

Research awareness in nursing and midwifery: a workbook

Commissioned by:
**WHO Collaborating Centre for Reference, Education and
Research in Primary Health Care Nursing**

Compiled by:
**Kate Wilson
Tony Butterworth**

Critical readers:
**Jackie Oldham
Heather Waterman**

EUROPEAN HEALTH21 TARGET 18

DEVELOPING HUMAN RESOURCES FOR HEALTH

By the year 2010, all Member States should have ensured that health professionals and professionals in other sectors have acquired appropriate knowledge, attitudes and skills to protect and promote health

(Adopted by the WHO Regional Committee for Europe at its forty-eighth session, Copenhagen, September 1998)

EUROPEAN HEALTH21 TARGET 19

RESEARCH AND KNOWLEDGE FOR HEALTH

By the year 2005, all Member States should have health research, information and communication systems that better support the acquisition, effective utilization, and dissemination of knowledge to support health for all

(Adopted by the WHO Regional Committee for Europe at its forty-eighth session, Copenhagen, September 1998)

This workbook on research awareness has been produced for nurses and midwives who have little or no knowledge of research methods. It has been written for an international readership. It is not intended to provide sufficient knowledge to make readers proficient in conducting research, although advice on how to achieve this is included. The workbook comprises sections on carrying out a literature search, reading and evaluating research documents, further research-related activities, a glossary of research terms, and a discussion of the exercises.

After reading the workbook, undertaking the exercises and reading the discussion of the exercises, readers should be able to:

- conduct a literature search;
- identify sources of information available in their countries/local areas;
- make attempts at critical evaluation of research-based books, articles, reports and other publications; and
- consider whether specific publications have implications for their clinical practice.

Keywords

NURSING AND MIDWIFERY
EDUCATION, NURSING
RESEARCH EDUCATION
TEXTBOOKS

ISBN 92 890 1181 5

© World Health Organization 2000

All rights in this document are reserved by the WHO Regional Office for Europe. The document may nevertheless be freely reviewed, abstracted, reproduced or translated into any other language (but not for sale or for use in conjunction with commercial purposes) provided that full acknowledgement is given to the source. For the use of the WHO emblem, permission must be sought from the WHO Regional Office. Any translation should include the words: *The translator of this document is responsible for the accuracy of the translation.* The Regional Office would appreciate receiving three copies of any translation. Any views expressed by named authors are solely the responsibility of those authors.



CONTENTS

	<i>Page</i>
Foreword.....	i
1. Introduction.....	1
1.1 Learning objectives.....	1
1.2 What is research?.....	1
1.3 The relevance of research to nursing practice.....	3
1.4 What is research awareness?.....	4
2. Carrying out a literature search.....	5
2.1 Identify a topic relevant to your clinical practice.....	5
2.2 Identify a suitable library.....	5
2.3 Find indexes, abstract journals and other works of reference.....	5
2.4 Obtain selected articles, books and reports.....	6
2.5 Keep a record of each document.....	6
3. Reading and evaluating research reports.....	9
3.1 Why critical evaluation is important.....	9
3.2 Factors for critical consideration.....	9
3.3 Biographical details of the researchers.....	10
3.4 Source(s) of research funding.....	10
3.5 Details given in the abstract.....	11
3.6 The aims of the study.....	11
3.7 The literature review.....	13
3.8 The choice of research methods.....	15
3.9 Pilot work.....	21
3.10 Sampling procedures.....	21
3.11 Data analysis.....	27
3.12 Presentation of data.....	34
3.13 Presentation of quantitative data.....	34
3.14 Discussion of data.....	37
3.15 Conclusions and recommendations.....	39
3.16 References.....	40
3.17 Ethical considerations.....	40
3.18 Rigour.....	42
3.19 Relevance to practice.....	45
4. Where to go from here.....	47
5. Glossary of research terms.....	49
6. Discussion of exercises.....	55
6.1 Exercise 3. The aims of the study.....	55
6.2 Exercise 4. The literature review.....	55
6.3 Exercise 5. The choice of research methods.....	57
6.4 Exercise 6. The choice of research methods.....	58
6.5 Exercise 7. Pilot work.....	58

6.6	Exercise 8. Sampling procedures	58
6.7	Exercise 9. Sampling procedures	59
6.8	Exercise 10. Data collection methods	60
6.9	Exercise 11. Data collection methods	60
6.10	Exercise 12. Data collection methods	60
6.11	Exercise 13. Data collection methods	61
6.12	Exercise 14. Data analysis.....	61
6.13	Exercise 15. Data analysis.....	61
6.14	Exercise 16. Data analysis.....	62
6.15	Exercise 17. Data analysis.....	62
6.16	Exercise 18. Presentation of data	63
6.17	Exercise 19. Presentation of data	63
6.18	Exercise 20. Discussion of data	64
6.19	Exercise 21. Discussion of data	65
6.20	Exercise 22. Conclusions and recommendations	65
6.21	Exercise 23. Ethical considerations.....	67
6.22	Exercise 24. Rigour	67
6.23	Exercise 25. Rigour	69
7.	References	70
Annex 1.	Further relevant materials from the LEMON project.....	73
Annex 2.	Research articles.....	75

Foreword

I welcome this document, and commend it as of relevance to nurses and midwives across the WHO European Region, who must now – and increasingly in the future – seek an evidence base for their practice. To seek evidence is but the important first step, however; to understand and be able to evaluate the worth of research evidence is equally vital. This workbook will assist nurses and midwives to think analytically and critically, to weigh up such evidence as exists and then, if appropriate, to apply it in their practice.

Inevitably, and rightly, this workbook requires the reader to consult other relevant texts. In some countries where nurses and midwives do not easily have access to research and library resources in their own language, this will be a major hurdle and challenge. Nevertheless, quite a number of such countries, particularly in the CCEE and NIS, have translated the LEMON materials, and also have a LEMON coordinator and/or LEMON group. For nurses and midwives in these countries, LEMON will be a very valuable aid in using this workbook. Annex 1 gives some specific references to help the reader link the text of the workbook with some of the chapters in LEMON.

I should like to express the grateful thanks of the Nursing unit at the WHO Regional Office for Europe to the WHO collaborating centre in Manchester for making this workbook available to WHO.

Ainna Fawcett-Henesy
Regional Adviser for Nursing and Midwifery
WHO Regional Office for Europe

1. Introduction

“What is research?”, “How do I begin to learn about research?” and “How do I begin to learn about the relevance of research to my clinical practice?”. The information contained in this document goes some way to answering these and similar questions. It is also designed to help nurses¹ who wish to explore the knowledge that underpins certain aspects of nursing practice, with a view to maintaining or improving quality of care.

Every nurse should become research-aware. Only a minority may wish or be able to conduct a research project, but it is important that all nurses know how to search for relevant publications and to evaluate them critically, so that they can clarify the rationale behind the care they give.

1.1 Learning objectives

By the time you have read this workbook you should be able to:

- conduct a literature search;
- identify sources of information available in your country/local area;
- make attempts at critical evaluation of research-based books, articles, reports and other publications; and
- consider whether specific publications have implications for your clinical practice.

After reading this document and doing some critical reading of research articles, you may feel ready to join or form a journal club. This is a group of people who meet informally to critique journal articles and discuss their relevance to clinical practice. You may also feel better equipped to adopt a critical attitude to any research projects taking place in your clinical situation.

While this document may give you an awareness of research issues, however, it will not provide sufficient knowledge to make you competent in carrying out research. Before attempting to design or conduct your own research project you should, if possible:

- take a research methods course; and/or
- undertake further distance learning on research; and/or
- read research textbooks; and/or
- obtain guidance and support from an experienced researcher; and
- (for quantitative research) get advice from a statistician.

These matters are discussed further in Section 4, page 47.

1.2 What is research?

“Research is a means of understanding, assessing and evaluating what we do as nurses” (1).

¹ Throughout this document “nurse” and “nursing” are used as general descriptive terms. Terminology and definitions vary from country to country, and the term “nurse” should be understood to include all health workers doing nursing-related work, including nurses, midwives and feldschers.

This quotation defines research in terms of its purpose or usefulness. According to this definition, research is important because it increases knowledge and encourages critical thinking about practice. Nursing research can enable nurses to examine and challenge established practices and to explore the effectiveness and value of new or proposed practices. Research findings may have implications for nurse education as well as nursing practice.

Research can also be characterized as a process: “... research is a systematic way of asking questions, a systematic method of enquiry” (2).

The key word in this quotation is “systematic”. Research must be conducted as a thorough and logical process if it is to produce credible findings that can be considered as an addition to existing knowledge. In particular, researchers need to demonstrate that they have considered and/or ruled out alternative explanations for their findings. The words “scientific” and “rigorous” are sometimes used in this context.

The following World Health Organization (WHO) definition (3) highlights topics that are appropriate to nursing research.

Nursing research focuses on developing knowledge of the care of people in health and illness. It is directed towards understanding the fundamental genetic, physiological, social, behavioural and environmental mechanisms that affect the ability of individuals and families to maintain or enhance optimum function and minimize the negative effects of illness.

Nursing research also emphasizes the generation of knowledge of policies and systems that effectively and efficiently deliver nursing care; the profession and its historical development; ethical guidelines for the delivery of nursing services; and systems that effectively and efficiently prepare nurses to fulfil the profession’s current and future social mandate.

In addition, principal investigators who are nurses often conduct scientific inquiry into the cause, diagnosis and prevention of diseases; the promotion of health, the processes of human growth and development, and the biological effect of environmental contaminants. These foci, however, are not unique to nursing.

This is a wide-ranging definition of the scope of nursing research, appropriate to an organization that takes an international perspective. Hockey (4), however, notes that what is defined as nursing research is likely to vary from country to country and also with time.

Because nursing encompasses a wide range of activities, because it is interpreted differently in different parts of the world, and because it changes over time, a definition of the term nursing research and an explanation of its nature should make provision for these variations.

In summarizing these points about the nature of research, and in particular nursing research, we can say the following.

- The purpose of research is to add to the sum of human knowledge. The knowledge generated by research may have practical implications. In nursing, these implications may be for nurse education and/or clinical practice.
- Research must be conducted in a systematic fashion if its findings are to be considered credible and an addition to knowledge.

- Nursing research encompasses a wide range of topics. These include clinical care, the maintenance of health, the causes of ill health and the education of nurses.
- Appropriate topics for nursing research will vary with place and time.

1.3 The relevance of research to nursing practice

Several authors and organizations have asserted that research is central to good nursing practice. In the United Kingdom, the Briggs Report (5) was one of the earliest documents to emphasize this point, stating that “nursing should become a research based profession ... a sense of the need for research should become part of the mental equipment of every practising nurse and midwife”.

This extract advocates that nursing practice should be built on research-based knowledge and that practitioners should give their support to researchers and make use of relevant findings. The International Council of Nurses (6) places the emphasis rather differently by exhorting researchers to make their research relevant to clinical and educational practice and to make their findings readily available.

The International Council of Nurses is convinced of the importance of nursing research as a major contribution to meeting the health and welfare needs of people. The continuous and rapid scientific developments in a changing world highlight the need for research as a means of identifying new knowledge, improving professional education and practice and effectively utilizing resources.

ICN believes that nursing research should be socially relevant. It should look to the future while drawing on the past and being concerned with the present.

Research findings should be widely disseminated and their utilization and implementation encouraged when appropriate.

Clark & Keeble (7) note that nurses may be engaged in research-based activity at different levels. In line with the Briggs Report (5), they advocate that all nurses “develop research awareness”. They state that all nurses should “contribute to the research agenda and make use of research in practice”. Some nurses, particularly senior practitioners, nurse managers and educationalists, may be more involved in identifying relevant research topics, in supporting research projects and in implementing findings. Some of these nurses may undertake research studies themselves. A few nurses become “full-time research workers” (7).

Clark & Keeble state, however, that the use of research evidence is only one of several factors that may contribute to good nursing practice. Other important factors include: “clinical skills; communication skills; individual needs and circumstances of the patient or client experience; education; awareness of the resources available for understanding the environment in which you practise” (7).

It should also be remembered that all knowledge, whether gained through research, education or experience, is both limited and provisional: there is always more to know about a subject and what is considered to be “true” may change over time. An article by Sadler (8) on cot death illustrates this point by showing how ideas about the safest sleeping position for babies have changed over the years.

In addition, Rundell (9) warns that “we must guard against blindly accepting research findings”, just as we should guard against unthinkingly carrying out the instructions of a senior member of staff or doing things in accordance with tradition: “... we need to view research as a necessarily flawed item of work which one would respond to ... with some cynicism and distrust. Research should be subject to stringent criticism”.

