence interval of this estimate. A vertical line represents "line of no effect". If the horizontal line of any trial does not cross the line of no effect, there is a 95% chance that there is a "real" difference between the groups (Greenhalgh, 1997).

A typical example of the value of meta-analysis studies is the meta-analysis of seven trials of the effect of giving steroids to mothers who were expected to give birth prematurely. Only two of the seven trials showed a statistically significant benefit. But when the results of the seven studies were pooled together, the strength of the evidence in favour of the intervention was demonstrated. The meta-analysis showed that infants of mothers given corticosteroids were 30% to 50% less likely to die. The Cochrane Collaboration adopted this example as its logo (Greenhalgh, 1997).

Assessment of the quality of meta-analysis is discussed in Chapter 14.

11.16 Writing a paper on qualitative research

Since the 1990s, qualitative methods of research have been increasingly used in health research. This has led to a corresponding rise in the reporting of qualitative research studies in medical and related journals. The following are examples of papers on qualitative research studies recently published in the *British Medical Journal*:

- Patients' views about taking anti-hypertensive drugs
- Young women' s accounts of factors influencing their use and non-use of emergency contraception: in-depth interview study
- Patients' unvoiced agendas in general practice consultations: qualitative study
- A qualitative study of evidence-based leaflets in maternity care
- A qualitative study of barriers to uptake of services for coronary heart disease
- Why do general practitioners prescribe antibiotics for sore throat? Grounded theory interview study.
- Doctor's perceptions of palliative care for heart failure: focus group study
- Knowledge and perceptions of general practitioners about impaired glucose tolerance
- Why general practitioners do not implement evidence: qualitative study
- Relation between private health insurance and high rates of Caesarean section: qualitative and quantitative study
- Qualitative analysis of psychosocial impact of diagnosis of *Chlamydia trachomatis*

Writing a paper based on qualitative research does not need to differ from the framework used for quantitative research: introduction, methods, results and discussion (Kirsti, 2001). Quotes from participants are often used in the Results section of papers on qualitative research. These should not repeat what is in the text. It is not necessary to include more than one quote to illustrate a point. In translating quotes to English, this should be done in appropriate style, reflecting the sense of the quote, and not just a literal translation. As a general rule, authors should use verbatim quotes, wherever possible, and keep them down to short segments of text.

11.17 The dissertation or thesis

Different from a scientific paper submitted for publication, a dissertation or thesis is written and submitted as a partial or complete requirement for an academic degree,

a master or a doctorate. The thesis is meant to: present and defend the results of a scientifically sound piece of research; display good knowledge of the field of study; show familiarity with the scientific method; and demonstrate the intellectual ability of the candidate. The simple acquisition of voluminous data is not enough. In most cases, this acquisition could have been done equally well by a technician.

The steps in the preparation of a thesis follow the same lines outlined in previous chapters on what research to do, planning of the research and selecting a research design, writing the research protocol, implementing the study, describing and analysing the results, and their proper interpretation. Writing the thesis also follows the same guidelines and format for writing a research paper. Although space is not a constraint, brevity is always a virtue. The following are some additional remarks for the different sections.

The introduction is generally expanded or replaced by a comprehensive review of the literature. This review is meant to display not only good and up-to-date knowledge of the field, but also the intellectual ability of the candidate. It should not include information already available in textbooks. It should include only information relevant to the work done. It should be analytical and critical. It should show the ability of the candidate to synthesize and put together information from different sources. It should properly recognize the work of previous researchers.

The objectives should be carefully stated. The thesis will be judged against how each objective was achieved.

The information in the methods section should be adequate to allow other researchers to replicate the study. Already established methods do not need to be described in any detail. Quality control of the measurements should be explained.

The results section should give equal emphasis to negative and positive findings, and should be presented in adequate detail to allow other investigators to replicate the findings.

Discussion should be limited to the results of the study. The limitations of the study should be brought up. Conclusions should not go beyond what the candidate did and found.

Acknowledgements should be generous and give credit to all who have helped the investigator.

It is not the number of references that matters but their relevance. They should include original articles and not be largely based on reviews. They should be up to date, indicating that the candidate was following the literature during and after the study. References from national sources or regional sources should be included together with those from the international literature. It is assumed that the candidate has read all the references. The references should be carefully checked against original documents.

The thesis should be checked for style. Spelling and grammar mistakes indicate sloppiness on the part of the candidate, and may lead the examiner to suspect sloppiness in the work itself. Word processors can help the candidate to recognize and correct these mistakes but contain hidden dangers and should not be relied on blindly.

In presenting the thesis, the same guidelines for scientific presentations outlined in the next chapter should be followed. Unlike a presentation to a scientific meeting, questions to the candidate will take more than the time of the presentation. The candidate has to explain his/her findings and display general knowledge in the field. Defending the work does not mean trying to cover up weaknesses in the study.

References and additional sources of information

Baker P.N. How to set about writing your first paper. In: O'Brien PMS, Pipkin FB, eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 225–230.

Byrne DW. *Publishing your medical research paper*. Baltimore, Lippincott Williams & Wilkins, 1998.

Crowley P. *Corticosteroids prior to pre-term delivery*, (updated January 1996). Cochrane Database of Systemic Reviews. London, BMJ Books, 1996.

DeLacey G, Record C, Wade J. How accurate are quotations and references in medical journals. *British Medical Journal*, 1985, 291: 884–886.

Docherty M, Smith R.. The case for structuring the discussion of scientific papers (editorial). *British Medical Journal*, 1999, 318:1224–1225.

Forgacs J. How to write a review. In: Hall GM, ed. *How to write a paper*, 2nd edition. London, BMJ Books, 1998: 77–82.

Greenhalgh T. *How to read a paper: the basics of evidence-based medicine*. London, BMJ Books, 1997:122; 119–123.

Hall GM, ed. *How to write a paper*, 2nd edition. London, BMJ Books, 1998.

Halsy MJ. Revising prose structure and style. In: Hall GM, ed. *How to write a paper*. 2nd edition. London, BMJ Books, 1998: 109–136.

Herod JJO. How to prepare a thesis. In: O'Brien PMS, Pipkin FB, eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 241–247.

Hill B. The reason for writing. *British Medical Journal*, 1965, 2:870.

Huth EJ. *How to write and publish papers in the medical sciences*. 2nd edition. Baltimore, Williams & Wilkins, 1990.

Hymes KB, Cheung T et al. Kaposi's sarcoma in homosexual men: A report of eight cases. *Lancet* 1981, 2: 598–600.

International Committee of Medical Journal Editors. *Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication*. Updated November 2003. (http://www.icmje.org accessed 24/2/2004)

Malterad K. Qualitative research: standards, challenges, and guidelines. *Lancet*, 2001, 358: 483–88.

Lester JD, Lester JD, Jr. *Writing research papers: a complete guide*, 10th edition. New York, Longman, 2002.

Pearce N. Style: What is it and does it matter? In: Hall GM, ed. *How to write a paper*, 2nd edition. London, BMJ Books, 1998: 116–121.

Pipkin FB. How a thesis or dissertation is assessed or examined. In: O'Brien PMS, Pipkin FB, eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 248–253.

Skelton JR, Edwards SJL. The function of the discussion section in academic medical writing. *British Medical Journal*, 2000, 320: 1269–1270.

Strunk W Jr. *The elements of style*. 4th edition. New York, Longman, 2000.

Wildsmith JAW. How to write a case report. In: Hall GM, ed. *How to write a paper*. 2nd edition. London, BMJ Books, 1998: 70–76.

Chapter 12

Publishing a scientific paper

12.1 Introduction

Publication of research work is essential in order to advance science and to improve health. It is also essential for people pursuing a scientific career. Their recognition as researchers depends on their publications and contributions to scientific progress. Scientists live in a culture of "publish or perish". Researchers should learn not only how to write a scientific paper, but also how to get it published. Scientific journals have technical requirements, and authors should make themselves familiar with these requirements. Researchers deserve to have the credit for their work, but only if they have contributed intellectually to it. Ethical standards apply to scientific publication and should be observed by authors, and ensured by editors.

12.2 How to get your paper published

The editor's decision to accept or reject a paper is generally based on the following:

- the message of the paper: how clear, important and new is the message?
- the relevance of the paper to the journal's scope and its audience; the journal's backlog of accepted papers is also a factor in the consideration;
- scientific validity of the evidence supporting the paper's conclusions;
- quality of the manuscript.

The message

The paper must have a message. A good message can be put in one sentence. Some journals now require this one sentence, beneath the title of the paper, in order to put it in the table of contents. A second issue is the "so-what" test: Do the findings have implications? Whether a journal accepts a paper often hinges on whether its message is new, expands on, confirms or rejects a previously published message.

Matching the topic and the journal

A decision on which journal to submit the paper to must be made before the paper is finally written. The paper must be written in conformity with the style of the journal. The list of journals indexed in PubMed/MEDLINE of the US National Library of Medicine includes over 2600 peer-reviewed journals grouped by subject field. A peer-reviewed journal is one that submits most of its published articles for review by experts who are not part of the editorial staff. It is important to ensure that the topic of the paper falls within the scope of the journal selected. The format of the paper should also be one that is accepted by the journal.

High prestige journals have high rejection rates, sometimes as high as 90%. Rejection does not necessarily mean that the paper is not good. Journals cannot publish all the good papers they receive. For the authors, rejection means loss of weeks or months before the paper is submitted again to another journal. Publication lag is the interval between acceptance and publication; the average lag is seven months. Even so, it is not acceptable to send the same paper simultaneously to more than one journal. The journal considers the paper on the assumption that it has not been submitted elsewhere. Among the principal considerations that have led to this policy are the potential for disagreement when two journals claim the right to publish the same manuscript, and the possibility that two or more journals will unknowingly and unnecessarily do the work of peer review and editing of the same manuscript, and even publish the same article.

A single paper is more likely to be accepted than one in a series. (Arbitrary carving up of clearly related aspects of one study is referred to as “salami science” and is not encouraged.)

Scientific validity

Internal validity refers to the degree to which the investigator’s conclusions correctly describe what actually happened in the study. It means that within the confines of the study, results appear to be accurate, the methods and analysis used stand up to scrutiny, and the interpretation of the investigators appears to be supported.

External validity (also called generalizability) refers to the degree to which the findings of the study may be generalized to the population from which the sample for the study was drawn. Poor methods and inadequate results are most often responsible for rejection.

Quality of the manuscript

This has been discussed in detail in Chapter 11.

12.3 Uniform requirements for manuscripts submitted to biomedical journals

A group of editors of general medical journals met informally in Vancouver, British Columbia, Canada, in 1978 in order to establish guidelines for the format of manuscripts submitted to their journals. The group became known as the Vancouver Group. Its requirements for manuscripts, including formats for citing bibliographic references, were first published in 1979. The Vancouver Group expanded and evolved into the International Committee of Medical Journal Editors (ICMJE), which meets annually, and has gradually broadened its concerns.

The Committee has produced several editions of the Uniform Requirements for Manuscripts Submitted to Biomedical journals: Writing and Editing for Biomedical Publication. Over the years, issues have arisen that go beyond manuscript preparation. Some of these issues have been covered in subsequent editions; others are addressed in separate statements. Each statement has been published in a scientific journal. In the latest revision (November 2003), the committee revised and re-organized the entire document and incorporated the separate statements in the text (http://www.icmje.org). The total content of the Uniform Requirements for Manuscripts Submitted to Biomedical Journals may be reproduced for educational, not for-profit purposes without regard for copyright. The Committee encourages distribution of the material.

The Uniform Requirements are instructions to authors on how to prepare manuscripts, not to editors on publication style. (But many journals have drawn on them for elements of their publication styles.) If authors prepare manuscripts in the style specified in these requirements, editors of the participating journals will not return the manuscripts for changes in style before considering them for publication. In the publishing process, however, a journal may alter accepted manuscripts to conform to details of its publication style. Authors sending manuscripts to a participating journal should not try to prepare them in accordance with the publication style of that journal but should follow the Uniform Requirements.

Authors must also follow the instructions to authors in the journal as to what topics are suitable for that journal and the types of papers that may be submitted, for example, original articles, reviews or case reports. In addition, the journal's instructions are likely to contain other requirements unique to that journal, such as the number of copies of a manuscript that are required, acceptable languages, length of articles, and approved abbreviations.

Participating journals (over 500 internationally) are expected to state in their instructions to authors that their requirements are in accordance with the Uniform Requirements for Manuscripts Submitted to Biomedical Journals and to cite a published version.

The following sections are largely based on these uniform requirements.

12.4 Summary of technical instructions for submission of papers

Type or print out the manuscript on white bond paper, 216 × 279 mm, or ISO A4 (212 × 297 mm), with margins of at least 25 mm. Type or print on only one side of the paper. Use double-spacing throughout, including for the title page, abstract, text, acknowledgements, references, individual tables, and legends. Number pages consecutively beginning with the title page. Put the page number in the upper or lower right-hand corner of each page. Begin each section or component on a new page

Place each table on a separate page. Illustrations and unmounted prints should be no larger than 203 × 254 mm. Authors should submit the required number of paper copies and are advised to keep copies of everything submitted.

The title page should carry:

- title of the article, which should be concise but informative;
- name by which each author is known, with his or her highest academic degree(s) and institutional affiliation;
- name of the department(s) and institution(s) to which the work should be attributed;
- disclaimers if any;
- name and address of the author responsible for correspondence about the manuscript; the name and address of the author to whom requests for reprints should be addressed, or a statement that reprints will not be available from the authors;
- source(s) of support in the form of grants, equipment, drugs, or all of these;
- short running head or footline of no more than 40 characters (count letters and spaces) at the foot of the title page.

An increasing number of journals require electronic submission of manuscripts, whether on disk, as attachment to electronic mail, or by downloading directly onto the journal website. Electronic submissions save time as well as postage costs, and allow the manuscript to be handled in electronic form throughout the editorial process, for example when it is sent out to reviewers. Authors can follow the course of their paper by accessing the website of the journal. Authors should consult the journal's instructions to authors for acceptable word processing formats, conventions for naming files, and other details.

When paper manuscripts are submitted, journals commonly require authors to provide a copy in electronic form (on a disk) when the papers are close to final acceptance. The disk should be clearly labelled with the format of the file and the file name.

12.5 Sending the manuscript to the journal

The required number of copies of the manuscript should be sent in a heavy-paper envelope, enclosing the copies and figures in cardboard, if necessary, to prevent the photographs from being bent. Photographs and transparencies are better put in a separate heavy-paper envelope.

Manuscripts must be accompanied by a covering letter signed by all co-authors. This must normally include:

- information on prior or duplicate publication or submission elsewhere of any part of the work;
- a statement of financial or other relationships that might lead to conflict of interest;
- a statement that the manuscript has been read and approved by all the authors, that the requirements for authorship have been met, and that each author believes that the manuscript represents honest work; and
- the name, address, and telephone number of the corresponding author, who is responsible for communicating with the other authors about revisions and final approval of the proofs.

The letter should give any additional information that may be helpful to the editor, such as the type of article in the particular journal that the manuscript represents and whether the author(s) would be willing to meet the cost of colour illustrations.

Copies of any permission to reproduce published material, to use illustrations or report information about identifiable people, or to name people for their contributions must accompany the manuscript.

A transfer of copyright may be required at this stage, or after the paper has been accepted for publication.

12.6 After submitting the manuscript

Acknowledgement of receipt of the manuscript is usually received within 2–3 weeks. A decision regarding publication is usually made within 6–8 weeks, depending on

reviewers' responses. Rejection rates of the best journals are over 50%. Probably only 5% or so of papers are accepted without change recommended as a result of peer review.

The reviewers' responses may suggest that that the paper can be made more acceptable by revisions. The investigators do not need to make all the changes suggested automatically. They should adopt revisions that will satisfy the reviewers' criticisms wherever possible and justify any decision not to do so. It is good to indicate in a separate page the criticism made and how the revised paper responds to them. This will facilitate a decision by the editor.

12.7 Authorship in scientific papers

An "author" is generally considered to be someone who has made substantial intellectual contribution to a published study. The International Committee of Medical Journal Editors issued the following guidelines about authorship.

- Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content.
- Authorship credit should be based on substantial contributions to:
  (1) conception and design, or acquisition of data or analysis and interpretation of data; and
  (2) drafting the article or revising it critically for important intellectual content; and
  (3) final approval of the version to be published.
  Authors should meet conditions 1, 2 and 3.
- Acquisition of funding, collection of data or general supervision of the research group does not justify authorship.

To provide information on the work done by authors and to resolve the inconsistency between the information provided for those named in the byline versus those listed in the acknowledgements, some journals require authors to indicate the specific contributions of all those involved. Consequently, authors are required to describe their specific contributions as well as the contributions of those acknowledged but not listed in the byline. While many individuals may contribute to the work of an article, the contributors must decide for themselves what their contributions have been, and what level of contribution merits a place on the by-line. It is suggested that those in the by-line should be listed in order of actual contribution made, as decided by the authors.

Any contributors who do not meet the criteria for authorship should be listed in the acknowledgement section. Because readers may infer their endorsement of the data and conclusions, all persons listed must give written permission to be acknowledged.