Rundell also notes that there may be a particular danger of uncritical acceptance of research findings “where the research appears to back up our own prejudice”.

1.4 What is research awareness?

Research awareness is concerned with the knowledge nurses need to understand research issues. It is not primarily concerned with educating nurses to a level at which they are competent to conduct research themselves. “Research appreciation” and “research-mindedness” are alternative terms. In a report on nurse education from the Royal College of Nursing (10) the latter term is used.

We understand the term research-mindedness to imply a critical, questioning approach to one’s own work, the desire and ability to assess its value to the situation and apply it as appropriate. It also implies a recognition of the importance of research to the profession and to patient care, and a willingness to support nurse researchers in their work. All qualified nurses should be research-minded.

This quotation highlights several points:

- Research awareness (like research) involves critical thinking about issues in nursing.
- Research awareness involves an acknowledgement that knowledge gained through research may have important implications for both nurses and patients.
- Clinical nurses/nurse educationalists should give their support to nurse researchers.
- All qualified nurses should develop research awareness.

Hockey (4) states that awareness of and involvement in research is likely to develop at different rates in different countries. She places an emphasis on “a recognition that questions can and should be asked, that reading of research articles and research based texts is essential, and that research findings should be assessed in terms of their usefulness, their relevance, and their potential for implementation”.

In addition, as noted by Rundell (9), research studies and research documents vary in terms of quality. Readers need to develop the ability to identify flaws that undermine the claims that researchers make for their findings.

Section 3 on “Reading and evaluating research reports” will help to develop your critical skills and research awareness. The next section, however, gives information on how to identify and obtain research documents for critical scrutiny.

2. Carrying out a literature search

A literature search is carried out to find out what has been written about a chosen subject. This may be with a view to performing further research on the subject, or informing and/or questioning the knowledge base of certain aspects of practice. This section shows you how to carry out a literature search.

2.1 Identify a topic relevant to your clinical practice

Try to think of different terms that might define your topic of interest, such as “pressure area care” or “pressure sores”.

Identify related terms that might be relevant, such as “infection control”, “dressings” and “wound care”.

Try brainstorming the topic with colleagues and/or use a dictionary to identify synonyms. Time used on identifying alternative and related terms may be well spent, as it may lead to a thorough literature search.

2.2 Identify a suitable library

You may be able to gain access to a local university, college or hospital library or a national library where you can do your search. Some libraries have staff who will show you how to use the library and its various resources.

Some libraries may allow you to borrow books and reports, while others will restrict you to reading material in the library.

2.3 Find indexes, abstract journals and other works of reference

Indexes are works of reference that allow you to search in a subject index for details of material relevant to your chosen topic. You can use your research terms (e.g. “pressure sores” and/or “wound care”) to do this. You can also search in an author index for documents by specific authors.

Abstract journals are similar to indexes, but a short and often useful summary of each document (an abstract) is also given. Indexes and abstract journals may be in book form, although in the United Kingdom and some other countries they are now available on CD-ROM for computer use. In libraries where CD-ROM facilities are available, a librarian will usually show new users how to use it.

You may also find relevant articles by scanning through copies of journals. Some journals are more likely to provide relevant articles than others. For example, the *Journal of wound care* might be expected to contain articles relevant to the topic identified above.

2.4 Obtain selected articles, books and reports

Abstracts will help you to identify which articles, books and reports are most relevant and worth obtaining. The documents obtained may provide further references that can be followed up.

A librarian may be available to help you find your way around the library, to understand its classification system and to obtain other materials not kept in the library.

Books and reports will usually have to be read in the library, although some libraries may allow you to borrow them. You may find it useful to make notes on them.

The United Kingdom operates an inter-library loans system, which enables copies of books and reports and photocopies of articles to be obtained from other libraries. The librarians at major libraries in your country should also be able to obtain journal articles through this system.

2.5 Keep a record of each document

Build up your own index of relevant documents. It is a good idea to write the details of each document on a separate card or sheet of paper and file them alphabetically (by author's name) in a box, folder or other suitable holder. Use only one side of the card or sheet. The details could also be kept in a computer database.

You need to include specific details that identify the document and its source. These details constitute a "reference". For a journal article you need:

- the name(s) and initial(s) of the author or authors;
- the year of publication;
- the title (and any subtitle) of the article;
- the name of the journal in which it was published;
- the volume and part numbers of the journal in which it was published; and
- the inclusive page numbers.

For a book you need:

- the name(s) and initial(s) of the author or authors;
- the year of publication;
- the title (and any subtitle) of the book;
- (for any book which is a second or subsequent edition) the number of the edition;
- the place of publication; and
- the name of the publisher.

Reports, theses and specific chapters in books are referred to in a similar way. References are discussed in greater detail on page 14.

Examples and Exercise 1

Look at the reference list on page 71. Ref. 9 is an example of a reference to an article. Ref. 1 refers to a book, Ref. 4 to a chapter in a book and Ref. 16 to a report.

Look at the other references in the list. Decide whether each reference refers to an article, book, chapter or report. Can you identify all the essential details of each reference, as given above?

On your card or sheet of paper or in your database it may also be useful to record where you found the reference (e.g. in an abstract journal or in the reference list at the end of a specific article), where you can find a copy of the document (e.g. in which library, in which section of a particular library, or in your own collection of photocopies) and whether you have read it. For a book, it is also useful to include the International Standard Book Number (ISBN), a ten-digit number usually found on the back cover of a book and (along with other details) on the reverse of the title page. This may be particularly useful if you decide to order a book through a bookshop or by mail order.

After you have read the document you could write details about its content and your critical evaluation on the reverse of the card or sheet, or add them to your database.

3. Reading and evaluating research reports

Exercise 2

Using the information given in Section 2, obtain the three following journal articles.

ALLEN, M. ET AL. Effectiveness of a preoperative teaching programme for cataract patients. *Journal of advanced nursing*, **17**: 303–309 (1992).

MCBRIDE, A. Health promotion in hospitals: the attitudes, practices and beliefs of hospital nurses. *Journal of advanced nursing*, **20**: 92–100 (1994).

KING, K.M. & JENSEN, L. Preserving the self: women having cardiac surgery. *Heart and lung*, **23**(2): 99–105 (1994).

Attempt to read these before you read the rest of Section 3, but do not worry if you find them difficult to understand at present. You will be asked to refer to these articles as you progress through Section 3. You will also be given some exercises relating to them. By the end of Section 3 you should have a better understanding of the three articles and of research.

3.1 Why critical evaluation is important

As noted in Rundell's article (9), not all documents that report research findings are of high quality; some have major flaws that weaken or invalidate the claims made by the authors. Some minor flaws may, however, be unavoidable and may not undermine the worth of otherwise rigorous research. An ability to identify both the strengths and weaknesses of research documents will enable you to:

- identify good quality research documents relevant to your clinical situation;
- recognize credible material that supports or challenges insights that you may have gained through rational thought, intuition, education or clinical experience;
- make informed decisions about appropriate care, rather than relying on traditional practices or instructions from others, and implement these in consultation with patients; and
- discuss the worth and relevance of research findings with nursing colleagues and other health workers in your clinical area.

3.2 Factors for critical consideration

When reading documents critically you should pay particular attention to the following:

- the biographical details of the researchers
- the source(s) of research funding
- details given in the abstract (if any)
- the aims of the study
- the literature review
- the choice of research methods
- pilot work
- sampling procedures

- data collection methods
- data analysis
- presentation of data
- discussion of data
- conclusions and recommendations
- references
- ethical considerations
- rigour
- relevance to clinical practice.

These factors are discussed individually below. Terms shown in ***bold italic*** are defined in the Glossary (Section 5, page 49).

3.3 Biographical details of the researchers

Details about the researchers' qualifications, current posts and/or academic background are sometimes provided. This is good research practice, as it facilitates critical consideration of the claims made in the document. For example, a team of doctors researching nurse education may employ different research methods and/or come to different conclusions than a team of nurses or a mixed team of doctors and nurses. Similarly, a member of the public and a nurse may be likely to take different research approaches and interpret data differently.

3.4 Source(s) of research funding

Some small research projects conducted by practitioners do not have any outside source of funding. Much medical and nursing research, however, is funded by a variety of external agencies, including drug and other commercial companies, government departments and charities. Sources of funding should be explicitly stated in research documents so that readers can assess whether the nature of the funding agency might have influenced the choice of methods and/or the interpretation of the data. For example, it might be reasonable to be sceptical about research that claims that low-fat diets are unhealthy if it has been funded by a manufacturer of dairy products.

Example

If you look at the first page of each of the three articles you will see that all the authors have nursing qualifications, such as RN and RGN. Some also have academic qualifications such as PhD, MA and MN. The latter suggest that the authors have had training in research methods and/or experience of carrying out research. Some of the authors' current posts are also given. Only article 2 by McBride gives enough detail to enable us to see how the current post of the author relates to the subject of the research.

Only article 3 by King & Jensen acknowledges a source of research funding. We do not know whether the research projects described by Allen et al. (article 1) and McBride had any external funding and, if they did, which organizations provided it. Nevertheless, there are no obvious ways in which the source of funding might have affected the conduct or findings of the three research projects. If, for example, the project described by Allen et al. had tested a commercially available teaching package, we might have wondered whether the study had been funded by the company producing the package or by a commercial rival.

3.5 Details given in the abstract

Not all research documents include an **abstract** (a short summary near the beginning). Abstracts are, however, always included in articles published in good quality health journals such as the *Journal of advanced nursing* and the *British medical journal*. University theses include an abstract, while some reports open with an executive summary (similar to an abstract). Books do not usually contain an abstract, although multi-author works sometimes have an introductory chapter by the editor(s) giving an overview of the material in the book.

An abstract should provide an accurate summary of the content of the research document and should outline the main features of the research, such as the aims, methods, findings, conclusions and recommendations. Abstracts are useful in two main ways.

1. An abstract can often provide a quick assessment as to whether a document is relevant to your topic of search or your clinical practice. An abstract can therefore save time (and sometimes money), as it can help you to decide whether it is worth obtaining and/or reading a document.
2. An abstract may provide indications that a document has serious methodological flaws, or that the authors have reached far-fetched conclusions. Such suspicions can then be confirmed or rejected by careful reading of the main body of the document.

Example

All three articles reproduced at the end of this workbook have an abstract at the beginning. Read these three abstracts through. There are obvious differences in the style of presentation and the amount of detail given.

The abstract of the article by Allen et al. is extremely short. It does not make the aims of the research or sampling procedures clear. Research design and data collection methods are not mentioned. It suggests that the findings are unequivocal (this will be discussed later).

The abstract provided by McBride is much longer and this is reflected in the greater detail provided about aims, sampling, data collection, findings and recommendations. It does not, however, give a specific name to the research design.

The abstract of the article by King & Jensen is the clearest of the three. This is partly because it is presented under separate headings (a presentational feature required by the journal publishing the article). We are given clear details under the headings Objective, Design, Setting, Participants, Results and Conclusion.

You will better understand the issues involved in producing a good abstract as you read through the rest of Section 3.

3.6 The aims of the study

The aims of the research should be made clear early in the document (usually in a section entitled Introduction or Background). The aims will probably have been influenced by the researchers' previous experience and/or study of relevant literature and may be stated in terms of objectives, research questions and/or hypotheses.

Aims

The aims of a piece of research are often expressed in a broad statement of intent, such as a description of the factors that may influence clinical outcome and satisfaction in patients treated for urinary incontinence.

Objectives

Objectives break aims down into specific elements to determine:

- what treatments patients have received for their incontinence;
- whether each treatment was given by a doctor, nurse, physiotherapist or other worker;
- the perceived effectiveness of each treatment;
- patients' satisfaction with each treatment;
- factors contributing to satisfaction or lack of satisfaction with each treatment; and
- the clinical outcome of and patients' satisfaction with their overall treatment.

Research questions

The term “research question” is sometimes used synonymously with “research topic” or “research problem”. More often, however, it is used to describe a specific question to be answered by the research. There may be a number of these questions for any one research project. The aims and objectives above could be expressed in terms of research questions of the following type.

- What are the factors that influence clinical outcome and satisfaction in patients treated for urinary incontinence?
- What treatments have patients received for their incontinence?
- Was the treatment carried out by a doctor, nurse, physiotherapist or other worker?

Hypotheses

Certain questions may be phrased as research hypotheses. A *hypothesis* is a tentative statement about the relationship between two or more variables. The formulation of hypotheses is characteristic of quantitative rather than qualitative research (these types of research are defined in the discussion on choice of research methods below).

A *variable* is any characteristic of a person, thing or situation that can vary. Age, weight, type of medication taken, smoking behaviour, life experiences and beliefs, for example, are all variables. In a hypothesis, a variable may be described as independent or dependent. The *independent variable* is the “presumed cause” and the *dependent variable* is the “presumed effect” (11). In the study on incontinence described above, factors related to treatment would be independent variables and clinical outcome and patient satisfaction would be dependent variables. In some studies, researchers manipulate the independent variable in order to produce an effect on the dependent variable.

Examples of research hypotheses are that:

- wounds heal more quickly by using lotion A than by using lotion B; and
- cigarette smokers have higher rates of lung cancer than nonsmokers.

In the above examples, type of lotion and smoking behaviour are the independent variables, while speed of wound healing and rates of lung cancer are the dependent variables. In

the first example, the researchers aim to manipulate the independent variable by using two different types of lotion that can then be compared. Certain variables may need to be defined. For example, when is a wound to be considered healed; and is a woman who has smoked one cigarette in her entire life to be classified as a smoker?

If a hypothesis is to be subjected to statistical testing, a *null hypothesis* is sometimes formulated. This process is described in the discussion on data analysis below.

Exercise 3

Read the Introduction to Allen et al., the subsection marked Aims in McBride and the opening (untitled) section of King & Jensen.

How easily can you find the aims of each study and how clearly are the aims expressed? Are the aims stated in terms of objectives, research questions or hypotheses?

Make some notes on your answers to these questions and then read the relevant discussion in Section 6.

3.7 The literature review

Reviewing the literature allows researchers to set their research topic in the context of existing data and ideas. This enables them to demonstrate how their own work has added to knowledge on the subject. The literature review is usually presented in a specific section of the research document, but sometimes forms part of an introductory section or chapter. The review should discuss material relevant to the aims of the research.

Occasionally, a literature search may reveal insufficient (or no) relevant nursing literature. In such cases the researchers should make this lack of nursing literature explicit, and may choose to review parallel literature from another discipline such as medicine or education. Similarly, international literature may be reviewed when there are no relevant documents from the researchers' own country. Matters of context, such as differing work practices in medicine and nursing or the differences in health systems existing in different countries, should be acknowledged and their possible effects discussed.

A good quality literature review should concentrate on reviewing original research documents (known as "primary sources") rather than research described in other literature reviews ("secondary sources"). If all the references are more than five years old this may suggest that the researchers have failed to carry out a recent literature search or, more worryingly, that they have ignored recent documents because they provide new data that conflict with the researchers' beliefs. It may, however, be difficult to determine the thoroughness of a review unless you are an expert on the research topic or you conduct your own search to check for comprehensiveness. A thorough literature review should attempt to present a balanced view by discussing the conflicting ideas of different researchers and authors.

A literature review should not be just a list of documents published on a certain subject. As noted above, the ideas in different documents should be compared and contrasted. There may be gaps in certain areas of the literature that the authors should acknowledge. Sometimes the

researchers announce an intention to fill some of the gaps with their research findings. In addition, there should be a critique of the quality of the documents reviewed; significant methodological flaws, for example, may detract from the worth of data, conclusions and recommendations. The ability to critique research documents depends largely on skills developed through research awareness.

Referencing systems

There are two main systems in use for citing and referencing other work: the Harvard system and the numerical system.

In the *Harvard system*, references are cited in the text by giving the name(s) of the author(s) followed by the year of publication. This may be done in one of two ways:

- Polit & Hungler (1989) pointed out that ...
- It has been observed (Briggs, 1972) that ...

In the reference list, the references are set out in alphabetical order of the first author. If there are two or more references by the same author(s), these are given in chronological order. If there are two or more by the same author(s) in the same year, these are designated (and also in the text) as, for example, 1998a, 1998b, etc.

In the *numerical system* (as used in this workbook), the references are numbered consecutively as they occur in the text. The names of the author(s) may or may not be given:

- Polit & Hungler (11) pointed out that ...
- It has been observed (5) that ...

In the list, the references are set out in numerical order.

Note that where there are several authors many publishers prefer to use “et al.” in the text to denote that there are more than a certain number, commonly two or three. Some publishers also use this system in the list of references.

You may also come across the use of the term *Vancouver system* or *Vancouver style*. This is not a referencing system as such (although it uses the numerical system) but part of a set of standards, formulated by the International Committee of Medical Journal Editors, on how authors should prepare manuscripts for publication.

Exercise 4

Can you locate the literature review in each of the three articles? Do the reviews discuss literature directly relevant to the aims of the research? Are the documents reviewed evaluated in terms of quality? Are the ideas in different documents compared and contrasted? Are most of the references recent (from the last five years)? Do the three articles use numerical or Harvard referencing systems?

Write a few notes in response to these questions and then turn to the relevant discussion in Section 6.

3.8 The choice of research methods

The aims, research questions and hypotheses should be restated at the start of the Methods section of a research document. This should help you to assess whether the methods chosen by the researchers were appropriate to meet the aims of the study.

Quantitative or qualitative research?

One of the choices that researchers may need to make is that between quantitative and qualitative methods. Put simply, quantitative research seeks to answer questions of “how much?” and “how many?” and is concerned with the relationships (especially causal relationships) between variables, whereas qualitative research is concerned with a matrix of “who, why, when and where?” questions and with the exploration of issues.

Quantitative research often takes the form of experiments, quasi-experiments and surveys, which generate numerical data for tabulation and/or statistical testing. Qualitative research utilizes unstructured or semi-structured interviews, participant observation and/or the analysis of documents and conversations, often in combination. Qualitative data are in the form of words, such as interview transcripts and researchers’ field notes. Data are assigned to thematical categories, which are used as a basis for discussion.

Quantitative research is associated with a philosophy known as positivism, which is based on ideas about the existence of objective reality and the ability of human beings to discover facts if they set aside their personal biases and take a scientific approach. Positivist philosophy and quantitative method underpin most research into the natural sciences (physical, chemical and biological experiments) and medical research (drug trials) and also some social science research (psychological experiments and social surveys).

Qualitative research draws on a number of theories that can be loosely grouped under the term “interpretivism”. Interpretivists assert that, as human beings, we can never achieve complete objectivity and that we have to rely on a process of subjective interpretation in order to construct knowledge. Researchers working within an interpretivist framework believe that it is impossible for human beings to set aside their biases but that they can still be scientific (systematic and rigorous). Qualitative studies are sometimes described as ethnographic (exploring social/cultural interaction and meanings) or phenomenological (focusing on individuals’ experiences).

Quantitative researchers are sometimes disparaging about qualitative research and vice versa. Some quantitative researchers describe qualitative research as unscientific, biased or anecdotal, and claim that it produces no “real knowledge”. Qualitative researchers, on the other hand, may believe that quantitative research seeks to reduce complex issues to simplistic explanations, fails to explore contextual matters and takes insufficient account of the effects of the researcher on the process and products of research. In addition, qualitative researchers might comment that, since all knowledge is based on subjective deliberations, there is no such thing as “real knowledge” but that quantitative researchers fail to acknowledge this.

In general, doctors have tended to concentrate their efforts on quantitative research, while nurses have undertaken a mixture of quantitative and qualitative work. Many nursing researchers are in favour of multi-method research (sometimes known as triangulation), which has both

quantitative and qualitative elements. Some, notably Leininger (12), do not favour this because of the philosophical gulf between the two approaches.

A qualitative exploration of a topic may generate hypotheses that can then be subjected to statistical testing in a quantitative study. Conversely, statistical relationships demonstrated in an experiment or a survey may suggest that a number of interrelated factors need qualitative explication. Some of the major differences between quantitative and qualitative research are presented in Table 1. You will learn more about these differences as you progress through Section 3.

Table 1. Key differences between quantitative and qualitative research

	Quantitative research	Qualitative research
Underlying philosophy	Positivism	Interpretivism
Underlying ideas about reality	Reality is single and stable	Realities are multiple, complex and ever-changing
	Objective facts exist	What is "real" is subjective
	Competing explanations for facts must be eliminated through control or acknowledged as limitations and flaws	Complexity and contradictions cannot be explained away but should be explored
Aims	To make predictions and laws	To explore phenomena and relationships To promote understanding
	To produce data that are generalizable to other situations	To produce descriptions and theories that readers may find relevant to their own situations
Research techniques	Control Standardization	Flexibility Consistency
Research relationships	Relationships with "subjects" must be limited and standardized to promote control and objectivity	Relationships with "participants" are interactive The process and effects of interaction are often explored
Sampling	Statistical Random	Purposive: to aid exploration
Order of events	Sampling Data collection Data analysis	Sampling, data collection and data analysis are concurrent
Data	Numbers Statistics	Words Ideas

Some of the major forms of quantitative and qualitative research are described below. Experimental, quasi-experimental and survey work, all forms of quantitative research, are described first.

Examples

The articles by Allen et al. and McBride describe quantitative research studies, while that by King & Jensen describes a qualitative study.

Experimental research

Experiments are a powerful way of testing cause and effect relationships between variables. Polit & Hungler (11) describe the three characteristics of a true *experiment*. (Note: people participating in experiments are sometimes described as “subjects”.)

1. Manipulation. The experimenter must do something to at least some of the subjects in the study.
2. Control. The experimenter must introduce one or more controls over the experimental situation, including the use of a control group.
3. Randomisation. The experimenter must assign subjects to a control or experimental group on a random basis.

True experiments are sometimes known as randomized controlled trials.

In the discussion on the aims of the study (page 11), you read about independent and dependent variables. In experimental research the researchers manipulate the independent variable (the presumed cause) in order to observe whether this has an effect on the dependent variable. The independent variable may involve a treatment, an information leaflet, a training course or any other device (sometimes termed the *intervention*) that might produce an effect on the dependent variable. For example, a treatment is given to some patients (known as the *experimental group*) and withheld from others (known as the *control group*) in order to see whether those who receive the treatment demonstrate a greater improvement than those who do not. Relevant data are collected from all the patients both before and after any treatment is given (*pre-test* and *post-test*).

Since improvement may occur spontaneously, without any treatment being given, an experiment must involve some form of comparison. The ability to compare is provided by the control group (11): “The term *control group* ... refers to a group of subjects whose performance on a dependent variable is used as a basis for evaluating the performance of the *experimental group* (the group that receives the treatment of interest to the researcher) on the same dependent variable”.

The control group may be given the conventional treatment for the condition or a placebo (such as tablets with no active ingredient or a half-hour visit but no therapy) or no treatment at all. Some study designs involve more than one control group such as:

- experimental group vs. placebo group vs. no treatment group; or
- experimental group vs. conventional treatment group vs. placebo group.

The latter design has the ability to demonstrate that the experimental treatment is not only more effective than the placebo but is also more effective than the conventional treatment.

Other factors that could affect the dependent variable need to be controlled. For example, the temperature in a laboratory or room might affect biochemical, physical or psychological test results. For this reason, the temperature needs to be constantly controlled throughout the experiment. In a drug trial, patients might take other medicines that could affect the medical condition under study. Such patients must either refrain from taking these medicines or be excluded from the study. A study that has body temperature as a dependent variable must take account of the fact that this varies with the time of day. Body temperature should, therefore, be

measured at the same time (or times) of day. These additional factors that can affect the dependent variable are known as *extraneous variables*.

Randomization, the third essential element of an experiment, is performed to prevent the researcher(s) from distorting the results through “systematic bias”. The random assignment of individuals to groups aims to make the experimental group and the control group comparable in terms of characteristics such as age, gender, educational status and other factors that could affect the dependent variable. Randomization procedures can be performed by simple means, such as drawing names out of a hat or generating random numbers by computer.

There are several types of experimental design. Two common ones are:

- the comparison of independent groups; and
- the comparison of repeated measures carried out on the same group.

In the latter case the experimental group acts as its own control group. The “cross-over” study is a common form of the repeated measures design. In this, participants are divided into two groups. Group I receives treatment A followed by treatment B and group II receives treatment B followed by treatment A. The advantages of repeated measures studies are that they require fewer participants (as each individual is used twice) and that the characteristics of the experimental group and the control group are the same (as they are composed of the same people). The major disadvantage is that it can be claimed that data collected during the second treatment period have been influenced by the fact that there was a previous treatment period.

It is possible to carry out some experiments, such as drug trials under *double-blind* conditions. This means that neither the researchers nor the participants know which is the experimental group or treatment and which is the control group or treatment. Thus, in a double-blind cross-over drug trial, neither the researchers nor the participants know whether the experimental medication has been administered during the first or the second treatment period (placebo medication is given in the other treatment period). The study is only “unblinded” once data collection and analysis are complete.