12.8 Patents and publication

There is a growing understanding of intellectual property rights by academics and scientific institutions. The importance of the issue has already been discussed in Chapter 10. It is accepted that a private or public sponsor of the study has the right to review a manuscript for a defined period (for example 30 to 60 days) before publication to allow for the filing of additional patent protection if required. However, the sponsor must impose no impediment, direct or indirect, on the subsequent publication of the full results of the study.

12.9 Ethics in scientific publication

12.9.1 Credit

Researchers must get the credit for the research they have carried out. However, as explained in section 12.7, no one should get credit without having actively participated in the research. All authors should agree to have their names on the paper, and to take public responsibility for it. The order of authors should be by agreement among the authors.

The work of previous investigators on the topic in question should be cited. The investigator should not claim credit for an idea that has already been put forward or studied by others, and should indicate previous studies that may have shown different results and conclusions.

The contribution of others who have helped in the implementation of the research should be acknowledged, and the source of support for the research should be identified.

12.9.2 Respect of copyright

Copyright should be respected. The principle behind the copyright law is relatively simple. Copyright begins at the time a creative work is recorded in some tangible form. In scholarly work, there is seldom a financial compensation for copyright, as in other fields, but there is certainly the need for recognition. No figure or table from previously published work should be included without written permission from the publisher and author. Full credit to the source should be included in the paper ("Reproduced with permission from…").

Plagiarism is a major ethical offence. Using and claiming the words or ideas of another person as one's own, without acknowledging their contribution, is not acceptable.

12.9.3 Conflict of interest

"Disinterestedness" is a norm of science. When investigators have vested interests in the research, this should be explicitly disclosed. A statement on conflict of interest is now required by many journals before considering a paper for publication. As commerce and academia work closer together, there is the potential for financial and funding ties to distort the work. Some studies which reviewed published reports of clinical trials have suggested that clinical trials were more likely to reach conclusions that were favourable to the intervention, when supported by for-profit organizations (Als-Nielsen, 2003).

12.9.4 Redundant or duplicate publication

Redundant or duplicate publication is publication of a paper that overlaps substantially with one already published by the same authors.

Readers of primary source periodicals should be able to trust that what they are reading is original unless there is a clear statement that the article is being republished by the choice of the author and editor. This position is based on international copyright laws, ethical conduct and cost-effective use of resources.

Most journals do not wish to receive papers on work that has already been reported in large part in a published article or is contained in another paper that has been submitted or accepted for publication elsewhere, in print or in electronic media. This policy does not prevent the journal from considering a paper that has been rejected by another journal, or a complete report that follows publication of a preliminary report, such as an abstract or poster displayed for colleagues at a professional meeting. Nor does it prevent journals from considering a paper that has been presented at a scientific meeting but not published in full, or that is being considered for publication in proceedings or similar format.

When submitting a paper, the author should always make a full statement to the editor about all submissions and previous reports that might be regarded as redundant or duplicate publication of the same or very similar work. The author should alert the editor if the work includes topics about which a previous report has been published. Any such work should be referred to and referenced in the new paper. Copies of such material should be included with the submitted paper to help the editor decide how to handle the matter.

If redundant or duplicate publication is attempted or occurs without such notification, authors should expect editorial action to be taken. At the least, prompt rejection of the submitted manuscript should be expected. If the editor was not aware of the violation and the article has already been published, then a notice of redundant or duplicate publication will probably be published with or without the author's explanation or approval.

Acceptable secondary publication

Secondary publication in the same or another language, especially in other countries, is justifiable and can be beneficial, provided all of the following conditions are met.

- The authors have received approval from the editors of both journals; the editor concerned with secondary publication must have a photocopy, reprint or manuscript of the primary version.
- The priority of the primary publication is respected by a publication interval of at least one week (unless specifically negotiated otherwise by both editors).
- The paper for secondary publication is intended for a different group of readers; an abbreviated version could be sufficient.
- The secondary version carefully reflects the data and interpretations of the primary version.
- The footnote on the title page of the secondary version informs readers, peers and documenting agencies that the paper has been published in whole or in part and states the primary reference. A suitable footnote might read: "This article is based on a study first reported in [title of journal with full reference]."

Permission for such secondary publication should be free of charge.

12.9.5 Protection of patients' rights to privacy

Patients have a right to privacy that should not be infringed without their informed consent. Identifying information should not be published in written descriptions, photographs and pedigrees unless the information is essential for scientific purposes and the patient (or parent or guardian) gives written informed consent for publication. Informed consent for this purpose requires that the patient has the right to be shown the manuscript to be published.

Identifying details should be omitted if they are not essential, but patient data should never be altered or falsified in an attempt to attain anonymity. Complete anonymity may be difficult to achieve, and informed consent should be obtained if there is any doubt. For example, masking the eye regions in photographs of patients is inadequate protection of anonymity.

The requirement for informed consent is normally included in the journal's instruction for authors. When informed consent has been obtained it should be indicated in the published article.

12.9.6 Release of results to public media

Researchers should look for recognition primarily among their peers. It is not considered ethically acceptable for researchers to break news of their findings to the public or the media before they have been communicated to their peers in scientific press or meetings. Preliminary release of scientific information described in a paper that has been accepted but not yet published violates the policies of many journals. In exceptional (and rare) cases, and only by arrangement with the editor, preliminary release of data may be acceptable, for example, if there is a public health emergency. Some journals issue press releases about important findings to coincide with publication.

12.9.7 Scientific fraud

Research misconduct can be regarded as a continuum ranging from errors of judgement (that is, mistakes made in good faith) such as inadequate study design, bias, self-delusion and inappropriate statistical analysis, to what may be regarded as misdemeanours (also called "trimming" and "cooking") such as data manipulation, data exclusion, suppression of inconvenient facts, through to blatant fraud, usually categorized as fabrication, falsification and plagiarism (Farthing, 1998). The culture of science is based on trust. When a researcher presents his/her data in public, the data are taken at face value. One may interpret data differently, question the study design or disagree with the statistical analysis. However, if we cannot trust the data, the whole atmosphere of science is poisoned. There is also the impact on the public. Every single case of fraud and misconduct reduces public confidence in science. It also indicates that public and charitable funds may have been abused. Cases of scientific fraud are causes for embarrassment and frustration to the vast majority of honest scientific researchers.

In the United Kingdom, a Committee on Publication Ethics (COPE), made up of medical journal editors was established in July 1997, some two years after a senior gynaecologist was struck off the medical register by the General Medical Council for fabricating evidence that was published, including a claim to have successfully relocated an ectopic pregnancy and also a three year trial of a hormone treatment for recurrent miscarriage. Neither the relocated ectopic pregnancy nor the trial had ever taken place. Information about the work of COPE and its periodic reports is available on the internet (http://www.publicationethics.org.uk)

12.9.8 Ethical responsibility of journal editors

Editors should take all reasonable steps to ensure the accuracy of the material they publish. Whenever it is recognized that a significant inaccuracy, misleading statement or distorted report has been published, it must be corrected promptly and with due prominence. If articles prove to be fraudulent or contain major errors that are not

apparent from the text then they should be retracted—and the word retraction should be used in the title of the retraction (to ensure that it is picked up by indexing systems). Cogent critical responses to published material should be published unless editors have convincing reasons why they cannot be. Some journals have created electronic means of responding, so that "lack of space" will not be a reason for not publishing a response. Editors should ensure that research material they publish has been approved by an ethics committee. In addition, they should satisfy themselves that the research is ethical as they can be held responsible for publishing "unethical" research even if it has been approved by an ethics committee.

Editors must protect the confidentiality of information on patients obtained through the doctor–patient relationship. If ensuring anonymity is not completely possible, written consent for publication from patients should be obtained.

References and additional sources of information

Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials. *Journal of the American Medical Association*, 2003, 290: 921–928.

Byrne DW. *Publishing Your Medical Research Paper. What they don't teach in medical school*. New York, London, Lippincott Williams & Wilkins, 1998.

Delamothe T. Is that it? How online articles have changed over the past five years. *British Medical Journal*, 2002, 325: 1475–1478.

Farthing MJG. Ethics of publication. In: Hall GM, ed. *How to write a paper*. 2nd edition. London, BMJ Books,1998: 122–131.

Grant JM and Laird A. How a paper is reviewed and why it might be turned down. In: O'Brien PMS, Pipkin FB, eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 231–240.

Huth EJ. *How to write and publish papers in the medical sciences*, 2nd edition. Baltimore, Williams & Wilkins, 1990.

International Committee of Medical Journal Editors. *Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication*. Updated November 2003. (http://www.icmje.org accessed 24/2/2004)

Jones J. UK watchdog issues guidelines to combat medical research fraud. *British Medical Journa*l, 1999, 319:660.

Weeks WB, Wallace AE. Readability of British and American medical prose at the start of the 21st century. *British Medical Journal*, 2002, 325: 1451–1452.

Chapter 13

Making a scientific presentation

13.1 Introduction

The quality of presentations in scientific meetings often leaves much to be desired. A number of sources are now available to help researchers improve their presentations. Some are listed under the references and additional sources for this chapter. A good scientific presentation must follow the following three "Ps". It should be: Planned with care, Prepared with care and Presented with care. The following sections provide some useful guidelines, particularly for beginners.

13.2 Planning of the presentation

In planning a scientific presentation, presenters need to ask the organizers of the scientific meeting about: the audience and their level of knowledge and interest in the subject since the planning of the presentation will be different for a specialist audience, a generalist audience or a mixed audience; the time available for the presentation; and the type of visual aids available. Presenters should ask themselves what the main message (or messages) is that they would like to convey and how it can be conveyed to the type of audience concerned in the time allotted.

The manuscript of an article (as submitted for publication) should not be used as such for a scientific presentation. The difference between speaking and writing is the same as the difference between hearing and reading. A reader chooses his own pace; the listener must accept the pace chosen by the speaker. Listening to the news on television is different from reading the news in a newspaper.

To change a written scientific paper into an oral presentation, the presenter must follow three "s words": Select, Synthesize, and Simplify. Select from the written article the points to present. Synthesize the information in the article to package it in the limited time available. Simplify the presentation of the data, so that it can be easily followed and understood by the audience.

In the planning stage, the title of the presentation has to be decided and an abstract has to be submitted to the organizers of the scientific meeting. A good title can be defined as the fewest possible words that adequately describe the contents of the presentation. The

abstract can attract or put off the audience. The abstract is the part of the presentation that will be published in the conference programme. A good abstract should be a miniature version of the presentation. The abstract should be sent to organizers before the deadline and in the format and length requested

13.3 Preparation

13.3.1 Preparation of text

In preparing the text of a scientific presentation:

- Avoid too much detail and resist the temptation to overload the presentation with information.
- Avoid jargon and abbreviations, unless they are clear to all the audience.
- Aim at the average person in the audience.
- Use plain English.

The structure of a presentation is different from the structure of a written paper. Normally, it should consist of three parts: introduction, main message and conclusions.

The introduction should tell the audience what the presentation will be about. Where possible, the opening sentences should capture the attention of the audience. It helps to have something like a “punch line”, which will alert the audience to the importance of the subject.

The main message should be clear and concise. The usual detail of a written paper is unsuitable for a presentation. It is generally unwise to introduce more than one new idea every 2 to 3 minutes.

The conclusion should summarize the main points. Try for a strong finish. Stopping speaking is not finishing. Leave the audience with a “take home message”.

13.3.2 Preparation of visual aids: speaking visually

Objectives for using visual aids

It has been said that we remember 20% of what we hear, 30% of what we see, but between 50% and 75% of what we see and hear (Sorgi and Hawkins, 1985). A Chinese proverb says “A picture is worth a thousand words”. Visual aids are not an objective in themselves. They are used to serve one or more of the following objectives:

- holding the attention of the audience
- presenting the data in a clear way
- delivering the presentation without having to read from notes.

Commonly used visual aids include slides, overhead transparencies and computer-assisted presentations.

Slides

Slides are the commonest visual aid used in scientific presentations. They can make or break the presentation. Until recently photographic film slides were very commonly used; now electronic slides presented as a data show have largely taken over. The basic rules for a good presentation are the same for film slides and electronic slides. There are three main types of slide: text slides, data slides (tables, graphs, flow charts) and figure slides. A mix of text, data and figure slides helps to maintain the interest of the audience.

Text slides are not meant to be read by the speaker, but by the audience. Lettering should generally be limited to 4 lines and should never be more than 7, including the title. It is advisable not to use more than 8 words per line.

Complicated tables are not visual aids. They have been described as instruments of torture for the audience. Tables of data suitable for written publication are highly unsuitable for a scientific presentation. The term "Railway Timetable slides" is sometimes used to describe the difficulty with slides showing complicated tables. Do not use more than seven lines (including title) and four columns in any table. The writing on a film slide should be easily legible by the naked eye. Use the whole area of the slide. There is no need to put the data in an outer box. Note in the design of the table that the transparent area in a film slide is not square but oblong. Columns are preferably separated by a space larger than the width of the column.

Graphs should replace tables where possible in a visual presentation. They are better in showing relationships. Preparation of graphs has now been made easy by computer programs. Four types of graphs are often used: bar or column charts; curves; pie-charts; and scatter graphs.

- Bar charts are better for lettering than column charts. Avoid overcrowding the slide. The number of bars should be limited to five to seven. An overcrowded column chart is sometimes called a "New York Skyline" slide, to emphasize that it is not suitable for presentation.
- No more than two or three curves can be shown on a slide. Space on the slide should not be wasted.
- The slices of a pie-chart must not be too numerous nor too small. Three to five divisions are ideal.
- Scatter graphs are good for slide presentation. They give a clear and simple overview of the scatter of the data to show relationship.

- Flow charts should not be complicated. A complicated flow chart looking like a "subway" map is not useful for a presentation. A complicated flow chart can be built up in a number of successive slides.

Figure slides of drawings and pictures, if meant for humour, should be selected with care and sensitivity to the type of audience. They should not offend the feelings of anyone in the audience.

Tips in slide preparation

- A common mistake is to try to put too much on one slide. As a general rule, no slide should be shown unless it can be read by the back row of the audience. As a general rule, lettering on a film slide should be large enough to be read by the naked eye without projection.
- The shape of a film slide is rectangular: 36 × 24 mm. The dimensions of the material on the slide should be prepared with this in mind.
- Upper case letters are less legible than lower case letters. This is why lower case is commonly used in direction signs on motorways and on the underground. Our eyes are more accustomed to small letters in books and newspapers.
- While choice of colour is a matter of taste and judgement to a certain extent, colour should not be used for decoration but to improve understanding. Select colours that project well. Popular combinations are blue and white, and green and yellow. Red text may be more difficult to read. The number of colours should be limited to what is really necessary for presenting the data in a clear way.

Computer software

Computer software is used for preparation of electronic slides for a data show. A widely used program is Microsoft Powerpoint. The same program can make the preparation of 35-mm film slides easier and better. The file of slides created on the computer can be sent as a floppy disk or via a modem to a bureau for creating film slides. Computer generation of electronic or field slides offers a number of additional advantages. The software guides you through the preparation, provides templates and recommends consistent colour schemes. Preparation of graphs is easy. Photographs and drawings can be imported from other software programs. The program allows each slide to have a text note attached and the slide and note can be printed out on the same paper page to serve as speaker's notes. A number of slides can be printed out on one page of paper to be used as audience handouts. The slides created for a presentation can be viewed and edited on the computer screen. Slides are saved and can be included in another presentation. The slides created for a presentation can be viewed in a timed fashion on the screen and the timing of the accompanying talk can be checked and adjusted.

Overhead transparencies

The overhead projector is a natural successor to the chalkboard. It is particularly useful in presentations to small groups. Overhead transparencies, as visual aids, have advantages and disadvantages.

The advantages of overhead transparencies are that:

- they may not need to have the room darkened;
- the speaker faces the audience, allowing better eye contact;
- they are inexpensive to make;
- they can be made quickly, using the copy machine or a computer printer with compatible transparent plastic sheets;
- overhead projectors are usually readily available, are easy to set up and are less likely to break down; a projectionist is not required;
- the speaker can write directly on the film by a marking pen;
- information can be built up in a dynamic way by either drawing directly on the transparency, or by adding transparent overlays;
- colour can be easily used.

The disadvantages of overhead transparencies are that:

- they are not suitable for large audiences;
- the projected image is not as sharp as the slide;
- the projector cannot be put in a projection booth;
- they can give the impression of being prepared in haste if not carefully revised and well presented.

Overhead transparencies are easily prepared. Handwriting does not produce an elegant transparency. It gives the impression of last minute preparation. It may be more acceptable if the writing or drawing is done during the presentation. A photocopy machine can produce a nice transparency from the printed output of a word processor. Only special transparency sheets suitable for a photocopy machine should be used. A computer printer can print directly on special transparency sheets suitable for either laser jet or colour ink-jet printers. Note that transparencies need longer drying times than regular paper. The computer software may allow printing the transparency as a flipped document, in which the text and pictures are reversed. The transparency printed in this way is projected face down. This allows the speaker to write on the back of the transparency during the presentation. It is easy then to wipe the writing off later without scratching the original.