Quasi-experimental research

A *quasi-experiment* does not possess all three essential attributes of a true experiment, as it lacks either randomization or a control group. There are numerous questions worth researching where randomization is not possible. For example, a researcher may wish to determine the physical and psychological effects of a new shift-work system on nurses working in a particular hospital. Since the new shifts are being introduced for all nurses throughout the hospital, randomization is not possible. It may, however, still be possible to collect pre-test and post-test data and to use the nurses at a similar hospital, which is keeping the old shift system, as a *comparison group*. In some studies, however, it may not be possible to find an appropriate comparison group.

Quasi-experimental designs do not lend themselves to the testing of hypotheses concerned with cause and effect because of the presence of too many uncontrolled factors. It is, however, possible to test whether variables are correlated, i.e. related to each other in a statistically significant way.

Survey research

Surveys can be used to collect information about variables and, sometimes, to test relationships between variables. There is no experimental intervention in survey work; rather information is

collected about (for example) people's health status, living conditions, health beliefs or satisfaction with services. The information is turned into numerical data, which can be subjected to statistical analysis.

Survey data are obtained by means of questionnaires or interviews, using a highly structured interview schedule. Interviews may be conducted face-to-face or by telephone. You can read more about questionnaires and structured interviews in the section on data collection methods on page 24.

Some surveys are designed to gather information from the entire *population* of (for example) a country or town or all the patients who attend a particular health clinic. A survey of an entire population is known as a census. Censuses may be costly and time-consuming to conduct, however, especially if the population under study is large or widely spread. For this reason, it is more usual for part of the population to be surveyed (known as *sampling*). Assumptions are then made about the health, living conditions, etc. of the entire population on the basis of the data provided by the sample. Issues relating to the size and representativeness of survey samples are discussed in the section on sampling procedures (page 21).

Exercise 5

Read the section entitled The study in the article by Allen et al. and that entitled The research project in the article by McBride. Do these describe experiments, quasi-experiments or surveys?

Write down your answers and then turn to Section 6 for discussion.

In the article by Allen et al., what were the dependent variables? Can you identify any uncontrolled extraneous variables? Was it a double-blind study?

In the article by McBride, what was the study population?

Write down your answers and then turn to Section 6.

Qualitative research

It is fairly easy to divide quantitative research designs into methodological types, such as surveys and experiments, but this is not the case with qualitative research. Qualitative research is associated with a number of overlapping philosophies and methodologies such as naturalism, interpretivism, ethnography and phenomenology. It is not uncommon for qualitative researchers to draw on all of these and it may not be easy, for instance, to determine why a study has been described as ethnographic rather than phenomenological.

Hammersley & Atkinson (13) give a definition of naturalism:

Naturalism proposes that, as far as possible, the social world should be studied in its 'natural' state, undisturbed by the researcher. Hence, 'natural', not 'artificial' settings like experiments or formal interviews, should be the primary source of data. Furthermore, the research must be carried out in ways that are sensitive to the nature of the setting.

Interpretivism incorporates the above ideas but also stresses that, even when the social world is studied in its "natural" state, researchers can only understand it through their own five senses and a process of subjective interpretation. Because of this, researchers never have direct

access to the experiences, understandings and meanings of the people under study. Thus, research documents can represent the views of the people under study but only as interpreted by the researchers.

As noted towards the beginning of this section, ethnography (a term with roots in sociology and social anthropology) is primarily concerned with social interaction and meaning, whereas the main focus of phenomenology are the experiences of individuals. Some researchers, however, use the terms “ethnographic” and “phenomenological” as synonymous with “qualitative”.

A qualitative approach can enable researchers to explore issues in depth. An experiment may show that a medication is clinically effective, but it is less suited to exploring the experiences, opinions and feelings of patients who have had to decide between having an operation or taking long-term medication for a chronic illness. Data from a quantitative survey may suggest that there is a statistical relationship between urinary incontinence and low self-esteem, but an in-depth qualitative approach may be required to draw out how these two factors relate to each other and how the lives of incontinent people might be improved in both physical and emotional terms. Relatively unstructured data collection methods, such as participant observation and semi-structured interviews, are used in qualitative research so that everyday experiences can be studied and research participants can put forward their own ideas in their own words.

Qualitative researchers (particularly ethnographers) often draw on *grounded theory* (14) when undertaking sampling procedures, data collection and data analysis. This ensures, as far as possible, that the theories that the researchers construct are grounded in the data obtained (rather than in the personal interests or the previous knowledge of the researchers). Grounded theory is discussed further below in the sections dealing with sampling procedures, data collection methods and data analysis.

Exercise 6

Read the Method section of the article by King & Jensen. Is the study described as qualitative, ethnographic or phenomenological?

Turn to Section 6 for discussion.

Other forms of research

Research can be described and categorized in various other ways, some of which cross the quantitative/qualitative divide. Two of these types of research will be discussed briefly: feminist research and action research.

Feminist research is particularly relevant to nursing because the vast majority of nurses are women. In addition, because women often take care of children and dependent relatives they come into contact with the health care system on behalf of others as well as themselves. Feminist research can be either quantitative or qualitative. It frequently focuses on gender relations between doctors and nurses, gender issues in the care and treatment of women patients, and issues related to caring.

Action research is an approach in which researchers and research participants work together to solve problems and effect change. Data are collected at the start of the study and used as a basis for planning change. The process of change is monitored and plans modified if necessary. At the end of the study a final assessment is made. The process (as well as the outcome) of change is important in action research: during it researchers enable participants to learn about research and to gain control of their work situation or their care. Action researchers often take a multi-method approach combining quantitative and qualitative methods.

Example

The article by King & Jensen provides an example of feminist research, although this is not explicitly stated in the article.

3.9 Pilot work

It is good practice in quantitative research (and sometimes in qualitative research) to carry out a *pilot study*. This can help to eliminate problems related to access, sampling and data collection instruments that might otherwise arise in the main study. Polit & Hungler (11) describe a pilot study as “a small-scale version, or trial run, of the major study”. Sometimes, if the problems arising are minor, data from piloting work can be included in the main study data. This frequently happens in qualitative studies where, in any case, data collection instruments tend to evolve as the main study progresses. Research documents should state whether a pilot study was performed and give some idea of any problems encountered and subsequent modifications made to plans for the main study.

Exercise 7

Refer to the three articles. Did the researchers involved in the three studies conduct any form of pilot work? If so, what modifications were made for the main studies?

Note down your answers and then turn to the discussion in Section 6.

3.10 Sampling procedures

Sampling procedures tend to be quite different in quantitative and qualitative research. Whereas the aim of sampling in quantitative research is to select a sample that is representative of the study population, the aim in qualitative research is to select participants who will allow a detailed and in-depth exploration of the issues under study.

Sampling in quantitative research

Probability sampling is the most effective way of selecting a sample that is representative of the population from which it is drawn. In probability sampling every element (person or thing) in the population under study has an equal chance of being selected for the sample. There are three main types of probability sampling.

1. **Simple random sampling** uses a computer, a **random number table** or numbers drawn at random to select a sample from the **sampling frame** (a numbered list of the entire population under study).
2. **Stratified random sampling** divides the population into two or more **strata**, such as male/female or nursing assistant/student nurse/qualified nurse, and a random sample is then drawn from each stratum. Stratified sampling can be proportional or disproportional to the population under study. Proportional stratified sampling is used to make the sample more representative, whereas disproportional sampling can ensure that minority groups are represented in sufficient numbers to allow meaningful statistical analysis.
3. **Cluster sampling** can be used in large-scale studies where compiling a sampling frame of the entire population might prove costly, difficult or impossible. Compiling a sampling frame of all the diabetic patients in the United Kingdom, for example, would be difficult and expensive. It would be easier and cheaper to select a random sample of health clinics, thus giving access to “clusters” of diabetic patients, all of whom (or a random sample of whom) could be selected depending on the sample size required.

It is important, when using a sampling frame, to be sure that it indeed represents the population that the researchers wish to study. For example, a register of nurses registered *to practise* is not a register of nurses *in practice*. As implied above, there are no existing sampling frames for some populations. Compiling a nationwide database of diabetic patients or nursing assistants is a major piece of work in itself.

Because of the expense, time and difficulty that may be involved in probability sampling, **non-probability sampling** is often employed. This has the disadvantage of being less representative. There are three main types of non-probability sampling.

1. **Convenience sampling** adopts the most readily accessible people or things as the sample. Thus, a nurse performing a survey may decide to distribute questionnaires to the first 200 patients with leg ulcers seen at home or in a clinic. Some experiments, such as drug trials, may utilize convenience samples: all available patients meeting the inclusion criteria and agreeing to participate are recruited. Randomization occurs only at the point of assignment to the treatment or control group.
2. Snowball sampling employs a questionnaire or a request to participate. This is sent to known contacts, who are asked to provide names of other people who will agree to participate. These, in turn, are asked to provide further names and so on.
3. **Quota sampling** involves the researchers identifying strata in the population and selecting a convenience sample that reflects the proportions of those strata. Thus, if a particular population is composed of 400 males and 600 females, researchers might aim to survey 40 males and 60 females. It might also be composed of 500 nurses, 250 physiotherapists and 250 doctors, in which case 50 of those chosen would be nurses, 25 physiotherapists and 25 doctors. The aim of quota sampling is to make more representative a sample that would otherwise be a convenience sample.

When critiquing non-probability sampling designs in research documents, you should consider whether the researchers might have been able to select a more representative sample by using an existing sampling frame (such as doctors’ lists of patients or a national register of nurses) and employing random sampling techniques.

Determining a suitable sample size can be difficult. The larger the sample the more likely it is to be representative of the population from which it is drawn. However, the larger the sample the greater the resources needed for data collection and analysis. Where limited resources have necessitated the adoption of a small sample (and/or of non-probability sampling) this should be acknowledged in the research document. If a small, non-probability sample has been used, researchers should be extremely cautious about claiming that their findings can be extrapolated to the entire population.

Whenever possible, and prior to sample selection, researchers should take expert advice about sample size from a statistician (this too should be noted in the research document).

Exercise 8

What types of sample were used in the studies described by Allen et al. and McBride? Do you think that the sizes of the samples were adequate? Do the authors say whether they took statistical advice on sample size?

Make some notes and then turn to the discussion in Section 6.

Sampling in qualitative research

Qualitative researchers rarely employ random sampling techniques, but neither do they claim that their data and interpretations are generalizable to larger populations. They do, however, attempt to highlight and explore issues that may be interesting and relevant to people in settings and contexts other than that in which the research was carried out. Thus, a study of the morale of staff at one hospital may be of interest to staff at a second hospital; an exploration of doctors' communication skills may throw up interesting points about communication in nursing; cancer patients may identify with some of the experiences, thoughts and feelings presented in a study of AIDS patients.

In qualitative studies, sampling is often described as *purposive* or *judgement sampling*. This means that sampling has been conducted with a particular purpose in mind or on the basis of the researchers' judgement. In other words, both terms mean that researchers choose a sample that should enable a thorough exploration of issues relevant to the aims of the study.

Sampling decisions in qualitative research include the selection of a setting or settings that will allow exploration of the relevant issues. For example, researchers might be interested in how various aspects of policy, practice and staff relationships affect the recovery of patients who have undergone surgery for cancer. They may decide to study just one hospital ward or two or more wards at different hospitals, which can then be compared and contrasted in respect of the issues under study. Once a setting or settings have been chosen, there are decisions to make about who to interview and who or what to observe within the setting(s). This may involve decisions about, for example, which grades of nursing staff to interview, which doctors to interview, whether patients with particular types of cancer might have different things to say, whether to observe all ward rounds, or whether to observe staff in conversation during meal breaks.

"Theoretical" sampling is a specialized form of purposive/judgement sampling adopted by researchers who take a grounded theory approach to their work. In theoretical sampling, once a setting has been chosen, researchers make most of their sampling decisions as the research

progresses. This allows them to select interviewees, for example, who may facilitate exploration of theoretical concerns identified in the data already collected and analysed. Sampling can cease when no new concepts are being generated from the data or, to put it in grounded theory terms, when theoretical categories are “saturated”. This is one way in which sampling, data collection and analysis are linked in grounded theory.

When writing research documents, qualitative researchers should make the thinking (including any theoretical influences) behind their sampling decisions explicit.

Exercise 9

How are sampling procedures described by King & Jensen? Do you think there are any flaws in the sampling procedures described?

Make some notes and then turn to the discussion in Section 6.

Data collection methods

Quantitative and qualitative researchers also tend to utilize different data collection methods to pursue their aims. Quantitative methods are designed to promote standardization and control, whereas qualitative methods aim for a balance between flexibility and consistency (15).

Quantitative data collection methods

Data collection techniques (such as tests) and devices (such as questionnaires and interview schedules) are sometimes known as *instruments*. The most commonly used instruments in the collection of quantitative data are:

- tests and measurements (often used in experimental studies);
- questionnaires and structured interview schedules (often used in surveys); and
- structured observation schedules.

In well designed quantitative studies, data are recorded in tightly structured formats, which may be in the form of questionnaires or interview/observation schedules. In the case of questionnaires, data forms are completed by research participants (sometimes known as *respondents*), while in other cases the researchers complete the forms. It is good practice to include a copy of the survey questionnaire, observation schedule or other data form in the research document although, in the case of research articles, the need for brevity may prevent this. Inclusion of the data form may aid comprehension of the research document and will allow readers to evaluate the quality of data collection procedures and the *validity* and *reliability* of the study (see the section on rigour, page 42).

Experiments and quasi-experiments may involve the collection of several different types of data. For example, data collected during a clinical drug trial might include objective biophysiological measurements such as blood pressure levels and body weight, patients' subjective reports of their symptoms, and the presence or absence of any confounding variables such as other medication taken. Patients may be weighed and have their blood pressure taken each time they visit the person managing the trial. They may record their symptoms daily on a structured diary card (using, for example, the terms “mild”, “moderate” or “severe” to describe pain), which they hand in at each visit. They may be verbally questioned about their use of other

medication, concurrent illness and other potentially confounding variables. In other words, the data collected during the course of a clinical trial may involve biophysiological measurement, questionnaire and interview techniques. Data will be recorded on a data form at each visit. In a quasi-experiment designed to assess the effect of an educational package, data on nurses' knowledge may be collected pre- and post-test using questionnaires or structured interviews.

When critiquing experimental and quasi-experimental data collection techniques, you should attempt to assess whether data have been collected on potentially confounding variables. It is also important to consider ethical issues (discussed on page 40). This may be particularly relevant where biophysiological measurements are concerned, as some tests may be painful, distressing or even dangerous.

Exercise 10

What research instruments were used in the study described by Allen et al.? Are copies of these included in the article?

Turn to Section 6 for discussion.

Questionnaires may be distributed to a survey sample by post or, for smaller studies, by hand. In either case, attempts should be made to preserve respondents' anonymity. Individual questionnaires are usually identified by a code number rather than a name and, even where distribution has been by hand, questionnaires are often returned in a prepaid envelope. Questionnaires should be designed with a high proportion of closed ("yes/no" or "tick box") rather than open ("give details below") questions. Questions should be arranged in a logical order and should avoid complexity and ambiguity. Pilot work (as discussed above) can help researchers to identify and eliminate questionnaire design flaws.

Structured interviews are similar to questionnaire surveys, except that researchers (sometimes called "data collectors") obtain the data from participants by reading questions from and recording answers on an interview schedule. This can be done face-to-face or over the telephone. To preserve a degree of anonymity, completed schedules usually bear a code number rather than a name. The issues involved in designing an interview schedule are very similar to those involved in questionnaire design. Although data collectors are able to clarify questions that participants do not understand, they must be careful that they do not bias the data when doing so. Each participant should be asked the questions in a standardized way. This is extremely difficult to achieve in practice, particularly when there are two or more data collectors. Training may help to minimize some of these problems.

Some research topics might be better dealt with by questionnaire, some by telephone and some face-to-face. For example, people might be willing to complete a questionnaire about a potentially embarrassing subject, such as incontinence or impotence, but might not wish to talk about it. As factors such as this may affect the *response rate* (see under Data analysis, page 27) as well as the quality of the data, they should be taken into consideration when critiquing research documents.

Exercise 11

Is the research instrument included by McBride? What conclusions can we draw about the quality of the instrument?

Turn to Section 6 for discussion.

Observational research takes two main forms: participant observation, which is usually employed in qualitative studies (discussed below), and non-participant observation, which is more common in quantitative work and may be known as “structured observation”. Structured observation can take several forms, all of which involve the assignment of behaviour to categories recorded on an observation schedule. Activity sampling is a common form of structured observation: for example, every 15 minutes the observer records what each member of the ward staff is doing. When selecting a “time frame” for activity sampling, researchers must avoid choosing an interval that is too short to allow the observer to record all the necessary data, or is too long so that important events are missed during non-observation periods. Other difficulties with observation include the problem of **reactivity**: people may alter their behaviour if they know they are being observed (reactivity is also an issue in experiments and interviews and is sometimes known as the **Hawthorne effect**). On the other hand, it may be unethical to observe people without their consent, because they have a right to decide whether they wish to participate (see the section on ethical considerations, page 40). All these factors should be taken into consideration when critiquing observational studies.

Exercise 12

Did the study by Allen et al. involve an observational component in the assessment of skills?

Turn to Section 6 for discussion.

Qualitative data collection methods

Qualitative researchers, particularly ethnographers, often “draw on a wide range of sources of information” and collect “whatever data are available to throw light on the issues” (13). Some qualitative studies may rely primarily on participant observation or semi-structured interviews, but researchers often use a combination of the two and may also draw on conversations, written documents (such as nursing care plans or policy statements) and other sources.

Participant observation is, perhaps, the most important data collection technique in ethnographic research. Observers may participate in the work, social activity or culture under study (for example, nursing work) to a variety of degrees. In addition, the degree to which they are “insiders” or “outsiders” to the research setting may vary. Complete insiders study their own work or social settings; complete outsiders are strangers to the types of setting they study. Researchers who study settings similar to their own can be considered as relative insiders/outsiders.

There are numerous difficult issues that can arise in relation to participant observation. One such issue is whether it is ethical to do covert research (pretending to be a participant and not a researcher). There is also a problem in uncritically identifying with and even adopting the culture

of a group of people under study. It may also be very difficult to manage relationships: some people, for example, can be so unhelpful or threatening that researchers choose to avoid them.

The research setting (sometimes known as the field) will have been selected as an appropriate place for exploring the issues of interest and, thus, fulfilling the aims of the research. Quite commonly, however, the issues under study (which are initially determined by researchers' previous experiences and/or their knowledge of the literature) undergo some degree of change as field work progresses. This is particularly the case with researchers who take a grounded theory approach: theories emerging from the data already collected determine who and what should be observed next. Participant observers record their experiences in the field, and their reflections on them, in *field notes*.

In ethnographic studies, data gained through observation may influence decisions about who to interview and the themes to be pursued at interview. Qualitative interviews are commonly described as *unstructured*, semi-structured, conversational or *focused*. They differ from structured interviews in that interviewers work from an interview schedule (or interview guide), which has topics for discussion or a few open questions that will get interviewees talking about the relevant issues. Interviewers aim for a balance between flexibility and consistency (15), allowing interviewees to introduce their own ideas and concerns while trying to pursue themes central to the aims of the research. The interview guide usually evolves during the course of a series of interviews. Lines of inquiry that have proved unproductive are dropped, while key themes introduced by previous interviewees are added. This, again, accords with a grounded theory approach. Qualitative interviews are often recorded to save interviewers from attempts to listen and write down large amounts of data simultaneously (each interview may last an hour or more).

In order to demonstrate rigour, a document describing a qualitative study should make the decisions involved in data collection procedures explicit. Issues such as failure to manage relationships in the field may affect the rigour of the work: this should be acknowledged by the researchers. In addition, there may be ethical issues to consider.

Exercise 13

What data collection methods were used in the study described by King & Jensen?
Are there any flaws in the data collection methods or the reporting of them?

Make some notes and then turn to the discussion in Section 6.

3.11 Data analysis

As data are analysed very differently in quantitative and qualitative studies, the two types of research will, again, be discussed separately.

Quantitative data analysis

Quantitative data are usually analysed using statistics. A full account of statistical analysis, (including the relevant mathematical calculations) cannot be given in this document. You may need someone with existing statistical knowledge to help you to understand research studies that have employed complicated statistical techniques such as multiple regression or factor analysis.

Statistics can be divided into two main types: *descriptive statistics* and *inferential statistics*. Before describing different types of statistics, however, it is necessary to discuss the concepts of levels of measurement and *frequency distribution*.

Measurement is the assignment of categorization codes to data. There are four levels of measurement that can be applied to data: nominal, ordinal, interval and ratio.

Nominal measurement is used for data that can be categorized by name only. In nominal measurement, although each category is assigned a numerical code or value, no category ranks higher or lower than another. Some examples are given below.

Sex: 1. male; 2. female.

Blood group: 1. O positive; 2. O negative; 3. A positive; 4. A negative; 5. B positive; 6. B negative; 7. AB positive; 8. AB negative.

Ordinal measurement categorizes data by rank as well as name. The distance between the ranks cannot, however, be quantified precisely. The categorization of pain as mild = 1, moderate = 2 and severe = 3 can be used as an example: although severe pain is greater than moderate pain, we cannot say exactly how much greater it is.

The next level is *interval* measurement. This is where there is a fixed numerical interval between each point on a scale but no meaningful zero point. The Intelligence Quotient test (like many educational and psychological tests) is based on an interval scale. A person scoring 120 scores higher than a person scoring 110 who, in turn, scores higher than a person scoring 100. The score of 110 is equidistant between the higher and the lower score but there is no meaningful zero on the scale (average intelligence is set at 100).

The highest level of measurement is *ratio* measurement. This is the same as interval measurement except that there is a meaningful zero on the scale. Age can be categorized in terms of ratio measurement, a person who is 35 being half the age of someone who is 70. Ratio comparisons of this type cannot be made about interval data.

Distinctions between different levels of measurement are important because they determine which statistical tests can be used on the data. This will be discussed in greater detail below. When critiquing research documents, you should decide whether the researchers have identified the levels of measurement appropriate to certain types of data. In particular, four- or five-point attitude scales ranging from “strongly agree” to “strongly disagree” are sometimes treated as if they are based on equal intervals. The resulting data may, thus, be treated as interval data when it is ordinal data.

Exercise 14

Read the Results section of Allen et al. and the rest of McBride up to the Conclusion. Can you determine the levels of measurement appropriate to the data collected in each of the two studies?

Write a few notes and then turn to the discussion in Section 6.

The frequency distribution is important because it determines whether *parametric* or *non-parametric statistics* should be applied to the data (these terms are discussed further below). A frequency distribution is the systematic arrangement of numerical values (category codes, scores, ages, etc.) showing how often each value occurs.

Table 2. Frequency table showing age distribution

Value (age in years)	Frequency	Percentage
20	2	8
21	0	0
22	3	12
23	6	24
24	2	8
25	0	0
26	0	0
27	4	16
28	5	20
29	3	12
Total	25	100

Table 2 is a frequency table. Frequency data can, however, be displayed in a number of other formats, the most important of which is the distribution curve. Only data that follow a *normal distribution* curve can be subjected to parametric tests. A normal distribution curve represents data in which most values are clustered around the centre of the distribution. Many human attributes, such as height, weight, head circumference and grip strength, have been found to conform to this description. A normal curve is bell-shaped, symmetrical (rather than “skewed”) and not too sharply peaked (Fig. 1).

Fig. 1. Normal distribution (age)



Having explained the importance of levels of measurement and frequency distribution, we can now turn to the two main types of statistics, descriptive and inferential statistics.

Descriptive statistics is used to describe and summarize the data collected. The main ways of doing this are through:

- percentages;
- measures of *central tendency* (the *mean*, the *median* and the *mode*); and
- measures of dispersion (the *range*, the interquartile range and the *standard deviation* (SD)).

For example, the age and sex of respondents to a survey might be summarized in the following way: 114 (29%) of respondents were male and 276 (71%) were female. The mean age of the respondents was 58 years (SD = 17.2, range 21–92 years).

Measures of central tendency are types of average.

- The mean or arithmetic average is calculated by adding up all the values and dividing by the number of cases, e.g. $2 + 4 + 3 + 4 + 7 + 4 = 24/6 = 4$.
- The median is the exact middle value in a distribution of values, e.g. for 3, 5, 6, 9, 11, 11, 12 the median is 9 and for 4, 6, 7, 10 it is 6.5.
- The mode is the value that occurs most frequently in a distribution of values, e.g. for 8, 12, 13, 13, 13, 17, 19, 19, 20, 21 the mode is 13.

Decisions about whether to use the mean, the median or the mode are governed by rules relating to levels of measurement and frequency distribution. Researchers should follow the following rules:

- for nominal measurement use the mode;
- for ordinal measurement use the mode or the median;
- for interval or ratio measurement use the mode, median or mean; and
- for data that do not follow a normal distribution curve use the median rather than the mean.