Computer-assisted presentation

The same computer software that is used to produce 35-mm slides can produce a screen show of slides, with manual or automatic control over timing between individual slides. The slide show can be projected directly to the audience. Notebook computers have a port to allow connection to an external monitor or to a special projector. The equipment is rather expensive but it produces a very elegant presentation, including the use of moving text and images. However, do not overuse the animation features as they can distract the audience and become annoying. A good presentation is also a simple presentation. Slides can be easily sorted and their order re-arranged. The technology is rapidly becoming the standard for the use of visual aids.

Before preparing a computer-assisted presentation check about the availability of the equipment. Since the new technology is prone to equipment failure, it is advisable to have a backup of slides or overhead transparencies. It is better not to try using this new technology for the first time in an important meeting, particularly in settings which may not have experience with it.

13.3.3 Rehearsal

The preparation of the text and of the slides has to take the allotted time into consideration. Rehearsal is the key to making sure that you will deliver the presentation without exceeding the time. Even very experienced speakers rehearse their presentations. You can rehearse on your own, or with the help of colleagues.

A pleasant average rate of delivery is not more than 120 words a minute. A word processor can give the exact word count of a written presentation. A double spaced typewritten page is about 240 words. For a ten minute presentation, plan on no more than five pages of double-spaced text.

A general rule is one slide per minute if the slide contains information, and one slide every 5-10 seconds if the slide contains only titles, key words, or is designed just to remove another visual from the screen. The exact time for the non-information slide will also depend on the amount of script to be covered while it is displayed. Having to skip slides during the presentation, because the slides are too many, means that preparation of the presentation was poor.

13.4 Presentation

The challenge to the speaker is to hold the attention of the audience. Particularly when the lights are dim, the audience can have sweet dreams during a boring presentation (Harvey et al., 1983).

- Get ready
- Speak well
- Manage your slides
- Keep to the time
- Be prepared to answer questions.

Getting ready

It is always advisable to check the room where the presentation will be given, in advance. Check the podium for the microphone, the remote control for the slide projection, the slide pointer and the lights. Provide your slides, properly arranged, or diskette to the technician for projection.

Speaking well

Perfection in speaking is acquired. It is acquired by practice, by observing good speakers, and by learning from your own mistakes as well as the mistakes of other speakers. If you are excited and eager to share, others will warm to you. If the microphone is to be attached, attach it to the lapel of the jacket or dress, and not to a movable part such as the necktie. It can produce a distracting background noise when you move. Look the audience in the eye.

It is more effective not to read your presentation. If, however, you read from a script, the script should be written for hearing not reading. Prompter cards or prompter slides can help the speaker to deliver the presentation without having to read. The generally accepted rate for easy hearing and understanding is not more than 120 words-a-minute, as indicated above. Pauses in speaking replace punctuation in writing: comma: break of one second; semicolon: break of two seconds; period/full stop: break of three seconds; paragraph: break of four seconds. Varying the tone, pitch and volume helps to maintain the attention of the audience.

Managing slides

Mark and number film slides. If a slide is projected upside down, there are seven possible ways of showing it again wrongly, before the correct orientation is discovered. The international convention calls for a spot to be placed in the lower left-hand corner as the slide is viewed by the naked eye. This should be visible at the upper right corner when the slide is inserted. Check your slides before the presentation. Well organized conferences usually have a preview room where this can be done.

Remember the saying that if anything can go wrong, it will. Be prepared for the possibility of breakdown of visual equipment. It is generally advisable to start the presentation with the lights on. Keep the lights off till you complete showing the slides.

Use "filler" slides if needed, to avoid having lights on and off during the presentation. But, it may be good to conclude while the lights are on, to make a strong finish.

Do not read the slides. You can safely assume that the audience is literate and is not blind. An exception can be made in case of simultaneous translation, so that the translators can translate the slide which is read. Better still, provide translators with a copy of your text notes. Do not go back to a previous slide. Insert a copy.

The use of two projectors in parallel, with two screens (dual projection), and two sets of slides is really only useful when you want to show changes that are difficult to demonstrate unless two slides are compared side by side. The audience must be given time to look at both slides. A good rule is never to show two text slides at the same time.

Keeping to time

The speaker who exceeds his allotted time is guilty of gross bad manners. He imposes not only on his audience, but also on all the speakers who come after him. It is a sign of poor preparation.

Answering questions

Answer politely: Do not answer questions in a dismissive or confrontational manner. Answer knowledgeably. Remember that "I do not know" is a good answer.

13.5 Guide to how to give a "bad" presentation

(Based on a humorous piece by Richard Smith, editor of the *British Medical Journal*, 2000)

- Forgetting altogether that you agreed to speak is a good way to make a mess of your presentation. A variant is to arrive late. Don't arrive too late because they will simply have cancelled your session, probably sending a thrill of pleasure through an audience facing the prospect of five consecutive speakers.
- One way to prepare for a bad presentation is not to prepare at all. Step up to the platform, open your mouth, and see what comes out. This is, however, a high-risk strategy because spontaneity may inspire both your audience and you. Inspiration must be avoided at all costs.
- A really bad presentation needs careful preparation. A good piece of advice is to prepare for the wrong audience. It is much the best strategy to give an overcomplicated presentation than an oversimplified one.

- Be sure to prepare a presentation that is the wrong length. Too long is much the best. Most of the audience will be delighted if your talk is too short. But something that is too long always depresses an audience, even if what you are saying is full of wit and wisdom.
- Another trick is to ignore the topic you are given, and speak on a completely different subject.
- You may be able to enhance your bad presentation by sending the organizers in advance a long and dull curriculum vitae to read before your presentation.
- Bad slides are the traditional aid of a bad presentation. They must be far too many, contain too much information and be too small for even those in the front row to read. Flash them up as fast as you can, ensuring that they are in the wrong order with some slides upside down. Ideally there should be little connection between what you are saying and what is on the slide.
- The essence of a bad presentation is to be boring. Anything that isn't boring will detract from your bad presentation.
- Never look at the audience. Mumble your presentation, and preferably read it. A presentation that is read will usually be satisfyingly bad, but for the full effect you should have long complicated sentences with dozens of sub-clauses.
- A truly bad presentation rarely produces any questions. Most people will just want to get away. If you do get questions, you may have failed in giving a bad presentation. But all is not lost. By sticking to the basic rules of being boring and overcomplicated, and by speaking too long, you may still be able to rescue your bad presentation. The extra rule on answering questions is that under no circumstances should you really answer them. Once you have finished say, "Does that answer your question?" If the questioner has the effrontery to say no, then do it again, only at greater length.

References and additional sources of information

Harvey RF, Schullinger MB, Stassinopoulus A, Winkle E. Dreaming during scientific papers. *British Medical Journal*, 1983, 2: 1916–1919.

Hawkins C. Speaking at meetings. In: Hawkins C, Sorgi M, eds. *Research: How to plan, speak and write about it*. Berlin, Springer-Verlag, 1985: 60–84.

Hextall A, Cardozo L. Presenting a paper. In: O'Brien PMS, Pipkin FB, eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 218–224.

Lashford LS. Presenting a scientific paper, including the pitfalls. *Archives of Disease in Childhood*, 1995, 73: 168–169.

Smith R. How not to give a presentation. *British Medical Journal*, 2000, 321:1570–1571.

Sorgi M, Hawkins C. Illustrating talks and articles. In: Hawkins C, Sorgi M, eds. *Research: How to plan, speak and write about it*. Berlin, Springer-Verlag, 1985: 110–135.

Thompson WA et al. Scientific presentations. What to do and what not to do. *Investigative Radiology*, 1987, 22: 224–45.

Chapter 14

Assessment and evaluation of research

14.1 Introduction

Researchers need to have the skill to assess and evaluate the research papers they read, particularly those related to the research topic they are doing. This should be done before the research is planned, during the implementation of the project, and before discussing the results and preparing to communicate them. Researchers may also want to critically assess all accessible published papers on a particular topic in order to write a systematic review. They should bear in mind that science should not be admired; science should be questioned. The words "author" and "authority" come from a common English stock and run the danger of becoming synonyms in the minds of some. A good scientist should develop a sceptical attitude when reading scientific papers. Scepticism is an inherent part of the scientific approach. What defines any statement as being scientific is that it is verifiable in principle, or, as it is sometimes put, it should be "falsifiable" in principle. There is hardly any theory in science that ever achieves a degree of certainty beyond the reach of criticism or the possibility of modification. In science, there will always be more beyond.

Researchers may also be requested to peer-review a scientific paper submitted for publication by other researchers, or to assess the scientific output of candidates for academic posts.

The need to assess and evaluate research is not limited to researchers. Learning to evaluate and use research findings is an important and lifelong part of professional development for health professionals. They need to critically assess the value of new published research before considering its practical implications for their work. Health professionals need to be aware of the fact that there are different levels for scientific evidence. Health researchers should help in outlining these different levels of evidence.

Policy-makers should have the ability to assess research results and their implications for policy. In particular, they need to assess new technologies and also currently used technologies, to introduce what is new and cost-effective, discard what is not effective or potentially harmful, promote what is effective but under-utilized, and postpone a

decision where evidence is still lacking. Health researchers need to be aware of these considerations.

Research is an investment, and is becoming more and more expensive. Those who fund the research need to evaluate the return on their investment. Researchers need to be aware about how the investment in health research is evaluated by funding agencies, particularly governments, their public paymasters.

This chapter addresses the assessment and evaluation of research by researchers, health professionals, policymakers, and investors in health research. For additional information on the subject, the sources listed in the references and additional sources for the chapter can be consulted.

14.2 Assessment and evaluation by researchers

14.2.1 Reading a research paper

The title of the paper and the abstract give an indication of the novelty and relevance of the paper.

For the critical reader, the methods section should be the first part of the paper to assess. It will tell whether it is good science or bad science. It has been rightly said that a paper will sink or swim on the strength of its methods section (Greenhalgh, 1997). A good methods section should provide sufficient detail to allow other investigators to replicate the study and confirm the results. If it does not, the study results cannot be easily accepted.

In most papers, the two most important methodological issues relate to how the sample was selected and what measurements were made. The sample must be representative of the population studied. If two samples are compared, they must be selected to be identical for every relevant variable, except the one to be studied. The critical reader must question whether the measurements used have been assessed for their validity and their reliability. As discussed in Chapter 4, validity is an index of how well a test or procedure measures what it is intended to measure. Reliability assesses consistency of measurement. It relates to the reproducibility of measurements. When reliability is high, a test that is repeated on the same patient and under the same conditions will yield the same result, whether by different investigators (Inter-rater reliability), or by the same investigator (Intra-rater reliability). Where appropriate, the investigators should provide assurance about the quality control of their data. As an example of the importance of inter-rater reliability, one study looked at the agreement among four pathologists on the classification of cervical intra-epithelial neoplasia, compared with the index pathologist. Of 101 cases of carcinoma in situ (CIS), 6 were reported as mild dysplasia, 19 as moderate dysplasia, 54 as severe dysplasia, and 22 as CIS (deVet et al., 1990).

The critical reader of a scientific paper takes a close look at the results and their interpretation. Pitfalls in the interpretation of research results are discussed in detail in Chapter 9.

Statistical jargon should not put off the critical reader. Use and interpretation of statistics can be misleading. Disraeli is quoted as saying "There are three types of lies: lies, damn lies and statistics". One does not need to be a statistician to make some judgement about the statistical analysis of the research. Statistics is about common sense, before it is about mathematics. The first question to ask is whether the authors have used any statistical methods at all. If they have not, there is no reason to accept that the results are not being caused by chance alone. The second question is whether the authors have selected the right statistical methods to analyse their data. The third question is whether they have drawn the right conclusions from the statistical analysis. It is tempting to make wrong conclusions on the basis of statistical analysis. There is a limit to what statistics can tell us.

14.2.2 Peer review

Peer review is the critical assessment of manuscripts submitted to scientific journals by experts who are not part of the editorial process. The process of peer review helps editors to decide which manuscripts are suitable for publication, and helps authors to improve the quality of their papers. A peer-reviewed journal is a journal that submits most of its published research articles for outside review.

In the peer review process, editors generally provide reviewers with a format for the assessment of all components of the paper, from the title to the references. There is a common misconception that finding flaws is key to the high quality of peer review. The objective of the peer review process is not to find something to criticize. Finding flaws is certainly important, and scepticism is revered in scientific tradition. Authors can benefit from constructive criticism of good reviewers. However, responding to misguided comments may waste time and effort.

There are ethical considerations in the peer review process. Reviewers must disclose to editors any conflicts that could bias their opinions of the manuscript, and they should disqualify themselves from reviewing specific manuscripts if appropriate. Editors should avoid selecting external peer reviewers with obvious potential conflict of interest, for example those who work in the same department or institution. Reviewers must not use knowledge of the work before its publication to further their own scientific interests.

14.3 Assessment and evaluation by health professionals

14.3.1 Levels of evidence

Health professionals reading scientific papers for possible clinical application should recognize that there is a hierarchy of the level of evidence obtained from different study designs. In assessing the effectiveness of 169 interventions, the U.S. Preventive Services Task Force (1989), including a 20-member panel of scientific and medical experts, proposed the following guide for rating the quality of evidence for clinical effectiveness.

- Level I evidence: Evidence obtained from at least one properly designed randomized controlled trial
- Level II-1 evidence: Evidence obtained from well-designed controlled trials without randomization
- Level II-2 evidence: Evidence obtained from well-designed cohort or case-control studies. In these observational studies, the investigator has no role in assignment of study exposure but, rather, observes the natural course of events of exposure and outcome.
- Level II-3 evidence: This category includes cross-sectional studies, which are observational studies that assess the status of individuals with respect to the presence or absence of both exposure and outcome, at a particular time. The category also includes uncontrolled intervention studies. They may demonstrate impressive results, but in the absence of a control group the results may be attributable to factors other than the intervention or treatment. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the1940s) may, however, be difficult to dismiss.
- Level III evidence: This category includes descriptive studies, such as case reports and case series. It also includes expert opinion, often based on clinical experience.

14.3.2 Systematic reviews and meta-analyses

Results of scientific studies are often not uniform. To try to draw conclusions from these studies, systematic reviews are undertaken by researchers. A systematic review, as outlined in Chapter 11, is an overview of primary studies that contains an explicit statement of objectives, materials and methods, and has been conducted according to explicit and reproducible methodology. It is different from a narrative review, which is an overview of primary studies that have not been identified or analysed in a systematic (standardized and objective) way.

The quality of systematic reviews should generally be judged by the following two criteria:

- Have the authors performed a thorough literature review or presented only selected research findings?
- Have they accepted the primary researchers' interpretation of study data uncritically, or do they include methodological commentary along with their content review?

A meta-analysis, as discussed in Chapter 11, is a special type of systematic review that combines results from more than one investigation to obtain a weighted average of the effect of a variable or intervention on a defined outcome. Combining data from a number of studies increases the sample size and the power of the study to provide statistically significant conclusions. A meticulously conducted meta-analysis, in which all the primary studies on a particular subject have been hunted out and critically appraised according to rigorous criteria, has a very high place in the hierarchy of evidence.

In reading a meta-analysis study, it should be recognized that a meta-analysis can only be as good as the quality of its individual components. Assessment of quality of a meta-analysis has to address the following questions:

- Is the pooling done only among studies where there is reasonable assurance that subjects and treatments are similar? Misleading conclusions can be drawn from pooling together heterogeneous data.
- Has care been taken to exclude publication bias toward positive results? Studies with positive results are more likely to be published, leading to problems with meta-analysis interpretation; many researchers are reluctant to pursue and publish negative results.

14.3.3 Cochrane Collaboration

The Cochrane Collaboration focuses on identifying reliable evidence and preparing systematic reviews of therapeutic interventions using randomized controlled trials (RCTs) (Bero and Rennie, 1995). Archie Cochrane was a Scottish epidemiologist who worked in Wales for most of his life. In 1972, he wrote a book in which he highlighted the absence of an adequate knowledge base for much of the health care provided. He made a strong case for the evaluation of new and current forms of care in controlled trials, which use randomization to generate unbiased comparison groups. Cochrane first challenged the profession of obstetrics to seek good evidence for its practice. The challenge was taken up, and the database of perinatal trials was the first to come out. Having demonstrated that the approach was possible with one specialty, the work was extended to other areas of health care. In 1992, the first Cochrane Centre was opened in Oxford, and the Cochrane Collaboration was launched internationally one year later. The Cochrane Library (http:

//www.update-software.com/cochrane/) is currently considered one of the best single sources of critical evidence for health care interventions. The library publishes a database solely of RCTs. It is published on a quarterly basis and made available both on CD-ROM and on the internet. It is easily accessible in a user-friendly format. It is the result of collaborative hand-searching efforts and electronic searching from many of the different review groups and centres of the Cochrane Collaboration. Collaborative review groups have evolved, which cover most areas of health care.