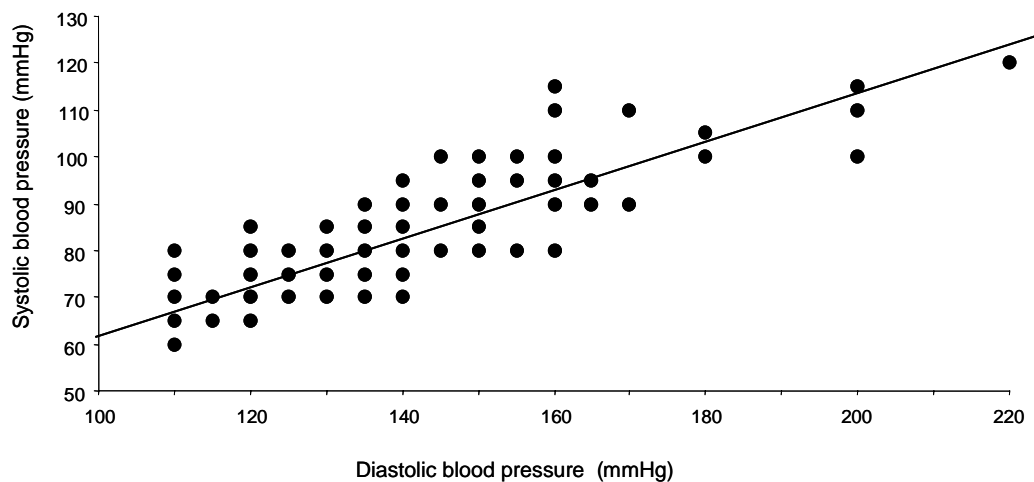
Similar rules govern the use of measures of dispersion (the range or spread of values):

- for nominal measurement use the range;
- for ordinal measurement use the range or the interquartile range;
- for interval or ratio measurement use the range, the interquartile range or the standard deviation; and
- for data that do not follow a normal distribution curve use the interquartile range rather than the standard deviation.

Measures of central tendency and dispersion are, like all statistics, based on mathematical calculations that can be done by hand or (as is more usual nowadays) with a computer package such as SPSS (Statistical Package for the Social Sciences).

Descriptive statistics can also be used to demonstrate the relationship between two sets of data, such as children's age and their head circumference. This can be done by plotting the two data sets along the two axes of a scatter diagram and seeing if they correlate in a positive or a negative way. The scatter diagram in Fig. 2 shows an upward trend indicative of a positive correlation, in this case between patients' systolic and diastolic blood pressures (a downward trend would indicate a negative correlation).

Fig. 2. Scatter diagram showing a positive correlation between systolic and diastolic blood pressure readings



The *correlation coefficient*, which expresses the degree to which two data sets are related, can be calculated by means of statistical tests.

Inferential statistics allows inferences to be made about a population on the basis of data obtained from a sample of that population. For instance, if researchers find that wounds heal up more quickly in a group of patients treated with lotion A than in a group treated with lotion B, they can use inferential statistics to test whether this is likely to apply to all wounds. They do this by testing the probability that faster wound healing occurred by chance rather than through the action of lotion A. Probability is a key concept in inferential statistics and is discussed in greater detail below.

Inferential statistical tests can be divided into parametric and non-parametric tests. Parametric tests have the advantage of being very powerful and very sensitive but cannot be used on data that do not follow a normal distribution curve. Non-parametric tests should be used in such cases.

Some commonly used tests are:

- the chi-squared (χ^2) test
- the *t*-test
- analysis of variance
- the Mann-Whitney U test
- the Kruskal-Wallis test.

Rules about which test to use are governed by factors such as frequency distribution, level of measurement, the number of groups in the sample and whether the groups consist of the same individuals (as in a crossover study), matched individuals or different individuals. They are thus too complex to discuss here. When critiquing a research document, however, you should note down which test the researchers used, along with details of the research design and the level of measurement and frequency distribution of the data. At some point you may get the opportunity to discuss the appropriateness of the test used with someone with statistical knowledge, or you may get access to a textbook on statistics. You should also note, when critiquing studies, whether the researchers took statistical advice on the research design and the tests used.

Exercise 15

Were descriptive and/or inferential statistical tests used in the studies described by Allen et al? Which specific statistical tests were used?

Write down a few notes and then turn to the discussion in Section 6.

Inferential statistics are frequently used for hypothesis testing. In the section entitled “The aims of the study” (page 11) you read about research hypotheses. When a hypothesis is to be tested, a negative statement known as the *null hypothesis* is often formulated. The null hypothesis states that there is no actual relationship between variables and that any apparent relationship is merely a function of chance. Thus a null hypothesis will state, for example, “The wounds of patients treated with lotion A will not heal any more quickly than the wounds of patients treated with lotion B”. Researchers hope that the results of statistical tests will allow them to reject the null hypothesis (or accept the research hypothesis).

Statistical significance is usually stated in terms of probability or *P* values. For researchers to be able to reject the null hypothesis there must be a low probability (such as 1 in 100) that any differences in the groups have occurred by chance. Researchers usually set the level of probability at which results of tests become statistically significant (the *significance level*) – and at which the null hypothesis can be rejected – at either 1 in 100 (usually expressed as 0.01) or 5 in 100 (0.05). To put it another way, the null hypothesis can be rejected when the probability that the differences occurred by chance is no higher than 1% or 5%. Where the probability is higher than 5% the null hypothesis must be accepted.

The *P* values of test results may be expressed in different ways. The actual value may be expressed, as in $P = 0.03$, or the value may be expressed in relation to the significance level, as in $P > 0.01$ or $P < 0.05$ ($>$ means greater than and $<$ means less than). You should remember that the lower the *P* value the greater the significance: $P = 0.003$ is highly significant, $P = 0.03$ is significant if the significance level has been set at 0.05, and $P = 0.09$ is not significant. When critiquing research documents you should also watch out for claims that statistical results are significant when they are not.

Exercise 16

Was a null hypothesis formulated in the study by Allen et al? At what level was the significance level set? Which results were statistically significant?

Write down your answers and then turn to Section 6 for discussion.

Qualitative data analysis

Qualitative data analysis does not involve statistical tests. Most qualitative researchers organize their data into theoretical or thematic categories. The data to be organized are usually in the form of field notes or typed transcripts of tape-recorded interviews (often both). Sometimes other documents such as policy statements or letters, and transcripts of everyday conversations, are also used.

In qualitative research, data organization does not occur before data analysis. The data must be partially analysed to enable them to be organized into categories. Once the data have been systematically organized, however, they can be analysed in greater depth. The organization and analysis of qualitative data are, therefore, closely linked. Categorized data may be organized manually on index cards or sheets of paper or may be entered into and stored on a computer with the aid of a program such as Ethnograph. For example, in a study on the telephone work of community nurses five thematic categories were developed. Each category was given a different colour, which was used to colour-code the data on interview transcripts. The five categories were:

- current practice (blue)
- future development (purple)
- communication (green)
- ideas about work (brown)
- ideas about being a community resource (red).

Data were later assigned to subcategories, using code numbers written in the appropriate category colour, and were then transferred to index cards filed by category and subcategory. Numbered subcategories in the green category included:

- being approachable
- the effect of facilities
- counselling/supportive work
- lack of nonverbal communication
- audible cues
- communication skills training
- on-the-job vs. formal learning
- telephone phobia (staff)
- telephone phobia (clients)
- talking to the telephone answering machine

There were 16 other subcategories.

Data could also be categorized using numbers for categories and letters for subcategories (e.g. 1a, 1b, 2a, 2b, etc.) and in a number of other ways.

With a grounded theory approach, data analysis is linked not only to data organization but also to sampling and data collection procedures (see above). Concurrent analysis or the “constant comparative method” is used: data analysis commences after the first interview or the first few days of observation. Decisions about who to interview, what to ask interviewees and what to observe are driven by themes, categories and theories emerging from the data already collected. This process of theory-driven sampling and data collection continues throughout the field work phase of the study. On any one day during field work a researcher may, for example, review data obtained in the previous day’s interview, compare them with data gained in earlier interviews, engage in participant observation and decide who to interview next and what to ask them. The aim of this process is to generate theories that are firmly grounded in the data collected. Even when a grounded theory approach is taken, data analysis is likely to extend beyond the data collection period.

Many qualitative researchers believe that the final stage of data analysis is the thinking involved in writing the data presentation and discussion sections of the research document.

Research documents based on qualitative studies should include an explicit description of the techniques used in data organization and analysis. If a grounded theory approach has been taken, have the researchers described and justified the decisions they made with regard to sampling/data collection/analysis? Can you see the ways in which these processes were linked?

Exercise 17

How did King & Jensen organize their data? Are the links between sampling, data collection and analysis clear? What major theoretical categories did the researchers develop?

Write some notes and then turn to Section 6.

3.12 Presentation of data

In documents on quantitative research, the results should be presented in a section entitled “Results” and then discussed in a section entitled “Discussion”. In qualitative research, the data are usually discussed as they are presented in a section or sections, which may have a variety of titles. Some qualitative researchers (especially those describing themselves as interpretivists) would argue that researchers construct qualitative data rather than collect them: in this scheme of things the discussion forms part of the data. For this reason, the presentation and discussion of qualitative data will be dealt with under “Discussion of data” (see page 37).

3.13 Presentation of quantitative data

It is generally deemed good practice to present all the results generated by a quantitative study. This prevents the researchers from hiding results that might run counter to the research hypothesis and/or the arguments advanced in the discussion. You may be able to determine whether all the data have been presented if a copy of the data collection instrument is included in the research document.

Example and Exercise 18

Note that in the article by McBride numerical data are presented and discussed concurrently. This is not generally considered good practice when reporting quantitative studies.

Do Allen et al. and McBride present all the data that were collected in their studies?

Turn to Section 6 for discussion.

Response rates to surveys should also be presented. A low response rate (lower than 50% for example) may threaten the **generalizability** of the findings, even if the sample selected was representative and of an adequate size. The characteristics of the population, the sample and the actual respondents in terms of (for example) age, sex and socioeconomic status may also be presented at this point.

Example

The response rates to the survey described by McBride varied between 70% and 88% depending on occupational group (see Table 2 in the article). These were all adequate (88% is very good).

The majority of data may be presented in the text of the document, thus:

- more respondents from Area 1 (264, 68%) than from Area 2 (216, 60%) had consulted a doctor about their bowel problems; *or*
- a significantly greater proportion of the interviewees from Clinic A (105, 88%) than from Clinic B (78, 77%) were satisfied with their consultations with doctors ($\chi^2 = 4.5$, $df = 1$, $P = 0.03$).

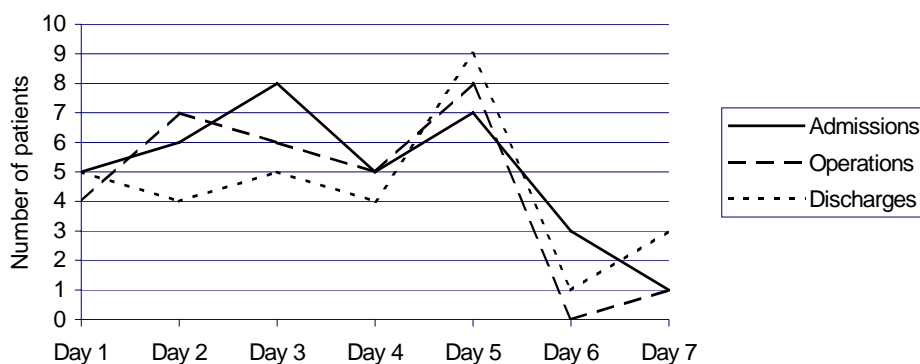
In the latter example χ^2 is the *test statistic* for the chi-squared test and df stands for degrees of freedom (a concept too complex for discussion here). Actual numbers should always be presented, along with percentages. The presentation of percentages alone can be misleading as it may obscure the fact that data relate to a very small and nonrepresentative subgroup of respondents/subjects.

Example

In the article by McBride, percentages only are presented. This masks the fact that the numbers of occupational therapists, physiotherapists and dietitians surveyed were very small. A response given by 72.9% of the occupational therapists (Table 3) sounds far more impressive than a response given by 10 occupational therapists. This is misleading. In both text and tables, actual figures should be given along with percentages, thus: 10 (72.9%).