14.4 Assessment and evaluation by policy-makers

There has been an explosion of technologies in the past few decades as an outcome of the expansion in health research. These technologies provide great opportunities in health care. The assessment of these technologies presents major challenges to health policy-makers. A major challenge is how these technologies can be assessed to determine their appropriateness. Assessment should not be limited to newly introduced technologies. There is a need also to assess technologies currently in use, which may not be effective or even potentially harmful. There are also beneficial technologies which may be under-utilized. Technology can be defined as the implementation of scientific knowledge in order to satisfy human needs. Health technologies include the drugs, devices, equipment and medical and surgical procedures used in the prevention, detection, diagnosis and treatment and rehabilitation of disease.

The responsibility for assessment of health technologies is ill defined. Drug regulatory authorities have responsibility for the approval of drugs for human use. Based on pre-clinical and clinical studies, the authority decides whether the drug is safe and effective to do what it is claimed to do. But it is not the business of the drug regulatory authority to compare the drug with other available drugs. It only ensures that the manufacturer makes no unjustified claims. This is the status of drug regulation, but health technologies include also devices, equipment and procedures. Devices are only regulated if they are used inside the human body. Medical equipment and medical and surgical procedures are not, in general, subject to regulation by authorities; not that such regulation is desirable in a rapidly advancing field.

The following four questions need to be carefully examined before any new technology is considered appropriate:

- Is the technology evidence-based?
- Is it good value for money?
- Is it culturally and ethically acceptable?
- Are the system requirements for its introduction available?

Is the technology evidence-based?

There is a need to critically assess the evidence before adopting any new technology. This is particularly important when there are strong commercial interests involved. The practice of medicine has been rapidly evolving from being authority-based to being evidence-based. The history of our medical practice is not short of examples of technologies which were widely used and subsequently proved not useful or even harmful.

There are ongoing efforts to assess currently available health technologies. In an ongoing assessment of reproductive health technologies, WHO classified these technologies into the following six categories: beneficial, likely to be beneficial, with a trade-off, of unknown effectiveness, likely to be ineffective, and likely to be harmful (WHO, 2002). In the UK, the National Institute for Clinical Excellence (NICE) was set up as a special health authority for England and Wales in 1999. Its role is to provide patients, health professionals, and the public with authoritative, robust and reliable guidance on current "best practice" (www.nice.org.uk).

Is the technology good value for money?

If the technology is evidence-based, the next question is whether it is good value for money. This is a different question from the issue of affordability. Economists have shown an increasing interest in what health professionals are doing, contributing a new discipline of health economics. With the increasing introduction of health technologies, health care has become too costly to be left to health care providers alone. Research on health economics is discussed in Chapter 4.

Health economists introduced two important concepts to consider in deciding whether a new technology is good value for money: cost-effectiveness and opportunity cost. Cost-effectiveness measures the net cost of providing a service as well as the effectiveness of the service. The result of cost-effectiveness analysis is expressed as the monetary cost per unit of effectiveness. To illustrate this concept, let us take the example of an assisted reproduction technology procedure. The cost is measured against the desired outcome, "a take home baby", not simply by the cost of the procedure. If the success rate is, say, 25%, then the cost per take home baby is four times the cost of the procedure. If a new technology is claimed to raise the success rate by 10%, but the procedure also has an additional cost, we need to bear in mind that, for each one additional "take home baby", ten patients must receive this new procedure. The additional cost of one "take home baby" will be ten times the additional cost of the new procedure.

The second economic concept in judging whether a technology is good value for money is the opportunity cost. The concept implies that if resources are used in one way, an opportunity to provide some other benefit has to be renounced. To illustrate

this concept, take the example of a health policy-maker deciding on whether to provide infertility patients with free assisted reproduction services. The issue is not simply about having enough budget. There are other health services which can be "bought" with the same level of resources. The issue is what opportunities the policy-maker will miss if resources are allocated to this service.

Is the technology culturally and ethically acceptable?

The next question to address is whether the health technology is culturally and ethically acceptable. The assessment has to be done in the context of each country and religion. This question is particularly important in reproductive health technologies. Assisted reproduction technologies and fertility control technologies are such examples.

Are the system requirements available?

"System requirements" have to be carefully checked before any new technology is considered. This term is used in computer jargon. If we want to install a new software program on the computer, we are asked to check that the system requirements are available, in terms of operating system, free memory, etc. If we do not have the system requirements and we still try to install the software, the attempt will be rejected. New health technologies have system requirements, in terms of facilities, qualified and trained personnel, maintenance and supply logistics. If we try to install a new technology where the system requirements are lacking, it will not be rejected by the system, but it will not perform as desired, and may even do more harm than good, wasting resources in the process.

Social concerns are often expressed about the proliferation of new health technologies. Health professionals need to be socially conscious and fully aware of these concerns. There is concern that the proliferation of health technologies is getting out of hand, contributing to escalating and soaring costs of health care. There is concern that the health divide between rich and poor may widen, if the new technologies are more responsive to the needs of the rich and are available only to those who can afford their high cost. Then, there is the concern that medicine may be moving too far away from its social roots, and that health professionals are becoming technicians rather than humane physicians. Hippocrates wrote in about 400 BC: "Whoever wishes to investigate medicine properly should proceed thus: in the first place to consider the seasons of the year. Then the winds ... In the same manner, when one comes into a city in which he is a stranger, he should consider its situation, the water which the inhabitants use … and the mode in which the inhabitants live, and what are their pursuits." Now medical teachers advise whoever wants to investigate medicine properly to study molecular biology, perhaps forgetting in the process that these molecules and cells make up a human being

with a social life of her or his own. Machines now stand between doctors and patients. With the obsession with the "technology fix", the humane physician may be in danger of becoming one day an endangered species (Fathalla, 2000).

14.5 Assessment and evaluation by investors in research

Investors in health research expect a return on their investment. It is inevitable that the unpredictable nature of much scientific research should invite questions about value for money. A commitment to evaluation and accountability on the part of the scientific community is fundamental if science is not to be marginalized in the public and political agendas. Research is an investment.

Three approaches can be pursued and are being used to evaluate the return on the investment in research: impact on advancement of science, impact on health promotion, and impact on wealth creation.

Impact on the advancement of science

Investment in research may be evaluated on the basis of the quantity and quality of the scientific output. These are the criteria commonly used for the evaluation of researchers and scientific institutions. Governments, on the basis of such measures, may allocate funding. Computers now allow bibliometric analysis to provide measurement of publication outputs. Scientific quality is generally based on originality of the subject, thought and method. Quantitatively, it may be measured as the contribution to the advancement of science, reflected on the number of times a paper has been cited as a reference by subsequent authors. This information is readily available from the Science Citation Index (SCI), produced by the Institute of Scientific Information (ISI) (www.isinet.com/isi/products/citation/sci/). The journal in which the paper has been published also matters. Journals are assigned "impact factors". The impact factor measures the frequency with which the "average article" in a journal has been cited in a particular year or period. It provides a way to judge the prestige and influence of a particular journal.

One of the primary objectives of research is to advance science. Science is advanced step by step, through the research efforts of successive investigators. From this perspective in the scientific community, the impact of research is not only about how widely it is disseminated and read; the impact is also about how much it contributes to the advancement of science by being used in subsequent work of other researchers.

Scientific journals are not ranked by scientists according to their circulation but by their impact factors. The impact factor for a journal is calculated by the Science Citation Index (Institute of Scientific Information www.isinet.com). Journal Citation Reports

calculate the number of times that articles from the journal have been cited during the previous two years divided by the total number of articles published by the journal during this period. The impact factor gives a clue to its relative intellectual influence. Some Journals with high impact factors have relatively small circulation. For example, the journal *Nature* has a circulation of about 30 000 and an estimated impact factor of 25; the *Journal of the American Medical Association* has a circulation of about 370 000 and an impact factor less than 7 (Byrne, 1998). The contribution of a scientist to the advancement of science is measured not by the number of publications, but by the impact of these publications. The impact of the publications is assessed indirectly by the impact factors of the journals in which they were published and by subsequent citation of the articles by other authors. Citation analysis tells us that between a third and one half of published papers are never cited even once in subsequent reference lists (Lock, 1984). Many articles are hardly read at all.

Too much emphasis has been put on impact factors, and this emphasis has several drawbacks (Seglen, 1997). The impact may be technically unrelated to the scientific quality of the publication. It should also be noted that citation impact increases as one moves from clinical to basic research (Dawson et al, 1998). Assessment of the impact factor does not do justice to areas of research directly applicable to improvement of health.

Impact on health promotion

The main aim of health research is to improve the health of the people. Scientific quality and impact on health do not always go together. Much research that scientists may judge to be of high quality has no measurable impact on health, often because there may be decades before it has an impact. In contrast, research that may not be judged as high quality by scientists, because of its lack of glamour, may have immediate health benefits, if it has important health policy implications. Evaluation of the investment in research, in terms of impact on health promotion, is not easy. However, this is not a reason for not doing it, with the application of qualitative as well as quantitative methodologies. It is needed and it is necessary for public and not-for-profit private investors in research.

In the evaluation of the impact of research on health promotion, there is an economic return, which should not be undervalued. Human lives are saved, and a human life has monetary worth, in its impact on economic productivity. Health is wealth. What may not be generally appreciated is that there are savings for the health service by using appropriate technologies and discarding ineffective procedures or interventions, and rational allocation of resources. Expenditure on research by the UK National Health Service (NHS) has been estimated to be more than 400 million pounds sterling every year (Wellcome Trust, 2000). In justifying a relatively high level of expenditure on health research, the NHS affirmed the truism that publicly funded research is as important in

the NHS to enable managers to save money, as it is in industry for making money on new products and services (Royal College of Pathologists, 1996).

Mary Lasker, a well-known philanthropist who played a central role in the rapid expansion of medical research and public health in the USA, has been quoted as saying "If you think research is expensive, try disease". In 1999, the Lasker Foundation, through its Funding First initiative, asked nine academic economists from the universities of Chicago, Columbia, Harvard, Stanford and Yale to focus on the economic value of the increase in life expectancy and the impressive decline in mortality. The report "Exceptional returns: the economic value of America's investment in medical research" (http://www.laskerfoundation.org/reports/pdf/exceptional.pdf) estimated the increase in life expectancy in the United States between 1970 and 1990 to be worth roughly US$ 2.8 trillion a year. Reduced mortality from cardiovascular disease alone was estimated to be worth US$ 1.5 trillion a year. Even if only a small percentage of this gain is attributed to advances in research, the return on the research investment would be enormous.

There are also cost savings to the health service, as a result of properly conducted health research. Cost savings include money saved from hospitalization avoided, and from production work gained, from medical procedures not required. For example, preventing hip fractures in postmenopausal women at risk of osteoporosis can save hundreds of millions of dollars annually in treatment costs, apart from loss of productivity. One study in the USA indicated that for every dollar invested throughout the public and private sectors, there was a return of at least three to one from cost savings alone (Rosenberg, 2002).

Impact on wealth creation

Health research may be viewed as an engine for economic growth in developed and also recently in some developing countries. The health industry is one of the fastest growing industries, and one of the most profitable. It has been estimated that companies in the health care market place contribute about 5% of the gross development product in the UK, and generate a trade surplus of some 2 billion pounds sterling (Royal College of Pathologists, 1996). Job creation in the private sector is another parameter. It has been estimated that there are more than 500 000 people employed in the US biopharmaceutical industry because of commitments to research and development (Rosenberg, 2002). These high-paying employment opportunities would not have existed if government was not priming the scientific pump by supporting research.

Governments encourage and support basic research that can provide promising leads for discovery, innovation and wealth creation. For impacts on wealth creation, patent citation indicators have been used to evaluate the investment in research. US patents cite papers as "prior art", that is, the research that has formed the basis for the development of a new and novel product. The Wellcome trust, for example, maintains TechTrac, an

in-house database to link publications in the UK with the US patent prior art information (Dawson et al., 1998).

The importance of health research for development has received increasing international attention over the past 10–20 years. In October 2000, an International Conference on Health Research for Development was convened in Bangkok, co-sponsored by the Council on Health Research for Development, the Global Forum for Health Research, the World Bank and the World Health Organization. The Conference issued a declaration (Annex 5). A ministerial summit on health research is planned by WHO for November 2004 in Mexico.

References and additional sources of information

Reading the medical literature. Applying evidence to practice. Washington, DC, American College of Obstetricians and Gynaecologists, 1998.

Bero L, Rennie D. The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care. *Journal of the American Medical Association*, 1995, 274: 1935–8.

Byrne DW. *Publishing your medical research paper*. Baltimore, Williams & Wilkins, 1998: 52.

Dawson G, Lucocq B, Cottrell R, Lewinson G. *Mapping the landscape. National biomedical research outputs 1988–95. Policy report No. 9*. London, The Wellcome Trust, 1998: 47.

de Vet HCW, Knipschild PG, Schouten HJA, et al. Interobserver variation in histopathological grading of cervical dysplasia. *Journal of Clininical Epidemiology*, 1990, 43: 1395–1398.

Fathalla MF. Appropriate technology in reproductive health. *Egyptian Fertility and Sterility Society Journal*, 2003, 7: 37–41.

Fathalla MF. When medicine rediscovered its social roots. *Bulletin of the World Health Organization*, 2000, 78: 677–678.

Farquhar CM. Evidence-based medicine and getting research into practice. In: O'Brien PMS, Pipkin FB, eds. *Introduction to research methodology for specialists and trainees*. London, Royal College of Obstetricians and Gynaecologists Press, 1999: 84–93.

Friedland DJ, ed. *Evidence-based medicine: A framework for clinical practice*. Stamford, Appleton & Lange, 1998.

Gehlbach SH. *Interpreting the medical literature*. 3rd edition. New York, McGraw-Hill Inc., 1993.

Greenhalgh T. *How to read a paper: the basics of evidence-based medicine*. London, BMJ Books, 1997: 53.

Lock S. Foreword. In: Hawkins C, Sorgi M, eds. *Research: how to plan, speak and write about it*. Berlin, Springer-Verlag, 1985: vii.

McAlister FA, Straus SE, Guyatt GH, Haynes RB, for the Evidence-based Medicine Working Group. Users' guides to the medical literature XX. Integrating research evidence with the care of the individual patient. *Journal of the American Medical Association*, 2000, 283: 2829–2836.

Rosenberg LE. Exceptional economic returns on investments in medical research. *Medical Journal of Australia*, 2002, 177: 368–371.

Sustaining the strength of the UK in healthcare and life sciences R&D: Competition, cooperation and cultural change. A report on the third SmithKline Beecham R&D Policy Symposium. Oxted, ScienceBridge, 1996: 12.

Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-based medicine: how to practice and teach EBM*, 2nd edition. Edinburgh, New York, Churchill Livingstone, 2000.

Seglen P.O. Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 1997, 314: 497.

U.S.Preventive Services Task Force. *Guide to clinical preventive services: An assessment of the effectiveness of 169 interventions*. Baltimore, Williams & Wilkins, 1989.

Wellcome Trust and National Health Service Executive. *Putting NHS research on the map: an analysis of scientific publications in England 1990–1997*. London, The Wellcome Trust, 2001. (http://www.wellcome.ac.uk. accessed 30/3/2004).

World Health Organization Department of Reproductive Health and Research, 2002. *The WHO Reproductive Health Library*, No. 5 [CD-ROM]. Geneva (WHO/RHR/02.1).

Annex 1

World Medical Association Declaration of Helsinki Ethical principles for medical research involving human subjects

Adopted by the 18th World Medical Association (WMA) General Assembly,
Helsinki, Finland, June 1964;
and amended by the
29th WMA General Assembly, Tokyo, Japan, October 1975
35th WMA General Assembly, Venice, Italy, October 1983
41st WMA General Assembly, Hong Kong, September 1989
48th WMA General Assembly, Somerset West, Republic of South Africa,
October 1996
and the
52nd WMA General Assembly, Edinburgh, Scotland, October 2000.
Note of clarification on paragraph 29 added by the WMA General Assembly,
Washington 2002

A. Introduction

1. The World Medical Association has developed the Declaration of Helsinki as a statement of ethical principles to provide guidance to physicians and other participants in medical research involving human subjects. Medical research involving human subjects includes research on identifiable human material or identifiable data.

2. It is the duty of the physician to promote and safeguard the health of the people. The physician's knowledge and conscience are dedicated to the fulfilment of this duty.

3. The Declaration of Geneva of the World Medical Association binds the physician with the words, "The health of my patient will be my first consideration," and the International Code of Medical Ethics declares that, "A physician shall act only in the patient's interest when providing medical care which might have the effect of weakening the physical and mental condition of the patient."

4. Medical progress is based on research which ultimately must rest in part on experimentation involving human subjects.

5. In medical research on human subjects, considerations related to the well-being of the human subject should take precedence over the interests of science and society.
6. The primary purpose of medical research involving human subjects is to improve prophylactic, diagnostic and therapeutic procedures and the understanding of the aetiology and pathogenesis of disease. Even the best proven prophylactic, diagnostic, and therapeutic methods must continuously be challenged through research for their effectiveness, efficiency, accessibility and quality.
7. In current medical practice and in medical research, most prophylactic, diagnostic and therapeutic procedures involve risks and burdens.
8. Medical research is subject to ethical standards that promote respect for all human beings and protect their health and rights. Some research populations are vulnerable and need special protection. The particular needs of the economically and medically disadvantaged must be recognized. Special attention is also required for those who cannot give or refuse consent for themselves, for those who may be subject to giving consent under duress, for those who will not benefit personally from the research and for those for whom the research is combined with care.
9. Research Investigators should be aware of the ethical, legal and regulatory requirements for research on human subjects in their own countries as well as applicable international requirements. No national ethical, legal or regulatory requirement should be allowed to reduce or eliminate any of the protections for human subjects set forth in this Declaration.

B. Basic principles for all medical research

10. It is the duty of the physician in medical research to protect the life, health, privacy, and dignity of the human subject.
11. Medical research involving human subjects must conform to generally accepted scientific principles, be based on a thorough knowledge of the scientific literature, other relevant sources of information, and on adequate laboratory and, where appropriate, animal experimentation.
12. Appropriate caution must be exercised in the conduct of research which may affect the environment, and the welfare of animals used for research must be respected.
13. The design and performance of each experimental procedure involving human subjects should be clearly formulated in an experimental protocol. This protocol should be submitted for consideration, comment, guidance, and where appropriate, approval to a specially appointed ethical review committee, which must be independent of the investigator, the sponsor or any other kind of undue influence. This independent

committee should be in conformity with the laws and regulations of the country in which the research experiment is performed. The committee has the right to monitor ongoing trials. The researcher has the obligation to provide monitoring information to the committee, especially any serious adverse events. The researcher should also submit to the committee, for review, information regarding funding, sponsors, institutional affiliations, other potential conflicts of interest and incentives for subjects.

14. The research protocol should always contain a statement of the ethical considerations involved and should indicate that there is compliance with the principles enunciated in this Declaration.
15. Medical research involving human subjects should be conducted only by scientifically qualified persons and under the supervision of a clinically competent medical person. The responsibility for the human subject must always rest with a medically qualified person and never rest on the subject of the research, even though the subject has given consent.
16. Every medical research project involving human subjects should be preceded by careful assessment of predictable risks and burdens in comparison with foreseeable benefits to the subject or to others. This does not preclude the participation of healthy volunteers in medical research. The design of all studies should be publicly available.
17. Physicians should abstain from engaging in research projects involving human subjects unless they are confident that the risks involved have been adequately assessed and can be satisfactorily managed. Physicians should cease any investigation if the risks are found to outweigh the potential benefits or if there is conclusive proof of positive and beneficial results.
18. Medical research involving human subjects should only be conducted if the importance of the objective outweighs the inherent risks and burdens to the subject. This is especially important when the human subjects are healthy volunteers.
19. Medical research is only justified if there is a reasonable likelihood that the populations in which the research is carried out stand to benefit from the results of the research.
20. The subjects must be volunteers and informed participants in the research project.
21. The right of research subjects to safeguard their integrity must always be respected. Every precaution should be taken to respect the privacy of the subject, the confidentiality of the patient's information and to minimize the impact of the study on the subject's physical and mental integrity and on the personality of the subject.

22. In any research on human beings, each potential subject must be adequately informed of the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliations of the researcher, the anticipated benefits and potential risks of the study and the discomfort it may entail. The subject should be informed of the right to abstain from participation in the study or to withdraw consent to participate at any time without reprisal. After ensuring that the subject has understood the information, the physician should then obtain the subject's freely-given informed consent, preferably in writing. If the consent cannot be obtained in writing, the non-written consent must be formally documented and witnessed.
23. When obtaining informed consent for the research project the physician should be particularly cautious if the subject is in a dependent relationship with the physician or may consent under duress. In that case the informed consent should be obtained by a well-informed physician who is not engaged in the investigation and who is completely independent of this relationship.
24. For a research subject who is legally incompetent, physically or mentally incapable of giving consent or is a legally incompetent minor, the investigator must obtain informed consent from the legally authorized representative in accordance with applicable law. These groups should not be included in research unless the research is necessary to promote the health of the population represented and this research cannot instead be performed on legally competent persons.
25. When a subject deemed legally incompetent, such as a minor child, is able to give assent to decisions about participation in research, the investigator must obtain that assent in addition to the consent of the legally authorized representative.
26. Research on individuals from whom it is not possible to obtain consent, including proxy or advance consent, should be done only if the physical/mental condition that prevents obtaining informed consent is a necessary characteristic of the research population. The specific reasons for involving research subjects with a condition that renders them unable to give informed consent should be stated in the experimental protocol for consideration and approval of the review committee. The protocol should state that consent to remain in the research should be obtained as soon as possible from the individual or a legally authorized surrogate.
27. Both authors and publishers have ethical obligations. In publication of the results of research, the investigators are obliged to preserve the accuracy of the results. Negative as well as positive results should be published or otherwise publicly available. Sources of funding, institutional affiliations and any possible conflicts of interest should be declared in the publication. Reports of experimentation not in accordance with the principles laid down in this Declaration should not be accepted for publication.

C. Additional principles for medical research combined with medical care

28. The physician may combine medical research with medical care, only to the extent that the research is justified by its potential prophylactic, diagnostic or therapeutic value. When medical research is combined with medical care, additional standards apply to protect the patients who are research subjects.

29. The benefits, risks, burdens and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods. This does not exclude the use of placebo, or no treatment, in studies where no proven prophylactic, diagnostic or therapeutic method exists. (See footnote.)

30. At the conclusion of the study, every patient entered into the study should be assured of access to the best proven prophylactic, diagnostic and therapeutic methods identified by the study.

31. The physician should fully inform the patient which aspects of the care are related to the research. The refusal of a patient to participate in a study must never interfere with the patient-physician relationship.

32. In the treatment of a patient, where proven prophylactic, diagnostic and therapeutic methods do not exist or have been ineffective, the physician, with informed consent from the patient, must be free to use unproven or new prophylactic, diagnostic and therapeutic measures, if in the physician's judgement it offers hope of saving life, re-establishing health, or alleviating suffering. Where possible, these measures should be made the object of research, designed to evaluate their safety and efficacy. In all cases, new information should be recorded and, where appropriate, published. The other relevant guidelines of this Declaration should be followed.

Footnote: Note of clarification on paragraph 29 of the WMA Declaration of Helsinki

The WMA hereby reaffirms its position that extreme care must be taken in making use of a placebo-controlled trial and that in general this methodology should only be used in the absence of existing proven therapy. However, a placebo-controlled trial may be ethically acceptable, even if proven therapy is available, under the following circumstances:

- Where for compelling and scientifically sound methodological reasons its use is necessary to determine the efficacy or safety of a prophylactic, diagnostic or therapeutic method; or

- Where a prophylactic, diagnostic or therapeutic method is being investigated for a minor condition and the patients who receive placebo will not be subject to any additional risk of serious or irreversible harm.

All other provisions of the Declaration of Helsinki must be adhered to, especially the need for appropriate ethical and scientific review.

Annex 2

International ethical guidelines for biomedical research involving human subjects

Prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO). CIOMS, Geneva 2002. The text of the guidelines reproduced here does not include the commentary provided in the full document. This can be found at http:\\www.cioms.ch\frame_ guidelines _nov_2002.htm

Guideline 1: Ethical justification and scientific validity of biomedical research involving human beings

The ethical justification of biomedical research involving human subjects is the prospect of discovering new ways of benefiting people's health. Such research can be ethically justifiable only if it is carried out in ways that respect and protect, and are fair to, the subjects of that research and are morally acceptable within the communities in which the research is carried out. Moreover, because scientifically invalid research is unethical in that it exposes research subjects to risks without possible benefit, investigators and sponsors must ensure that proposed studies involving human subjects conform to generally accepted scientific principles and are based on adequate knowledge of the pertinent scientific literature.

Guideline 2: Ethical review committees

All proposals to conduct research involving human subjects must be submitted for review of their scientific merit and ethical acceptability to one or more scientific review and ethical review committees. The review committees must be independent of the research team, and any direct financial or other material benefit they may derive from the research should not be contingent on the outcome of their review. The investigator must obtain their approval or clearance before undertaking the research. The ethical review committee should conduct further reviews as necessary in the course of the research, including monitoring of the progress of the study.

Guideline 3: Ethical review of externally sponsored research

An external sponsoring organization and individual investigators should submit the research protocol for ethical and scientific review in the country of the sponsoring organization, and the ethical standards applied should be no less stringent than they would be for research carried out in that country. The health authorities of the host country, as well as a national or local ethical review committee, should ensure that the proposed research is responsive to the health needs and priorities of the host country and meets the requisite ethical standards.

Guideline 4: Individual informed consent

For all biomedical research involving humans the investigator must obtain the voluntary informed consent of the prospective subject or, in the case of an individual who is not capable of giving informed consent, the permission of a legally authorized representative in accordance with applicable law. Waiver of informed consent is to be regarded as uncommon and exceptional, and must in all cases be approved by an ethical review committee.

Guideline 5: Obtaining informed consent: Essential information for prospective research subjects

Before requesting an individual's consent to participate in research, the investigator must provide the following information, in language or another form of communication that the individual can understand:

1. that the individual is invited to participate in research, the reasons for considering the individual suitable for the research, and that participation is voluntary;
2. that the individual is free to refuse to participate and will be free to withdraw from the research at any time without penalty or loss of benefits to which he or she would otherwise be entitled;
3. the purpose of the research, the procedures to be carried out by the investigator and the subject, and an explanation of how the research differs from routine medical care;
4. for controlled trials, an explanation of features of the research design (e.g. randomization, double-blinding), and that the subject will not be told of the assigned treatment until the study has been completed and the blind has been broken;

5. the expected duration of the individual's participation (including number and duration of visits to the research centre and the total time involved) and the possibility of early termination of the trial or of the individual's participation in it;
6. whether money or other forms of material goods will be provided in return for the individual's participation and, if so, the kind and amount;
7. that, after completion of the study, subjects will be informed of the findings of the research in general, and individual subjects will be informed of any finding that relates to their particular health status;
8. that subjects have the right of access to their data on demand, even if these data lack immediate clinical utility (unless the ethical review committee has approved temporary or permanent non-disclosure of data, in which case the subject should be informed of, and given, the reasons for such non-disclosure);
9. any foreseeable risks, pain or discomfort, or inconvenience to the individual (or others) associated with participation in the research, including risks to the health or well-being of a subject's spouse or partner;
10. the direct benefits, if any, expected to result to subjects from participating in the research;
11. the expected benefits of the research to the community or to society at large, or contributions to scientific knowledge;
12. whether, when and how any products or interventions proven by the research to be safe and effective will be made available to subjects after they have completed their participation in the research, and whether they will be expected to pay for them;
13. any currently available alternative interventions or courses of treatment;
14. the provisions that will be made to ensure respect for the privacy of subjects and for the confidentiality of records in which subjects are identified;
15. the limits, legal or other, to the investigators' ability to safeguard confidentiality, and the possible consequences of breaches of confidentiality;
16. policy with regard to the use of results of genetic tests and familial genetic information, and the precautions in place to prevent disclosure of the results of a subject's genetic tests to immediate family relatives or to others (e.g., insurance companies or employers) without the consent of the subject;
17. the sponsors of the research, the institutional affiliation of the investigators, and the nature and sources of funding for the research;

18. the possible research uses, direct or secondary, of the subject's medical records and of biological specimens taken in the course of clinical care;
19. whether it is planned that biological specimens collected in the research will be destroyed at its conclusion, and, if not, details about their storage (where, how, for how long, and final disposition) and possible future use, and that subjects have the right to decide about such future use, to refuse storage, and to have the material destroyed;
20. whether commercial products may be developed from biological specimens, and whether the participant will receive monetary or other benefits from the development of such products;
21. whether the investigator is serving only as an investigator or as both investigator and the subject's physician;
22. the extent of the investigator's responsibility to provide medical services to the participant;
23. that treatment will be provided free of charge for specified types of research-related injury or for complications associated with the research, the nature and duration of such care, the name of the organization or individual that will provide the treatment, and whether there is any uncertainty regarding funding of such treatment;
24. in what way, and by what organization, the subject or the subject's family or dependants will be compensated for disability or death resulting from such injury (or, when indicated, that there are no plans to provide such compensation);
25. whether or not, in the country in which the prospective subject is invited to participate in research, the right to compensation is legally guaranteed;
26. that an ethical review committee has approved or cleared the research protocol.

Guideline 6: Obtaining informed consent: Obligations of sponsors and investigators

Sponsors and investigators have a duty to:

- refrain from unjustified deception, undue influence, or intimidation;
- seek consent only after ascertaining that the prospective subject has adequate understanding of the relevant facts and of the consequences of participation and has had sufficient opportunity to consider whether to participate;

- as a general rule, obtain from each prospective subject a signed form as evidence of informed consent—investigators should justify any exceptions to this general rule and obtain the approval of the ethical review committee;
- renew the informed consent of each subject if there are significant changes in the conditions or procedures of the research or if new information becomes available that could affect the willingness of subjects to continue to participate; and,
- renew the informed consent of each subject in long-term studies at pre-determined intervals, even if there are no changes in the design or objectives of the research.

Guideline 7: Inducement to participate

Subjects may be reimbursed for lost earnings, travel costs and other expenses incurred in taking part in a study; they may also receive free medical services. Subjects, particularly those who receive no direct benefit from research, may also be paid or otherwise compensated for inconvenience and time spent. The payments should not be so large, however, or the medical services so extensive as to induce prospective subjects to consent to participate in the research against their better judgement ("undue inducement"). All payments, reimbursements and medical services provided to research subjects must have been approved by an ethical review committee.

Guideline 8: Benefits and risks of study participation

For all biomedical research involving human subjects, the investigator must ensure that potential benefits and risks are reasonably balanced and risks are minimized.

- Interventions or procedures that hold out the prospect of direct diagnostic, therapeutic or preventive benefit for the individual subject must be justified by the expectation that they will be at least as advantageous to the individual subject, in the light of foreseeable risks and benefits, as any available alternative. Risks of such "beneficial" interventions or procedures must be justified in relation to expected benefits to the individual subject.
- Risks of interventions that do not hold out the prospect of direct diagnostic, therapeutic or preventive benefit for the individual must be justified in relation to the expected benefits to society (generalizable knowledge). The risks presented by such interventions must be reasonable in relation to the importance of the knowledge to be gained.

Guideline 9: Special limitations on risk when research involves individuals who are not capable of giving informed consent

When there is ethical and scientific justification to conduct research with individuals incapable of giving informed consent, the risk from research interventions that do not hold out the prospect of direct benefit for the individual subject should be no more likely and not greater than the risk attached to routine medical or psychological examination of such persons. Slight or minor increases above such risk may be permitted when there is an overriding scientific or medical rationale for such increases and when an ethical review committee has approved them.

Guideline 10: Research in populations and communities with limited resources

Before undertaking research in a population or community with limited resources, the sponsor and the investigator must make every effort to ensure that:

- the research is responsive to the health needs and the priorities of the population or community in which it is to be carried out; and
- any intervention or product developed, or knowledge generated, will be made reasonably available for the benefit of that population or community.

Guideline 11: Choice of control in clinical trials

As a general rule, research subjects in the control group of a trial of a diagnostic, therapeutic or preventive intervention should receive an established effective intervention. In some circumstances it may be be ethically acceptable to use an alternative comparator, such as placebo or “no treatment”.

Placebo may be used:

- when there is no established effective intervention;
- when withholding an established effective intervention would expose subjects to, at most, temporary discomfort or delay in relief of symptoms;
- when use of an established effective intervention as comparator would not yield scientifically reliable results and use of placebo would not add any risk of serious or irreversible harm to the subjects.

Guideline 12: Equitable distribution of burdens and benefits in the selection of groups of subjects in research

Groups or communities to be invited to be subjects of research should be selected in such a way that the burdens and benefits of the research will be equitably distributed. The exclusion of groups or communities that might benefit from study participation must be justified.

Guideline 13: Research involving vulnerable persons

Special justification is required for inviting vulnerable individuals to serve as research subjects and, if they are selected, the means of protecting their rights and welfare must be strictly applied.

Guideline 14: Research involving children

Before undertaking research involving children, the investigator must ensure that:

- the research might not equally well be carried out with adults;
- the purpose of the research is to obtain knowledge relevant to the health needs of children;
- a parent or legal representative of each child has given permission;
- the agreement (assent) of each child has been obtained to the extent of the child's capabilities; and,
- a child's refusal to participate or continue in the research will be respected.

Guideline 15: Research involving individuals who by reason of mental or behavioural disorders are not capable of giving adequately informed consent

Before undertaking research involving individuals who by reason of mental or behavioural disorders are not capable of giving adequately informed consent, the investigator must ensure that:

- such persons will not be subjects of research that might equally well be carried out on persons whose capacity to give adequately informed consent is not impaired;
- the purpose of the research is to obtain knowledge relevant to the particular health needs of persons with mental or behavioural disorders;

- the consent of each subject has been obtained to the extent of that person's capabilities, and a prospective subject's refusal to participate in research is always respected, unless, in exceptional circumstances, there is no reasonable medical alternative and local law permits overriding the objection; and,
- in cases where prospective subjects lack capacity to consent, permission is obtained from a responsible family member or a legally authorized representative in accordance with applicable law.