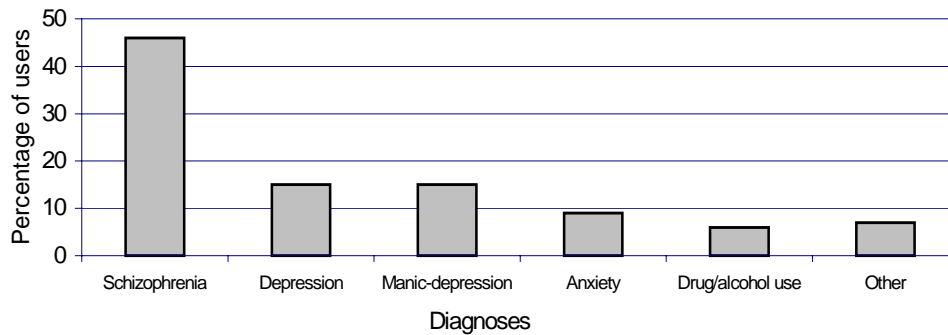
Data are sometimes presented in other formats of which graphs, bar charts, histograms and pie charts, as well as frequency tables and scatter diagrams (illustrated above), are the most commonly used. Examples of these, based on fictitious data, are given in Fig. 3–6.

Fig. 3. Line graph showing admissions, operations and discharges on a surgical ward over a 7-day period



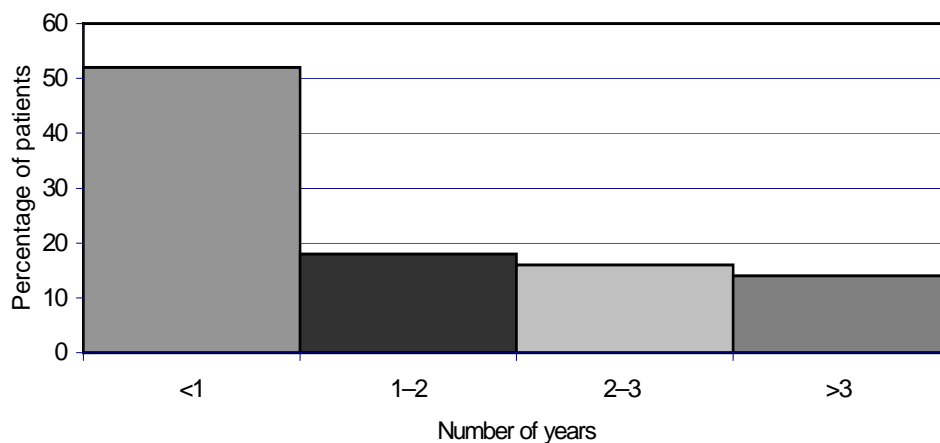
Line graphs are most often used to display changes over time, as in Fig. 3.

Fig. 4. Bar chart showing diagnoses of long-term users of a community mental health service



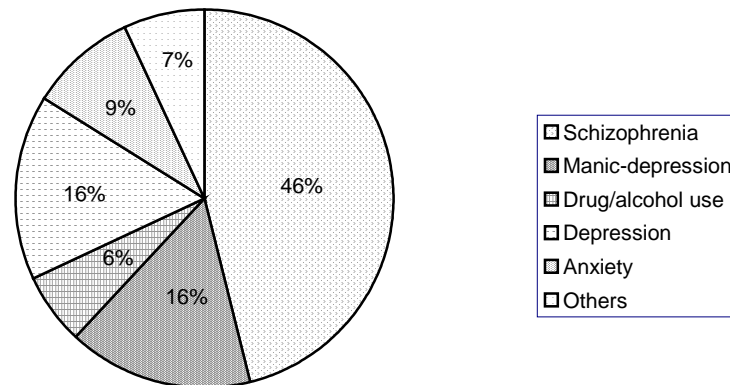
A bar chart (Fig. 4) and a histogram (Fig. 5) look quite similar, except that on a bar chart the bars are separate from each other whereas on a histogram they are touching. This reflects the fact that bar charts are used for discrete variables such as town of birth or grade of staff, whereas histograms are used for continuous variables, such as test scores or age.

Fig. 5. Histogram showing the number of years taken by patients to consult a health professional about urinary incontinence



It is often possible to present data in more than one format. For instance, the data in Fig. 4, a bar chart, could be presented in a pie chart (see Fig. 6).

Fig. 6. Pie chart showing diagnoses of long-term users of a community mental health service



Each table or figure should be numbered and titled clearly, and reference made to it in the text of the document. Well presented tables and figures can aid understanding of data and make research documents more interesting.

Exercise 19

What tables and figures are used in the articles by Allen et al. and McBride? Are these appropriate to the data being presented? Do they help to clarify the data?

Write a few notes and then turn to the discussion in Section 6.

3.14 Discussion of data

Discussion of quantitative data

The Discussion section of a quantitative research document is the part where the interpretative skills of the researchers become important. It is where the researchers attempt to make sense of the data that have been collected and analysed. Researchers must decide what correlations between variables mean and whether hypotheses should be accepted or rejected.

Readers of research documents should decide whether they agree with the researchers' interpretations. For example, could two variables be associated for a different reason than that suggested by the researchers? If you have already identified the presence of uncontrolled confounding variables in an experimental study, this could easily be the case. Are the researchers offering a simple explanation when a more complex one is needed? Do you agree with their decision to reject (or accept) the hypothesis when some results support it and some do not? Are researchers claiming that a relationship is causal when they should only claim that it is correlational?

In addition, the whole of the research document, and the Discussion in particular, should be written in a tentative style. It should be remembered that statistical significance indicates only that a high probability exists that relationships between variables did not occur by chance.

Expressions such as “these results suggest ...” are, therefore, more appropriate than “these results prove ...”. The document should also be written in a clear style, without ambiguity and with a minimum of jargon. Lack of clarity may suggest that the authors are trying to hide something or that there is little of interest at the core of their argument.

Lastly, in the Discussion the researchers should acknowledge the limitations and flaws of their work, such as limited generalizability to other settings and contexts or inadequate control of confounding variables.

Exercise 20

Read the Discussion section of the article by Allen et al. Do the researchers acknowledge the limitations of their study? Do you agree with the main points of their discussion? Do they write in a tentative style?

In the article by McBride discussion is presented concurrently with the data. Does the researcher recognize the limitations of her study? Is the discussion interesting and credible? Is the style of writing tentative?

Write some notes and then turn to Section 6.

Presentation and discussion of qualitative data

As noted above, qualitative data are frequently presented and discussed in the same section of the research document. It is not usually possible to present and discuss all the data collected, as they may be in the form of 10-, 20- or even 50-hour-long interviews plus copious field notes written over several weeks of participant observation. By the end of data collection some of the data collected early in the process may not be relevant to the theories that have emerged, particularly if a grounded theory approach has been taken. Data are therefore selected for presentation and discussion; this should be acknowledged by the researchers.

This issue of data selection can, however, lead to charges of subjective bias, particularly from researchers working within a quantitative framework (as noted above, quantitative researchers are expected to present all their data). However, qualitative researchers would argue that turning statistical data into meaningful ideas is a process based on subjectivity and interpretation, and that it too is bound to be based to some extent on the bias of the researchers.

Several techniques have been developed to either minimize the impact of or explore the issue of bias in qualitative research. Some of these (respondent validation, reflexivity and *inter-rater reliability*) are covered in the discussion on rigour (see page 42). Two more techniques are discussed here: the presentation of substantial extracts from interviews/field notes and the presentation and discussion of data from deviant cases.

Extracts from interview transcripts/field notes should be substantial in terms both of number and of length. These serve to illustrate the discussion and make it more interesting, but also provide evidence that the theories advanced by researchers are grounded in the data collected rather than purely in their own biases and beliefs. The presentation and discussion of data from deviant cases (participants who say or do different things from most of the other people interviewed or observed) may be particularly interesting, in that they can show that researchers are willing to examine and challenge their own theories. This allows researchers to acknowledge that issues are often complex and that human understanding of them will vary.

Data in qualitative research documents are usually presented and discussed theme by theme, and are supported and challenged by introducing references to publications cited in the literature review. Themes will probably reflect (to a greater or lesser extent) the categories formulated during data organization and analysis. Sometimes researchers provide a diagram to show how various theories or parts of a theory relate to each other.

When critiquing qualitative research documents, you need to decide whether the researchers have presented enough evidence to convince you of their arguments. Are there competing theories that the researchers have failed to recognize and discuss? If these problems do exist, might they be related to study design, sampling procedures or other methodological issues?

Exercise 21

Is the theory developed by King & Jensen credible? Can you think of any competing theories? Does the diagram help you to understand how parts of the theory relate to each other?

Make some notes and then turn to the discussion in Section 6.

3.15 Conclusions and recommendations

Not all research documents have a specific section for conclusions and recommendations; sometimes this type of material is included in the Discussion or in a section entitled Implications or similar. If a section entitled Conclusions and recommendations does exist, it is likely to include a summing up of the arguments presented in the Discussion, along with a discussion of the relevance of the research findings to clinical practice or nurse education, for example. The researchers should attempt to define the audience for the document. For quantitative research they should restate the population to which the findings are generalizable, and for qualitative research they should discuss the relevance to particular occupational groups and settings.

Recommendations are likely to be of two main types: for practice and for further research. Within one document, specific recommendations for practice may be directed at particular bodies or groups of people: the government, nurse educators, nurse managers/employers and clinical practitioners, for example. Recommendations for further research may include the performance of a study that addresses the limitations of the research, the qualitative exploration of issues identified in a quantitative study, the testing of hypotheses generated through qualitative research and the *replication* of a quantitative study in order to verify its validity and generalizability.

When critiquing a research document, you should satisfy yourself that the conclusions and recommendations arise out of the findings and discussion. Do you think the recommendations for practice are realistic? Can you think of any other recommendations that the authors could have made? Can you think of any pieces of further research that they have not suggested?

Exercise 22

Read the Conclusion in the articles by Allen et al. and McBride and the Discussion in the article by King & Jensen. What recommendations are made in the three articles? Can you divide them into recommendations for practice and for research? Do you think the recommendations are reasonable, given the findings and any flaws in the study?

Write some notes and then turn to Section 6 for discussion.

3.16 References

You have already read about references on pages 5 and 13. A reference list usually follows the conclusions and recommendations at the end of the article.

Research documents that have references to previous research findings, relevant theories and/or methodological precedents are usually deemed to carry more weight than unreferenced documents. Always make a note of whether a document contains references or not. How many are there? How recent are they? Are there any that you might like to follow up? If you do try to follow one up, you may find that not all the details of the reference are correct: this should be noted.

3.17 Ethical considerations

The first code of ethics relating to biomedical research, the Nuremberg Code, was produced over 50 years ago. In 1964 the World Medical Assembly adopted an influential document on the subject of ethics, the Declaration of Helsinki. Many professional associations and health organizations have since produced their own codes based on the Declaration, including the Royal College of Nursing (16) and the International Confederation of Midwives (17). In addition, in many countries there are local research ethics committees that scrutinize research proposals put forward by nurses, doctors and other health professionals. These committees try to ensure that proposed projects take sufficient account of ethical issues. There are numerous ethical issues that can arise during the process of conducting health research, three of the most important being *informed consent*, *anonymity* and *confidentiality*.

The issue of consent, and the related issue of physical harm, are obvious in research involving potentially dangerous tests or drugs. The dangers of causing emotional harm to interviewees or even to those responding to a questionnaire, however, are perhaps less obvious. The possibility of this occurring may be recognized where the research topic is obviously emotive, as in the case of bereavement or sexual assault. Problems can, however, arise with topics that researchers have expected to be neutral in emotional terms.

Consent must be freely given: researchers must not put any undue pressure on people to participate. In the case of minors or people with impaired mental function, it may be necessary for the parent or guardian to give consent. In experimental and quasi-experimental research the informed consent of all participants should be sought, regardless of whether they are assigned to the experimental or control/comparison group. Consent can be written (participants sign a form) or verbal (participants say they are willing). Written consent should always be obtained where research involves any physical intervention or tests. It may also be a good idea to obtain written

consent before carrying out interviews, particularly with patients. Postal surveys are a different matter, in that written consent may compromise the anonymity that may otherwise be afforded to respondents (see below). In any case, consent is implicit in a respondent's decision to complete and return a survey questionnaire.

With any form of research a detailed, written explanation should be given prior to consent being sought. This should include the following:

- the aims of the research;
- the research methods to be employed;
- assurances of confidentiality;
- details of the funding body (if any); and
- a statement that the participant can withdraw from the study at any time with no detriment to future care/treatment.

In the case of questionnaire surveys, an explanatory letter is sent or handed out with the questionnaire. The written explanation is usually scrutinized by the local ethics committee before the research is started.