Guideline 16: Women as research subjects

Investigators, sponsors or ethical review committees should not exclude women of reproductive age from biomedical research. The potential for becoming pregnant during a study should not, in itself, be used as a reason for precluding or limiting participation. However, a thorough discussion of risks to the pregnant woman and to her fetus is a prerequisite for the woman's ability to make a rational decision to enrol in a clinical study. In this discussion, if participation in the research might be hazardous to a fetus or a woman if she becomes pregnant, the sponsors/ investigators should guarantee the prospective subject a pregnancy test and access to effective contraceptive methods before the research commences. Where such access is not possible, for legal or religious reasons, investigators should not recruit for such possibly hazardous research women who might become pregnant.

Guideline 17: Pregnant women as research participants

Pregnant women should be presumed to be eligible for participation in biomedical research. Investigators and ethical review committees should ensure that prospective subjects who are pregnant are adequately informed about the risks and benefits to themselves, their pregnancies, the fetus and their subsequent offspring, and to their fertility.

Research in this population should be performed only if it is relevant to the particular health needs of a pregnant woman or her fetus, or to the health needs of pregnant women in general, and, when appropriate, if it is supported by reliable evidence from animal experiments, particularly as to risks of teratogenicity and mutagenicity.

Guideline 18:Safeguarding confidentiality

The investigator must establish secure safeguards of the confidentiality of subjects' research data. Subjects should be told the limits, legal or other, to the investigators'

ability to safeguard confidentiality and the possible consequences of breaches of confidentiality.

Guideline 19: Right of injured subjects to treatment and compensation

Investigators should ensure that research subjects who suffer injury as a result of their participation are entitled to free medical treatment for such injury and to such financial or other assistance as would compensate them equitably for any resultant impairment, disability or handicap. In the case of death as a result of their participation, their dependants are entitled to compensation. Subjects must not be asked to waive the right to compensation.

Guideline 20: Strengthening capacity for ethical and scientific review and biomedical research

Many countries lack the capacity to assess or ensure the scientific quality or ethical acceptability of biomedical research proposed or carried out in their jurisdictions. In externally sponsored collaborative research, sponsors and investigators have an ethical obligation to ensure that biomedical research projects for which they are responsible in such countries contribute effectively to national or local capacity to design and conduct biomedical research, and to provide scientific and ethical review and monitoring of such research.

Capacity-building may include, but is not limited to, the following activities:

- establishing and strengthening independent and competent ethical review processes/ committees
- strengthening research capacity
- developing technologies appropriate to health-care and biomedical research
- training of research and health-care staff
- educating the community from which research subjects will be drawn.

Guideline 21: Ethical obligation of external sponsors to provide health-care services

External sponsors are ethically obliged to ensure the availability of:

- health-care services that are essential to the safe conduct of the research;
- treatment for subjects who suffer injury as a consequence of research interventions; and,
- services that are a necessary part of the commitment of a sponsor to make a beneficial intervention or product developed as a result of the research reasonably available to the population or community concerned.

Annex 3

Searching the literature

1. The US National Library of Medicine

The US National Library of Medicine (NLM) is the largest library in the world. It has been indexing the biomedical literature since 1879 to help provide health professionals access to information necessary for research, health care, and education. It makes available to researchers a vast database that is updated and changed frequently.

MEDLINE

MEDLINE is NLM's premier bibliographic database covering the fields of medicine, nursing, dentistry, veterinary medicine, and the pre-clinical sciences. MEDLINE has been available for online searching since 1971. It has practically replaced the familiar bulky volumes of printed index medicus, the monthly subject/author guide to biomedical literature, formerly painstakingly hand-searched. Advantages of online search include: speed of retrieval; access to more journals; availability of abstracts for many references; searching by non-MeSH (Medical Subject Headings) terms in titles and abstracts; and short lag period since publication.

Free MEDLINE searching has been introduced since 1997. MEDLINE can be accessed through the NLM web site: http://www.nlm.nih.gov. No registration is required for access.

Journal articles are indexed for MEDLINE, and their citations are searchable. MEDLINE includes articles from more than 4600 international biomedical journals. Coverage is world-wide, of journals published in the United States and 70 other countries, but most records (86%) are from English language sources or have English abstracts. The file contains over 12 million records dating back to 1966. It is updated weekly; with about 40 000 new citations added each month.

PubMed

In addition to providing access to MEDLINE, PubMed provides access to other citations, and to citations that precede the date that a journal was selected for MEDLINE indexing. It also includes access to PubMedCentral.

PubMed Central

PubMed Central is an archive of life sciences journals. As of October 2003, it provides full text of over 100 000 articles from over 130 journals, and the number is increasing. It is linked to PubMed and is fully searchable. Access is free and no registration is required.

MeSH (Medical Subject Headings)

MeSH is NLM's controlled vocabulary used for indexing articles in PubMed. MeSH terminology provides a consistent way to retrieve information that may use different terminology for the same topic. MeSH terms are similar to the keywords used in other web searches. All MeSH terms (over 19 000) are arranged alphabetically as well as in subject groups. Within groups, MeSH terms are arranged in hierarchical levels known as "tree structures", in which all the terms are arranged from the most general to the most specific. The specificity rule is that papers are always indexed under the most specific MeSH headings available, and not under general all-embracing terms. NLM indexers examine articles and assign the most specific MeSH to describe it. The indexers will assign as many MeSH headings as appropriate to cover the topic of the article (generally 5–15). MeSH headings are constantly under review and new headings are regularly introduced. The MeSH database can be searched from the MEDLINE/ PubMed web page. Suggested MeSH titles will be provided for any search term entered by the searcher. Alternatively, the researcher can navigate the MeSH tree from the top down.

MEDLINE search

Search can be made by author, topic or journal title.

Logical (or "Boolean") operators AND, OR, AND NOT can be used to narrow the search, as in other web search, discussed below.

Limits can be set for the search. For example, the search may be limited by language, type of publication, date, and to whether an abstract is available.

Citations include the English abstract when published with the article (approximately 76% of the most current 5 years). Hyperlinks to related articles are also available.

2. Searching the internet

The internet and the World Wide Web

The internet is now a major source of research information. It offers instant access to millions of computer files relating to almost any topic. For easy access to this huge network, the World Wide Web (WWW) is used. It is a set of files connected by hypertext links and accessed by means of a browser, such as Microsoft Explorer or Netscape navigator. The WWW thus provides an interface to various forms of information on the Internet, The world wide web has been described as the universe of network-accessible information, the embodiment of human knowledge.

Sites on the internet have addresses, called Uniform Resource Locators (URL), written in a uniform style. An example of a URL may be: http://www.georgetown.edu/home/libraries.html

- "http" stands for "hypertext transfer protocol" which is used to transmit the data.
- "www" stands for the World Wide Web, the global Internet service that connects the multitude of computers and the Internet files.
- "georgetown.edu" is called the "domain" which names the organization feeding the information, in this case Georgetown University. The suffix in the domain indicates the type of organization, for example: .edu (educational); .com (commercial); .gov (government); .org (organization); .net (network organization).
- "home/libraries" represents the homepage of the website and the file to be searched for.
- "html": is the hypertext markup language, which is the computer language used to write the file.

Search engines

Several companies, buttressed by advertising on their sites, have created methods of instantaneously searching the content of every publicly accessible web site. These search engines are important because new web sites are continually added, and many change their location (URL). There are numerous search engines available. Each search engine operates differently and consequently has different strengths and weaknesses. Rather than a simple list of file names or URLs, many search engines provide a small extract or other information about the file. "Hits" can be ranked in order of relevancy to the search terms requested, calculated by the frequency with which the terms appear in a document, their proximity to one another, or their relative position in a web page. The

web browser usually has a list of search engines, and each has information about itself. All have help pages explaining the various search options.

Some search engines are human-driven (information compiled and indexed by people). Others are robot-driven. Meta-search engines will search for the topic in several of the search engines.

Search engines usually offer subject directories and keyword searches. Hypertext links will allow moving from a certain site to another. The indication that a hypertext link is available is shown when a text is underlined and/or coloured, and when the browser pointer changes its shape.

Search using subject directories

A subject directory takes the searcher through a sequence of topics. Subject-tree directories are hierarchical, moving from broader to narrower topics, and the general to the more specific. Several directories have entire organized sections of medicine- related information, often dividing resources by clinical specialty or general health and medical subject headings. Directories characteristically lend themselves to "casual browsing". The biggest advantage of manually created directories is the ability to include an annotation describing the resource—although not all directories choose to do this.

Using a keyword search

Using a keyword search needs to be carefully optimized. Selection of the keywords in the search process is similar to the selection of a screening test for diseases. One can select a test that is highly sensitive but will pick many false positives. Or one can select a test that is very specific but will miss many false negatives. Professional searchers refer to sensitivity by the term "recall ratio" and to specificity by "precision ratio". One keyword selection may be highly sensitive but low in specificity; it will pick up all sites relevant to the subject but also too many irrelevant sites. Another set can be insensitive and miss many relevant sites while being highly specific, picking only highly relevant sites.

Boolean logic

Some software packages allow the use of linkage of search terms to raise the specificity of references retrieved. More than one term can be used in the search, linked by logical (or "Boolean") operators AND, OR, AND NOT. They are named after Charles Boole, a British logician and mathematician of the 19th century. They indicate to the computer how you want the terms treated in relation to each other during the search.

AND: Both terms must be found in each reference

OR: Either term must be present in each reference

NOT: References containing the term are excluded

Some search engines allow also the use of means to narrow or facilitate the search.

"Proximity operators": SAME and WITH indicate how closely the terms they connect must be linked in a reference.

If a phrase is used as the keyword, rather than words, it should be limited by quotation marks.

Wildcards (truncation): When searching for a word, the search system looks for an exact match unless the word ends with an asterisk (*). The asterisk symbol acts as a "wildcard" and matches all words that begin with the string of characters before the asterisk. For example, entering arter* will retrieve documents containing words such as artery, arteries, arterial, etc.

Usually case (upper or lower) is ignored in keywords.

Health information on the web

Health information is often said to be one of the most retrieved types of information on the web. Because the Internet is an unregulated, constantly changing set of computers around the world, the quality of information varies substantially from site to site. Nonetheless, there are "official" web pages developed by organizations that do have a reputation at stake, including pages from high-quality, peer-reviewed journals, government institutions, and many educational institutions. These sites are the real backbone of health information on the web, but there are also commercial medical information sites, in competition with each other, that are more aggressively and regularly updated.

3. Free access to medical journals on the internet

"Open Access now", a newsletter campaigning for freedom of research information and published by BioMedCentral, provides information to researchers in the life sciences about organizations involved in Open Access publishing, and links to their web sites.

(http://www.biomedcentral.com/openaccess/contact.asp)

Open Access links

- **Open Access news:** http://www.earlham.edu/~peters/fos/ This site provides news and discussions on open access to research literature.
- **PubMed Central:** http://www.pubmedcentral.nih.gov A digital archive of life sciences journal literature with free and unrestricted access.
- **Directory of open access journal (Lund University):** http://www.doaj.org/ This service covers free, full text, quality controlled scientific and scholarly journals. The Directory aims to include open access journals in all subjects and languages.
- **Budapest Open Access Initiative:** http://www.soros.org/openaccess/index.shtml This initiative was created in connection with the Soros Foundation at a meeting in Budapest in December 2001 to accelerate progress in the international effort to make research articles in all academic fields freely available on the Internet.
- **Public Library of Science:** http://www.publiclibraryofscience.org/ A non-profit organization of scientists committed to making the world's scientific and medical literature freely accessible to scientists and to the public around the world.
- **SPARC:** http://www.arl.org/sparc/ SPARC is an alliance of universities, research libraries, and organizations built as a constructive response to market dysfunctions in the scholarly communication system. These dysfunctions have reduced dissemination of scholarship and crippled libraries. SPARC serves as a catalyst for action, helping to create systems that expand information dissemination and use in a networked digital environment while responding to the needs of scholars and academe.
- **SciELO:** http://www.scielo.br/ The scientific Electronic Library Online- SciELO is an electronic library covering a selected collection of Brazilian scientific journals.
- **Health InterNetwork:** http://www.healthinternetwork.org/scipub.php The Health InterNetwork was launched by the Secretary General of the United Nations and is led by the World Health Organization to bridge the "digital divide" in health. It aims to ensure that health information and the technologies to deliver it are widely available and effectively used by health personnel professionals, researchers, scientists, and policy makers.
- **FreeMedicalJournals.com:** http://www.freemedicaljournals.com/ Dedicated to the promotion of free access to medical journals over the Internet, the site carries listings of free full-text journals.

4. Searching the Health InterNetwork Access to Research Initiative (HINARI)

(http://www.healthinternetwork.org/scipub.php?lang=en Accessed 24/2/2004)

Background information on HINARI has been provided in Chapter 3, section 6.2.

Using journals through HINARI

Journal abstracts are available to all users, without registration. Registered HINARI users have full-text access to journal. Non-HINARI users may still have full-text access through the HINAR LOG IN menu, if their institutions subscribe to the journals.

Finding journals

From the menu page, you can find journals alphabetically by title or by subject. To find journals by title, click on a letter for an alphabetical list of journal titles starting with that letter, then select a journal title to go directly to that journal.

Finding articles

Detailed searching for articles can be done through Pubmed (Medline database) or by visiting the web sites of individual publishers. To search through Pubmed, users can click on "Search for articles through Pubmed (Medline)", and then have two options:

- Retrieve citations to all articles about a subject. Some of these articles may not be online, or may not be accessible to HINARI users.
- Retrieve citations only to articles that are accessible through HINARI.

Indexes to regional journals

The menu page provides links to the following:

- African Index Medicus (AIM)
- Index Medicus for the WHO Eastern Mediterranean Region (IMEMR)
- Latin American and Caribbean Center on Health Sciences Information (LILACS)
- Index Medicus for South-East Asia Region (IMSEAR).

The user can click for a shortcut to index information.

Reference sources full text

Links are provided to databases, encyclopedias, books and other full text resources.

Links to other free collections

These sites offer free access to journal collections:

- BioMed Central
- Free Medical Journal
- Free books for doctors
- PubMed Central
- SciELO

HINARI registration

Academic, government or research institutions located in one of the countries eligible for access to HINARI can register by completing a registration form. Once the registration form is received, a common username and password will be issued for all staff at the institution. It is suggested that the institution's librarian be the main contact point.

As of February 2004, the following countries in the WHO Eastern Mediterranean Region are eligible for access to HINARI:

Afghanistan, Djibouti, Iraq, Jordan, Morocco, Sudan, Syrian Arab Republic, Tunisia, West Bank and Gaza, Yemen.

5. Searching library resources of the WHO Regional Office for the Eastern Mediterranean (EMRO)

Information on the Eastern Mediterranean Region Index Medicus is provided in Chapter 3, section 3.6.2.

The following CD-ROM Databases are available in the EMRO library (http://www.emro.who.int/Library/LibraryDatabases.HTM Accessed 24/2/2004)

- **CDMARC** Bibliographic Library of Congress
- **The CD-ROM Directory** provides comprehensive details on CD-ROM titles commercially available.
- **Computer-related databases**

- **EMBASE**: **Drug and Pharmaceutical**. This database contains over 1 300 000 abstracts and citations from the last 10 years and covers comprehensively the drugs and pharmacology literature including effect and use of all drugs and potential drugs, clinical and experimental aspects and pharmacokinetics and pharmacodynamics. Also extensively covered are the side effects of and adverse reactions to drugs.
- **ERIC** (Educational Resources Information Centre)
- **ExtraMED** contains the contents of over 220 biomedical journals from all over the world, mainly from developing countries. It was established on the initiative of the World Health Organization. Users will be able to use indexing tools provided on the disk to locate relevant articles and can then print out those of particular interest.
- **Food and Human Nutrition** focuses on subjects from an international perspective, which includes over 135 participating countries covered in over a quarter of a million records.
- **Global Books in Print Plus**. This title contains bibliographical information from six English language databases.
- **LILACS/ CD-ROM** Latin American and Caribbean Health Sciences Literature. The only complete and updated database covering health related literature published in the Latin American and the Caribbean regions.
- **MEDLINE**. The MEDLINE database encompasses information from three printed indexes (Index Medicus, Index to Dental Literature and the International Nursing Index) as well as additional information not published in the Index Medicus.
- **Oxford English Dictionary** (Second Edition) on Compact Disc
- **POPLINE** is a bibliographic database containing more than 150 000 citations on population, family planning and related health care, law, and policy issues.
- **Ulrich's International Periodicals Directory**.

Annex 4

Guidelines on how to write references for scientific papers

1. General

This Annex supplements information provided in Chapter 11, section 11.10, by providing examples on how different types of references can be cited. It is based on the Uniform Requirements for Manuscripts submitted to Biomedical Journals issued by the International Committee of Medical Journal Editors. The Uniform Requirements style (the Vancouver style) is based largely on an ANSI standard style adapted by the NLM for its databases. (http://www.nlm.nih.gov/bsd/uniform_requirements.html). Authors should ensure that they follow any examples of style given by the journal to which they are submitting a paper.

2. Journal articles

Standard journal article

Vega KJ, Pina I, Krevsky B. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Ann Intern Med 1996 Jun 1;124 (11): 980–3.

If the Journal carries continuous pagination throughout a volume (as many medical journals do), the month and issue number may be omitted.

Vega KJ, Pina I, Krevsky B. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Ann Intern Med 1996;124: 980–3.

If more than six authors, list the first six authors followed by et al.

Vega KJ, Pina I, Krevsky B, et al. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Ann Intern Med 1996;124: 980–3.

Organization as an author

The Cardiac Society of Australia and New Zealand. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Ann Intern Med 1996;124: 980–3.

No author given

Heart transplantation is associated with an increased risk for pancreatobiliary disease [editorial]. Ann Intern Med 1996;124: 980–3.

Article not in English

Do not provide translation

Volume with supplement

Vega KJ, Pina I, Krevsky B, et al. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Ann Intern Med 1996;124 Suppl 1:980–3.

Issue with supplement

Vega KJ, Pina I, Krevsky B, et al. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Ann Intern Med 1996;124(1 Suppl 2):980–3.

Volume with part

Vega KJ, Pina I, Krevsky B. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Ann Intern Med 1996;124(Pt 3):980–3.

80–3.

Issue with no volume

Vega KJ, Pina I, Krevsky B. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Clin Ortho 1996;(124): 980–3.

No issue or volume

Vega KJ, Pina I, Krevsky B. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Curr Opin Gen Surg 1996: 980–3.

Pagination in Roman numerals

Vega KJ, Pina I, Krevsky B. Heart transplantation is associated with an increased risk for pancreatobiliary disease. Surg Clin North Am 1996 April;124(2):xi–xii

3. Books and other Monographs

Formal author(s)

Ringsven MK, Bond D. Gerontology and leadership skills for nurses. 2nd ed. Albany (NY): Delmar Publishers; 1996.

Editor(s), compiler(s) as author

Ringsven MK, Bond D, editors. Gerontology and leadership skills for nurses. 2nd ed. Albany (NY): Delmar Publishers; 1996.

Organization as author and publisher

Institute of Medicine (US). Gerontology and leadership skills for nurses. Washington; The Institute; 1996.

Chapter in a book

Phillips SJ, Whisnant JP. An introduction to gerontology. In: Ringsven MK, Bond D, editors. Gerontology and leadership skills for nurses. 2nd ed. Albany (NY): Delmar Publishers; 1996. p. 465–78.

Conference proceedings

Ringsven MK, Bond D, editors. Gerontology and leadership skills for nurses. Proceedings of the 10th International Congress of Nurses; 1996 Oct 15–19; Kyoto, Japan. Albany (NY): Delmar Publishers; 1996.

Conference paper

Phillips SJ, Whisnant JP. An introduction to gerontology. In: Ringsven MK, Bond D, editors. Gerontology and leadership skills for nurses. Proceedings of the 10th International Congress of Nurses; 1996 Oct 15-19; Kyoto, Japan. Albany (NY): Delmar Publishers; 1996. p. 1561–5.

Scientific or technical report issued by funding/sponsoring agency

Smith P, Golladay K. Payment for durable medical equipment in skilled nursing facilities. Final report. Dallas (TX): Department of Health and Human Services (US), Office of Evaluation and Inspection; 1994 Oct. Report No: HHSIGOEI69200860.

4. Unpublished material

In press

(Note: NLM prefers “Forthcoming” because not all items will be printed)

Leshner AI. Molecular mechanisms of cocaine addiction. N Engl J Med, In press 1996.

Electronic material

Journal article in electronic format

Morse SS. Factors in the emergence of infectious diseases. Emerg Infect Dis [serial online] 1995 Jan–Mar [cited 1996 Jun 5]; 1(1):[24 screens]. Available from URL: http://www.cdc.gov/ncidod/EID/eid.htm

Monograph in electronic format

CDI, clinical dermatology illustrated [monograph on CD-Rom]. Reeves JRT, Maibach H. CMEA Multimedia Group, producers. 2nd ed. Version 2.0. San Diego: CMEA; 1995.

5. How to order references

Most journals in medicine and the other medical sciences use the Vancouver, or citation-by-reference number, system in which the references in the reference list are numbered in the order in which they are first cited in the text.

Some journals still use the citation-by-author-and-date system (also known as the Harvard system) in which the paper cited is identified by author name and year of publication.

In a combined Alphabet-Number System, references are listed in alphabetical order according to the primary author's name and cited by numbers in the text.

Even if the journal to which the paper is submitted uses the citation-by-reference number system, it is advisable to use the Harvard system for the citations in the first and other early drafts. If numbers are assigned to references at this early stage, those numbers will very likely have to be changed in subsequent drafts. With a word-processing program, the "search-and-replace" function, one can place at the beginning of each citation a character not used elsewhere in the text, for example an asterisk (*).

Annex 5

Bangkok Declaration on Health Research for Development

The International Conference on Health Research for Development brought together more than 700 participants representing a wide range of stakeholders in health research from developing and developed countries. Conference particpants from over one hundred countries welcomed the interactive and participatory nature of the discussions.

Having reviewed the reports from the various regional and country consultations, and taking into account both the in-depth analysis of progress in health research over the past decade and the discussions before and during the meeting, We, the participants, make the following Declaration.

The Conference reaffirms that health is a basic human right. Health research is essential for improvements not only in health but also in social and economic development. Rapid globalization, new understanding of human biology, and the information technology revolution pose new challenges and opportunities. Social and health disparities, both within and between countries, are growing. Given these global trends, a focus on social and gender equity should be central to health research. In addition, health research, including institutional arrangements, should be based on common underlying values. There should be:

- a clear and strong ethical basis governing the design, conduct and use of research;
- the inclusion of a gender perspective;
- a commitment that knowledge derived from publicly funded research should be available and accessible to all;
- an understanding that research is an investment in human development;
- a recognition that research is an investment in human development; and
- a recognition that research should be inclusive, involving all stakeholders including civil society in partnerships at local, national, regional, and global levels.

An effective health research system requires:

- coherent and co-ordinated health research strategies and actions that are based on mutually beneficial partnerships between and within countries;

- an effective governance system;
- a revitalized effort from all involved in health research to generate new knowledge which addresses the problems of the world's disadvantaged, and increases the use of high quality, relevant evidence in decision-making.

It is the responsibility of an active civil society through their governments and other channels to set the direction for the health research system, nurture and support health research, and ensure that the outcomes of research are used to benefit all their peoples and the global community.

We the participants commit ourselves to ensuring that health research improves the health and quality of life of all peoples.

The work carried out in preparation for, and during, the Conference should continue, through a process that will allow all stakeholders to contribute to debate and decisions on the key issues for the future of health research for development.

Source: www.globalforumhealth.org\non-compliant_pages\forum4\declaration.htm

The International Conference on Health Research for Development was organized by:

- The Council on Health Research for Development COHRED (www.cohred.ch)
- The Global Forum for Health Research (www.globalforumhealth.org)
- The World Bank (www.worldbank.org)
- The World Health Organization (www.who.int)

Glossary of terms in health research

Abstract An abbreviated summary of a research paper, generally at the beginning of the paper.

Action research A style of research in which the researchers work with the people and for the people, rather than undertake research on them. The focus of action research is on generating solutions to problems identified by the people who are going to use the results of research.

Adjusted rates Terms used when results have undergone statistical transformation to permit fair comparison between groups differing in some characteristic that may affect risk of disease.

Analytical study An observational study that describes associations and analyses them for possible cause and effect.

Alternative hypothesis The hypothesis that the researcher is testing in the study. In scientific methodology, we start with the assumption that it is not true until proved otherwise, by rejecting the null hypothesis.

Anonymous linked information Information which cannot be linked to the person to whom it refers, ensuring that the investigator cannot know the identity of the person and there is complete confidentiality in a study.

Assignment The process in an experiment where the researcher allocates subjects to two or more groups, trying to achieve having groups as identical as possible to allow a valid comparison of the results. Matching and random assignment are the two most common methods.

Attributable risk An estimate to quantify the contribution which a particular risk factor makes in producing the disease within a population.

Audit of a trial A systematic examination, carried out independently of those directly involved in the clinical trial.

Bar or column charts A graphic method of describing the data, where the frequency of a particular category is reflected in the height of the bar in the graph.

Baseline A phase in an intervention study where the participants have not received any intervention.

Basic risk An expression of the likelihood that a particular event will occur within a particular population.

Before-and-after study A method of control in which results from experimental subjects are compared with outcomes from patients treated before the new intervention was available. These are called historic controls.

Bell-shaped curve The characteristic shape of the curve of a normal distribution, where the data are equally distributed around the mean.

Beneficence An ethical principle implying that every effort should be made to maximize the benefits to the subjects in health research.

Bias If the study sample is not representative of the population, the inference we make from the result may be misleading.

Blinding A randomized controlled trial may be blinded if participants in the trial are likely to change their behaviour in a systematic way that may influence the outcome of the study when they are aware of which intervention they receive. The term "masking" is often used instead of "blinding".

Case–control study A type of observational analytical longitudinal retrospective study in which a group of subjects with a specified outcome (cases) and a group without that outcome (controls) are identified. Investigators then compare the extent to which each subject was previously exposed to the variable of interest, such as risk factor, a treatment, or an intervention.

Categorical variables Data where each individual variable is one of a number of mutually exclusive classes.

Central tendency The average (mean), middle (median) or most common (mode) score for numerical data in a frequency distribution.

Chi-square (χ^2) A statistical test used for categorical data. It is based on a comparison of the frequencies observed and the frequencies expected in the various categories.

Cluster sampling A type of random sampling, based first on the random selection of certain subgroups, from which the sample can be taken.

Coding A method of analysis of qualitative data obtained for example in interviews, where categories are labelled to facilitate computer analysis and examination of relationships.

Cohort study The term used in clinical and epidemiological research to describe a longitudinal prospective observational study.

Confidence interval A statistic of the expected range in which the population value will be found, at a given level of confidence or probability.

Conflict of interest Investigators may have vested interests in the research. These may be intellectual property interests as well as commercial interests. Such interest should be explicitly declared.

Confounder In simple terms, confounders are all of the "other things" that could explain the result of the research. In technical terms, confounders are factors that are associated with both exposure and outcome.

Consecutive sampling A sampling procedure in which subjects are selected by taking every individual that presents over a specified period of time.

Continuous variables Data which are measured on a continuous scale. They are numbers that can be added, subtracted, multiplied and divided.

Correlation The strength and direction of the association between two variables. Correlation does not mean causation.

Correlation coefficient A statistic designed to measure the size and direction of the association between two variables. The value varies between 0 and ± 1 (1 means complete correlation).

Cost–benefit analysis A type of economic study design in which both costs and benefits of interventions are expressed in monetary units, allowing direct comparison of competing interventions.

Cost–effectiveness analysis A type of economic study design in which the net monetary costs of a health care intervention per unit measure of clinical outcome or effectiveness allows direct comparison of competing interventions.

Crossover study A special design of controlled trials in which half of the participants are randomly assigned to start with the placebo and then switch to active treatment, while the other half does the opposite.

Cross-sectional study An observational study design in which measurements are made on a single occasion.

Cross-tabulation tables Frequency distribution tables that examine the relationship between several of the variables at once, for better description of the data or in order to look for differences or relevant associations.

Crude rates Terms used when results have not been adjusted for confounding factors.

Dependent or output variables Responses or consequents that are contingent on independent variables.

Descriptive statistics Statistics designed to summarize and describe characteristics of the data. Descriptive statistics helps us to make sense of a large volume of data.

Descriptive study An observational study that simply describes the distribution of a characteristic.

Directional research hypothesis The research hypothesis outlining a relationship may be directional or non-directional. For example, a relationship between smoking and cardiovascular disease can only be directional. It is expected in the hypothesis that it will increase cardiovascular disease. The relationship between oral hormonal contraceptives and certain disease conditions can be non-directional. The disease conditions may increase or decrease as a result of oral hormonal contraceptive use.

Disability-adjusted life years (DALYs) lost An international measure of the burden of disease that expresses both time lost through premature death and time lived with a disability.

Discrete or discontinuous data Numerical variables that are not measured on a continuous scale.

Distributive justice An ethical principle implying that participation in the research should correlate with expected benefits. No population group should carry an undue burden of research for the benefit of another group.

Duplicate or redundant publication Publication of a paper that overlaps substantially with one already published by the same authors.

Effect size The amount of change associated with an intervention or risk factor. It is important in determining how significant the findings are in actual practice.

Ephemeral literature Literature judged to have a short period of usefulness and only for a small audience, not normally considered worth indexing or cataloguing. It may, however, be important. It includes reports, proceedings of conferences and other types of publication.

Essential national health research Each developing country should establish and strengthen an appropriate health research base to understand its own problems, improve health policy and management, enhance the effectiveness of limited resources, foster innovation and experimentation, and provide the foundation for a stronger developing country voice in setting international priorities.

Experimental or intervention study A study design in which the investigators test the effect of an intervention on the events taking place in the study.

External validity The extent to which the results of the study sample may be generalized to the population from which the sample was withdrawn; also called generalizability.

Focus group discussion A method of qualitative research used when information and insights will be better gained from the interaction of a group than from in-depth interviews with individuals.

Forced-choice format A format for closed-response questions used to elicit attitudes of the respondents to a certain statement. The respondent choices are limited to four: strongly agree, agree, disagree and strongly disagree. This format, different from the Likert format, does not allow an undecided answer.

Fraud Scientific fraud is deliberate deception and may take the form of fabricating data, inventing patients, or manipulating data to provide a desired answer.

Frequency distribution The way in which scores within a given sample are distributed.

Frequency distribution curve A graphic method for summarizing data and looking at them, in which each variable is plotted against the frequency with which it is found.

Frequency distribution table A table that gives the frequency with which a particular value appears in the data.

Gaussian distribution A bell-shaped frequency distribution curve, also described as “normal”.

Good clinical practice (GCP) Standard for clinical studies which encompasses the design, conduct, monitoring, termination, audit, analyses, reporting and documentation of the studies and which ensures that the studies are scientifically and ethically sound and that the clinical properties of the pharmaceutical product under investigation are properly documented.

Grantsmanship The ability to secure grants to support research projects.

Hawthorne effect An effect which results in the improvement of subjects’ performances through being observed and/or social contact. It is an example of a placebo effect.

Histogram A method of plotting frequency distributions.

Hypothesis The research hypothesis is a tentative statement that can be tested by a scientific research design.

Impact factor A measure of the frequency with which the "average article" in a journal has been cited in a particular year or period. It provides a way to judge the prestige and influence of a particular journal.

Incidence Incidence rates relate the number of new cases of a condition in a population within a time period.

Independent or input variables Variables that have values that are autonomous of the dependent or outcome variables. Because independent variables precede dependent variables, they are often called predictors. In epidemiology, independent variables are often called risk factors or exposure variables.

Inference A generalization made about a population from the study of a subset or sample of that population.

Informed consent An ethical requirement for participation in a research study, indicating that a competent person, in possession of all the relevant information, freely agrees to participate.

Internal validity The degree to which the investigator's conclusions correctly describe what actually happened in the study. It means that within the confines of the study, results appear to be accurate, the methods and analysis used stand up to scrutiny, and the interpretation of the investigators appears supported.

Inter-observer reliability The extent to which observers rating or measuring a particular phenomenon agree with each other.

Intra-observer reliability The extent to which an observer rating or measuring a particular phenomenon agrees with her/his rating or measurement when presented with the same task on two different occasions.

Interquartile range The distance between the scores representing the 25th and 75th percentile ranks in a distribution.

Likert format A format for closed-response questions used to elicit attitudes of the respondents to a certain statement. The respondent chooses from among five categories: strongly agree, agree, undecided, disagree, strongly disagree.

Literature Previous research done in the area under study.

Logistic regression Method commonly used by statisticians for multivariate analysis.

Longitudinal study An observational study design in which measurements are made over a period of time.

Longitudinal prospective study An observational study design in which the investigators follow subjects for future events.

Matching A sampling method to ensure that the two groups to be compared have similar characteristics. In an intervention study, pairs of similar "matched" subjects are formed and then

one member of the pair is randomly assigned to one group and the other member to the other group.

Mean The average of a group of scores. The mean is derived by summing up the individual values and dividing by the total number of measurements.

Measurement or information bias Measurement bias occurs when the methods of measurement are consistently dissimilar in different groups of patients.

Median The median of a distribution is a midpoint at which one half of the observations fall below and one half fall above the value.

MEDLINE A bibliographic database which provides details of articles and their abstracts, from peer-reviewed journals. MEDLINE is funded by the US National Institutes of Health.

Meta-analysis A methodology to critically review research studies and statistically combine their data to help answer questions that are beyond the power of single papers.

Mode The most frequent measurement in a distribution.

Multivariate analysis Assessment of the independent contribution of multiple independent variables on a dependent variable, to identify those independent variables most significant in explaining the variation of the dependent variable.

Negative correlation A negative correlation between two variables implies that as one variable gets bigger the value of the other variable becomes smaller.

Nominal categorical data Data in which the categories cannot be ordered one above another. Examples of categorical nominal variables are sex and marital status.

Non-maleficence An ethical principle implying that where research involves experimentation on human subjects, the subjects should suffer no harm.

Non-nominal linked information Information linked to the person by a code (not including personal identification) known to the investigator.

Non-parametric tests Statistical tests that can be applied when the data fall in a frequency distribution curve that is skewed. Also called "distribution free" statistics.

Normal distribution curve A bell-shaped curve of the frequency distribution of the data.

Null hypothesis In scientific methodology, we do not test the research hypothesis directly. Instead, we start with an assumption that there is no difference or association between the variables compared. This is called the null hypothesis (H_0). If statistical analysis rejects the null hypothesis, it means that the alternative hypothesis is probably true, and that there a difference between the group or a relationship between the variables.

Numerical variables Data expressed in numbers.

Objectivity Objective measures are made in a process involving a minimum amount of human interpretation, for example measurement of height.

Observational study A study design in which the investigators observe and record events taking place in the study.

Odds ratio Term used in case-control studies as a measure of the odds of having the risk factor among people with the disease divided by the odds of having the risk factor among people without the disease.

One-tailed test A statistical test where a difference between two groups, if true, is expected to be in one direction. For example, the difference between passive smokers and non-smokers in the occurrence of lung cancer is expected to be in one direction. It is not expected that smoking will protect from lung cancer, and so there is no need to test for it. A one-tailed test will need a smaller sample size than a two-tailed test.

Open-ended question A question asked without providing a pre-defined set of responses to select from.

Ordinal categorical data Categorical data in which the variables can be ordered one above another. An example of ordinal categorical data is the number of children a woman has.

***P* value** The probability that a difference or an association as large as the one observed could have occurred by chance alone.

Parametric tests Statistical tests that can be applied when the data fall in a normal distribution, that is, when they are spread evenly around the mean, and the frequency distribution curve is bell-shaped or Gaussian.

Peer-reviewed journal A journal in which the articles are vetted by independent referees for quality and interest, and is therefore more highly regarded.

Phase I clinical trials First trials of a new active ingredient or new formulation in humans, often carried out in healthy volunteers.

Phase II clinical trials Trials performed in a limited number of subjects and often of a comparative (e.g. placebo-controlled) design, to demonstrate therapeutic activity and to assess the short-term safety of the active ingredient in patients suffering from a disease or condition for which the active ingredient is intended.

Phase III clinical trials Trials including larger (and possibly varied) patient groups, with the purpose of determining the short-and long-term safety/efficacy balance of formulation(s) of the active ingredient, and of assessing its overall and relative therapeutic value.

Phase IV clinical trials Studies performed after marketing of the pharmaceutical product to discover rare and remote side-effects.

Pie chart A graphical method of representing the frequency distribution of a set of categorical data in the shape of a pie.

Pilot study A preliminary study to test the feasibility of the protocol, before implementing the study proper. It may also be called “pre-test”.

Placebo effect The phenomenon where, in an intervention study, subjects receiving, without knowing, an inert drug, show an improvement or perception of improvement in their condition, probably due to their expectations.

Population An entire set of persons, animals, objects or events which the researcher intends to study.

Positive correlation A positive correlation between two variables implies that as one variable gets bigger the value of the other variable also becomes bigger.

Power A statistic indicating the probability of rejecting the null hypothesis when the alternative hypothesis is true. Statistical power of a study is thus the probability of observing an effect (of a specified effect size) if one exists.

Predictive value The frequency with which a positive diagnostic test actually signifies disease.

Pre-test A preliminary study to test the feasibility and appropriateness of a questionnaire, before implementing the study proper.

Pre-test/post-test design An experimental research design in which measurements of the groups are made both before and after an intervention.

Prevalence The overall occurrence of a particular condition in a specific population at a specific point of time.

Probability The chance or likelihood of an event happening. Probability may vary in value from 0 (no chance) to 1 (certain). Researchers have to set the level of probability/certainty they are willing to accept for their findings.

Proportion The ratio of one value to another expressed as a fraction of one. For example, the proportion of women among patients with cardiovascular disease.

Proposal A document written for the purpose of obtaining funding for a research project.

Protocol The detailed written plan of the study. Any research study should have a protocol.

PubMed Central A public web-based archive offering barrier-free access to peer-reviewed primary research reports in the life sciences, funded by the US National Institutes of Health.

Quality assurance A system to ensure that the study is performed and the data are generated, recorded and reported in compliance with the protocol, good clinical practice and national regulations.

Qualitative methods A research approach that emphasizes the non-numerical data and interpretive analysis.

Quantitative methods A research approach that emphasizes the collection of numerical data or data than can be quantified, and statistical analysis.

Questionnaire A means of collecting data from people where they provide written responses to a set of questions, either in their own words (open-ended questions), or by selecting from among pre-defined answers (closed response questions).

Random sampling A sampling procedure in which a sample is drawn from a population such that each member of the population has had an equal chance of selection. Random sampling is not haphazard sampling.

Randomized controlled trials Intervention studies characterized by the prospective assignment of subjects, through a random method, into an experimental group and a control group.

Range In a group of scores, the range is the difference between the maximum and minimum scores.

Ratio A numerical expression of the relationship between one set of frequencies and another. An example is the ratio of males to females in a sample.

Rate A numerical expression of the frequency of a condition in a given population measured in a specified period of time.

Regression equation An equation to describe the correlation between two variables, meaning that when one of them changes by a certain amount the other changes on the average by a certain amount.

Regression line A line drawn on a scatter diagram, to illustrate the degree and direction of the correlation between two variables.

Regression coefficient The term used to signify the amount by which a change in one variable must be multiplied to give the corresponding average change in the other variable. It represents the degree to which the regression line slopes upwards or downwards.

Regression to the mean A phenomenon where, upon re-measurement, previous extreme (very high or low) scores tend to move towards (regress to) the average score.

Relative risk The ratio of the incidence of the outcome in the exposed group to the incidence of the outcome in the unexposed group.

Reliability The extent to which a test or measurement result is reproducible.

Representative sample A sample that accurately reflects the characteristics of the population from which it is drawn. It is a precise miniaturized representation of the proportion of elements of the population.

Retrospective study An observational study design in which the investigators study present and past events.

Risk factors A factor that is believed to increase the probability of a certain outcome or illness.

Rosenthal effect The phenomenon where the expectations of the researchers in a study influence the outcome.

Sample A subset selected for the study from the larger population.

Sampling error The discrepancy between the values obtained from the relatively small sample and the larger population from which the sample was drawn.

Scatter diagram A graph displaying the scatter of the relationship between two variables. The scatter diagram gives an indication of whether a correlation may exist and its direction.

Selection bias A systematic difference between people who are selected for a study and those who are not selected.

Sensitivity of a diagnostic test is the proportion of people who test as positive to a disease who really have the disease, i.e. they are true positive.

Skewed distribution A frequency distribution curve which is asymmetrical, with one side of the curve extending in an elongated fashion.

Specificity The proportion of people who test negatively for a disease.

Standard deviation A measure of the dispersion or variability of a group of scores.

Standard error A statistical measure of the probability that the finding in the sample will reflect the finding in the population from which the sample was drawn.

Statistical significance A statistic indicating that the result obtained is probably not due to chance but is real. A statistically significant result does not necessarily mean that it is important or interesting.

Statistical significance test A test to estimate the likelihood that an observed study result, for example a difference between two groups or an association, can be due to chance.

Stratified random sampling A sampling procedure in which the researcher tries to ensure that important subgroups in the population are adequately represented.

Structured interview An interview in which the questions are generally pre-defined, asked in a fixed order and recorded in writing.

Subjective measures Measures involving a substantial degree of human interpretation, for example ratings of pain.

Subjects Participants in a study. They should not be called material for the study.

Surrogate end point A variable that is relatively easily measured and that predicts a rare or distant outcome, but which is not itself a direct measure of either harm or clinical benefit.

Systematic sampling A sampling procedure in which subjects are selected by a simple periodic process, for example, selecting every second or third patient.

***t* test** Statistical test used for numerical data to determine whether an observed difference between the means of two groups can be considered statistically significant, i.e. unlikely to be due to chance.

The 10/90 gap While 90% of the global burden of disease is in developing countries, only an estimated 10% of the global resources are spent on disease problems of developing countries.

Transcript A verbatim written version of an interview.

True negative A diagnostic test correctly indicating that a person does not have the disease.

True positive A diagnostic test correctly indicating that a person has the disease.

Two-tailed test A statistical test where a difference between two groups is tested without reference to the expected direction of the difference, for example whether a risk factor, such as use of hormonal contraception will increase or decrease the incidence of a condition. A two-tailed test will need a larger sample size than a one-tailed test.

Type I error The error committed when, on the basis of a statistical test applied to the sample of data, a conclusion is made that there is evidence of an association between variables or difference between groups in the population, when in fact there is no difference or association. The probability of type I error is represented by the symbol alpha (α). Another name for alpha is the level of statistical significance.

Type II error A "miss", when, on the basis of a statistical test applied to the sample of data, a conclusion is made that there is no evidence of an association between variables or difference between groups in the population, when in fact there is a difference or association. The probability of type I error is represented by the symbol beta (β).

Unlinked information Information which cannot be linked, associated or connected with the person to whom it refers; confidentiality here is not at stake.

Univariate analysis A set of mathematical tools to assess the relationship between one independent variable and one dependent variable.

Validity The extent to which a test measures what it is intended to measure.

Variability The extent to which a group of scores varies or is spread out. This is usually described by a descriptive statistic such as the range or standard deviation.

Variable Statistical term for the score in data.

Variance A measure of the dispersion or variability of a group of scores.

Index

A

Absolute risk 111
Abstract 131
    Structured 131
Accessible population 50
Acknowledgements,
    Writing of 138
Action research 28
Adjusted rate 110
Alpha 95
Alternative hypothesis 94, 100
Analytical study 44
Attributable risk 113
Attrition 108
    Attrition bias 108
    Audit 83
    Trail 117
Authorship 156

B

Bangkok declaration 214
Basic versus applied research 27
Before-and-after study 47
Bell-shaped curve 91
Beneficence 21, 60, 61
Beta 95
Bias 107
Big science 28
Biological gradient 111
Biological plausibility 111
BioMed Central 123
Blinding 47, 80
Blobbogram 146
Boolean operators 204
Budget of research proposal
    Itemization 77
    Justification 77

C

Case–control studies 44
Case report 145
Categorical variables 99
Causation, making the case for 110
Central tendency 89
Chi-square test 102
CIOMS 21
Clinical practice guidelines 35, 125
Clinical trials, phases of 83
Cluster sampling 52
Cochrane Collaboration 124, 177
Cochrane Library 176
Coding of the data 103
Coding sort 104
Cohort study 44
Collaboration between industry and academia 35
Computer-assisted presentation 167
Confidence intervals 94, 96, 113
Confidentiality 21, 60, 61, 198
Conflict of interest 158, 174
Confounding 109
Consecutive sampling 52
Continuous variables 99
Controlled trials 46
Controls, historic 47
Copyright 157
Correlation 91
    Coefficient 92
Cost-benefit analysis 58
Cost-effectiveness analysis 58
Cross-sectional study 44
Crossover study 46
Cross-tabulations 88

D

Data 99, 100
    Distribution 100
    Management and analysis of 68
    Missing 81
    Types 99
Data immersion 103
Degrees of freedom 101
Dependent variables 91
Descriptive study 44

Descriptive statistics 87
Design of research
Types 44
Selection 47
Diagnostic suspicion bias 108
Diagnostic tests 114
Interpreting studies of 114
Discrete variables 99
Dissertation 147
Distributive justice 21
Dose-response relation 45, 83, 111
Double-blind studies 47
Dummy tables 88
Duplicate publication 155

E

Economics, health research in 28, 58
Effect size 54, 98
Efficiency in diagnostic tests 115
Electronic index 104
Ephemeral literature 36
Essential national health research 34
Ethics committees 22, 69, 76, 191
Ethical review of externally sponsored research 192
Ethics 20, 40, 133
General principles 21
International guidelines 191
Responsibility for 22
Ethics in
Epidemiological, field and qualitative studies 61
Experimentation on human subjects 59, 185, 191
Experimentation on animals 41, 62, 85
Implementation of the study 85
Research combined with medical care 189
Research design 59
Research involving children 99
Research protocol 68
Research proposal 76
Research involving pregnant women 198
Scientific publication 157
Selection of research topic 40
Writing a scientific paper 133
Evidence-based reviews 124
Experimental study 45
External validity 93, 152

F

Focus group discussion 56
Figures 91, 136
Fiscal honesty 86
Flow charts 104, 165
Fraud, scientific 85, 160
Frequency distribution curve 91
Frequency distribution tables 88
Funding
    Availability 32
    Sources 72

G

Gaussian distribution 91
Gender issues 70, 76
Generalizability 93
Good clinical practice GCP 35, 82
Graphs 91, 135

H

Helsinki declaration 20, 133, 185
HINARI 38, 208

I

Illustrations 135
Impact factor 121, 181
IMRAD 132
Independent variables 91
Index Medicus 121, 140, 209
    Eastern Mediterranean Region 39
Industry-sponsored research
    Concerns about participation 35
Inference 93
Inferential statistics 93
Information bias 108
Informed consent 188, 192, 194, 196, 197
Informed decision-making 69
Intellectual property 121
Inter-observer reliability 54, 81
Internal validity 152
International Committee of Medical Journal Editors 35, 130, 153
International health research
    Models for participation 32
    Concerns in developing countries 34
    Ethical concerns 34

Internet
Searching 203
Intervention studies 45
Interviews 56
Interquartile range 90
Intra-observer reliability 54, 81

K

Key words 131

L

Levels of evidence 175
Linked information 61
Longitudinal study 44

M

Matching 109
Measurement bias 108
Mean 89
Median 89
MEDLINE 201, 202
MEDLINEplus 127
MeSH 131, 202
Meta-analysis 145, 146, 176
Mode 89
Monitoring of the study 80
Multi-centre clinical trials 33
Multidisciplinary research 26
Multivariate analysis 100

N

Narrative review 145
National Library of Medicine 200
New pharmaceutical products
Research on 35, 83
Non-maleficence 21, 60
Non-parametric 100
Normal distribution 100
Null hypothesis 94
Number needed to treat 116

O

Observational studies 44
Odds ratio 113
Online searching 203
Opportunity cost 178

Outcome variables 92
Overhead projector 166

P

P value 96
Parametric tests 100
Patents and publication 157
Peer review 174
Percentages and proportions 90
Pharmaceutical company research
    Participation in 34
Pilot study 57, 80
Placebo controlled trial 46, 196
Post-marketing research 35
Predictive value, in diagnostic test 114
Pre-testing the protocol 80
Privacy, right to 159
Prospective study 44
Protocol 65
    Format for the 66
Public library of science 123
Public media 160
PubMed 152, 201
PubMed Central 38, 202, 206

Q

Qualitative research
    Description and analysis of results 102
    Ethics in 61
    Interpreting results of 116
    Planning of 55
    Quantitative versus 27
    Validation of result 82
    Writing a paper on 147
Quality assurance and quality control 81
Questionnaire design 57

R

Random sampling 51
Randomized controlled trials 46
Randomization 46
Range 90
Rates 91
Ratios 90
Recall bias 45

Receiver operator characteristic 115
Redundant publication 158
Relative risk 113
Reliability 54
References
    Citation of 139
    Guidelines on how to write 210
    How to order 213
Regression equation 92
Research question 49
Research design
    Types of 44
    Selecting 47
Research proposal 72
Research topic 36, 39
Research hypothesis
    Generating the 50
    Testing of 94
Response bias 108
Retrospective study 44
Risk factors 112, 113
    Interpreting results of 112
Rosenthal effect 47
Running title 131

S

Sample size 52
Sampling
    Types of 51
Scatter diagram 92
Science citation index 180
Scientific honesty 85
Scientific rigour 79
Search engines 203
Searching the literature 201
Secondary scientific papers, writing of 145
Selection bias 108
Sensitivity, in diagnostic tests 114
Skewed distribution 91
Slides 164
Specificity in diagnostic tests 114
Standard deviation 90
Standard error 93
Statistical power 53, 98
Statistical significance 53, 95, 96, 107

Interpreting 106
Statistical tests 95, 98, 101
Examples of 101
Selection of 98, 100
Statistics 133
In a scientific paper 134
Stratified random sampling 52
Submission of manuscript 155
Technical instructions 154
Uniform requirements 153
Surrogate endpoint 112
Systematic review 145, 176
Systematic sampling 52

T

t test 101
Tables
In scientific paper 134
Termination of the study 84
Time and motion study 55
Title 131
Of scientific paper 131
Running title 131
Transferability 103
Thesis 148
Type I error 95
Type II error 95

U

Unlinked information 61
Uniform requirements for manuscripts submitted to biomedical journals 153
Univariate analysis 100

V

Validity 54, 173
Vancouver Group 140, 153
System for citation of references 140, 213
Variables 99