When interviewing people about potentially emotive topics, it is good practice to allow some time for debriefing at the end; people may need to discuss personal issues that the research has raised. It is also wise to be aware that people may react in this way to almost any interview topic.

Confidentiality is an important issue because individual research participants may not wish details of their health status or their sexual habits, for example, to be divulged to all and sundry. In addition, they may fear that some of the information given could be used against them: if, for example, they have been critical of their employer or their family doctor. Anonymity is not usually achievable except in the case of postal surveys, where completed questionnaires are identified by a code number alone. Having given assurances of confidentiality (and, perhaps, anonymity) when seeking consent, researchers must ensure that they actually provide this.

Researchers also have a moral obligation to refrain from falsifying or intentionally distorting research results. In addition, they should ensure that the findings of the research are disseminated appropriately: to the funding agency, to relevant policy-makers and practitioners, and to research participants who request feedback.

Some research documents have specific sections relating to ethical issues, but many deal with such issues as they arise at various points throughout the document.

Exercise 23

In the three articles, do the researchers discuss informed consent, anonymity and confidentiality? Do they identify any other ethical issues relevant to the research? Are there other ethical issues that the researchers failed to identify?

Write some notes and then turn to Section 6 for discussion.

3.18 Rigour

Quantitative research

Both quantitative and qualitative research must be conducted rigorously if it is to be considered of worth. The worth of quantitative research is judged on three main criteria: **internal validity**, **external validity** and **reliability**.

Internal validity is concerned with the “truth” of a study and is largely dependent on the soundness of the research design. A study can be said to have internal validity when there is little doubt that significant associations between variables are real, as opposed to artefacts produced by uncontrolled extraneous variables. Experimental studies have a high potential for internal validity because they involve controlled situations. Although quasi-experiments and surveys lack this high level of control and internal validity, researchers can demonstrate rigour by acknowledging and discussing competing explanations for statistically significant results.

There are six types of extraneous variable that pose threats to internal validity.

- **Maturation** relates to the dependent variable being affected by the passing of time, as in wound healing or adaptation to living with a chronic illness.
- **History** relates to events that take place at about the same time as the manipulation of the independent variable. For example, if anxiety rates are being compared in two different towns, a bomb explosion in one town might affect results.
- **Reactivity** relates to awareness of being the subject of research, leading to changes in the types of behaviour being studied.
- **Selection bias** relates to pre-existing differences between groups that may confound the effects of the independent variable.
- **Mortality** relates to subjects dropping out of experimental and control/comparison groups at different rates. When testing a treatment for depression, for example, severely depressed members of a control group may drop out because of apathy or the onset of a crisis. If only mildly depressed patients remain in the control group at the end of the study it may seem as if the control group has improved significantly.
- Lack of reliability (see below).

The validity of research instruments is also an important issue. A valid instrument measures what it is supposed to be measuring. For example, a questionnaire that purports to measure self-esteem should measure all aspects of the subject. It should not, on the other hand, include questions that might elicit replies reflecting people’s religious beliefs or short-term psychological states. Validity can be assessed in three ways.

- The most common is an assessment of **content validity**, in which experts comment on the appropriateness and inclusiveness of the questions in the instrument.
- In an assessment of **criterion-related validity**, scores on the instrument are correlated with those on a valid and reliable measure of the same attribute (the criterion measure) and **validity coefficients** are calculated.

- **Construct validity** can be assessed in two main ways. In one, the instrument is tested on groups whose scores are likely to differ. For example, a scale to measure depression could be tested on patients with a diagnosis of depression and on people with no mental health problems; difference in scores suggests that the instrument is valid. The second way is to correlate scores on the depression scale with scores on an instrument that measures a related construct, such as grief or self-esteem. A positive correlation suggests that the depression scale is high in validity.

External validity refers to the generalizability and applicability of findings and is concerned mainly with issues relating to sampling and research settings. External validity is enhanced if the study sample is representative of the population from which it is drawn. Findings can then be generalized from the sample to the entire population. Random sampling and/or the presentation of findings relating to sample and population characteristics can help to establish external validity. It is interesting to note that research, such as experiments, conducted in highly controlled situations may be low in external validity, as it may be difficult to apply findings to everyday settings. Research such as quasi-experiments and surveys, which is conducted in a natural setting and subject to fewer controls, may by contrast be high in external validity. Internal and external validity are thus to some extent in conflict with each other.

Reliability is concerned with stability of measurement and the repeatability of testing procedures. It centres on various types of data collection instrument: technical devices such as weighing machines and thermometers used to measure physical variables; and structured observation schedules, interview schedules and questionnaires used to measure behavioural, attitudinal and psychological variables. Instruments should be designed so that they are capable of providing the same reading on or answer from the same participant in repeated tests by the same or different researchers. A **reliability coefficient** can be computed for the instrument, based on test–retest, **internal consistency** or inter-rater scores.

An instrument that is not reliable cannot be valid, as it will provide inconsistent and inaccurate results. Nevertheless, an instrument that is reliable is not necessarily valid, as it may be used to measure variables for which it is not designed. (A thermometer, for example, may be reliable but it is not a valid instrument for measuring blood pressure or self-esteem.) Instruments must therefore be appropriately designed to ensure validity and rigorously tested to ensure reliability.

Exercise 24

What problems relating to internal validity, external validity and/or reliability can you identify in the articles by Allen et al. and McBride?

When attempting to answer this question, you may wish to reread or refer to the discussion of these two articles in Section 6. The discussion on the choice of research methods, pilot work, sampling procedures, data collection methods, data analysis and discussion of data may be particularly helpful.

Make some notes on your answers and then turn to the relevant discussion on rigour in Section 6.

Qualitative research

Some qualitative researchers do not use the terms validity and reliability as they believe them to be imbued with inappropriate concepts such as control, standardization and objectivity. While these researchers propose a concept of rigour appropriate to qualitative research, along with the use of alternative terms, other qualitative researchers continue to use the original quantitative terms. In addition, whereas the majority of qualitative researchers accept that internal and external validity (or their qualitative equivalents) are concepts appropriate to qualitative research, there is less agreement about reliability.

Internal validity is renamed “credibility” by Sandelowski (18) and “truth value” by Guba & Lincoln (19). Some qualitative researchers have sought to increase the credibility of their work by employing techniques such as respondent validation. This means that they take their findings back to the research participants for confirmatory and/or critical comment, a process that may lead to modification of the findings. This technique has been found to have several drawbacks. For example, participants may not understand or be interested in the findings, or they may want researchers to change or omit reports of actions or words that, they feel, place them in a bad light.

Researchers working within an interpretivist framework tend to reject techniques such as respondent validation, recognizing that, ultimately, a research document represents the researchers’ interpretation of the participants’ accounts and actions. Interpretivists often rely on what is known as reflexivity to establish credibility. A reflexive approach necessitates the exploration of a number of factors, including:

- the context of the events and words that form the data;
- the effects of interactions between researchers and participants on the data; and
- the effects of researchers’ personal experiences and beliefs on the interpretation of the data.

Details of such matters are kept in reflexive notes (or a reflexive journal), which form part of researchers’ field notes. By adopting a reflexive approach, researchers aim to give a rigorous justification of their interpretations as credible, rather than to support claims that their accounts are true in absolute terms.

The presentation of substantial extracts from interview transcripts and field notes and the presentation and discussion of data from deviant cases (which you read about above) can also serve to increase the credibility of a qualitative research document.

External validity is renamed “fittingness” by Sandelowski (18) and “applicability” by Guba & Lincoln (19). Sandelowski suggests that, in qualitative research, fittingness has to be determined by a document’s audience. In other words, readers have to decide whether the ideas in a document can be usefully applied in practice. Sandelowski also states that fittingness is demonstrated when “findings can ‘fit’ into contexts outside the study situation”. Fittingness can also be demonstrated by “backfitting” the data to existing literature on the same subject, in order to show how they challenge and/or add to existing knowledge.

Reliability, renamed “auditability” by Sandelowski (18), is a rather more contentious issue than validity in qualitative research. Some researchers advocate the coding of qualitative data by two researchers in the hope that they will make similar coding decisions (inter-rater reliability). Some also claim other researchers should be able to repeat the research or scrutinize all the data and come to the same or similar conclusions. Interpretivists might argue, however, that this sort

of agreement and/or repeatability is neither possible nor relevant, since every situation, time, person and group of people studied during field work is unique and every researcher is bound to interpret data differently. They would argue that researchers can demonstrate rigour by providing a detailed, logical account of their decision-making processes during sampling, field work and data analysis.

Exercise 25

What attempts do King & Jensen make to establish the rigour of their work?

Make some notes and then turn to the discussion in Section 6.

Both quantitative and qualitative research can only be as rigorous as time and financial resources allow. These limits to rigour should be acknowledged and discussed by researchers.

As rigour is a very complex subject, you may find that you need to read a lot of research documents and/or undertake further learning about research methods before you feel confident in making judgements about what is valid, reliable, credible, applicable and so on. When evaluating research documents, however, you might like to make some preliminary notes on these matters and return to them later when you have gained more knowledge.

3.19 Relevance to practice

The relevance of research is closely related to its external validity or applicability. When evaluating quantitative work you should remember that just because a result is statistically significant it does not mean that it has any clinical or practical significance. A 5% reduction in patients' pain scores in the experimental group but not in the control group may be statistically significant, but the patients themselves may feel no improvement. In addition, the intervention may have side effects, be prohibitively expensive or have various other drawbacks that may render it unusable. These issues may or may not be discussed by the researchers. Some qualitative research studies may be carried out in settings that are so unusual that it is difficult to apply their findings widely.

It is not easy to assess the relevance of research documents that relate to disciplines or nursing specialties outside your area of experience or expertise. You will find this easier with documents dealing with your field of practice.

* * * * *

You have now read about the main factors that require evaluation when reading a research document. You have also made some preliminary attempts at evaluation by doing the exercises based on the articles provided. Section 4 tells you how you can build on your new knowledge and skills.

4. Where to go from here

- Obtain research documents relevant to your clinical practice/work situation (preferably via the routes suggested in Section 2).
- Make out a record card/sheet for each document as described in Section 2.
- Evaluate the documents using the framework in Section 3, and write the details of your critique on the reverse of your card/sheet.
- Form a journal club with other nurses who have read this document or who have some knowledge of research issues (see under Learning objectives, page 1).
- Some of the documents you obtain and read may have implications for your everyday practice. Do not make radical changes to your practice on the basis of one research document.
- Try to find other documents that support (or refute) research findings. Systematic reviews of the literature may be particularly useful.
- If you decide, on the basis of ample evidence, that you want to make changes, discuss them with other nurses in your field.
- Make changes for a trial period and then evaluate them (you may be able to use pre-test/post-test techniques to do this).
- Remember that introducing change may bring you into conflict with nursing colleagues, medical staff and others who may be affected by them. Always discuss plans for change with these people.
- Remember that the ultimate result of any change should be improved patient care.

Try to obtain copies of the following, which are recommended reading for those of you who wish to learn more about research methods. Together, they give a good overview of quantitative and qualitative research methods.

POLIT, D.T. & HUNGLER, B.P. *Essentials of nursing research: methods, appraisal, and utilization*. Philadelphia, J.B. Lippincott Co., 1989.

HAMMERSLEY, M. & ATKINSON, P. *Ethnography: principles in practice*, 2nd ed. London, Routledge, 1995.

Research awareness: a programme for nurses, midwives and health visitors. London, Distance Learning Centre, South Bank University (11 modules).

Module 1. *Health care in professional practice*, 1995.

Module 2. *Sources of nursing and midwifery knowledge*, 1995.

Module 3. *What is research?*, 1992.

Module 4. *Searching the literature*, 1989.

- Module 5. *Identifying and defining questions for research* (no date given).
- Module 6. *Ethics in nursing and midwifery research*, 1994.
- Module 7. *The ethnographic perspective*, 1987.
- Module 8. *The survey perspective*, 1988.
- Module 9. *The experimental perspective*, 1988.
- Module 10. *Evaluating research*, 1994.
- Module 11. *Using research findings*, 1994.

5. Glossary of research terms²

Abstract A brief description of a completed or proposed research investigation; in research journals, usually located at the beginning of an article.

Anonymity Protection of the participant in a study such that even the researcher cannot link him or her to the information provided.

Central tendency A statistical index of the “typicalness” of a set of scores that comes from the centre of the distribution of scores. The three most common indices of central tendency are the mode, the median and the mean.

Cluster sampling A form of multistage sampling in which large groupings (“clusters”) are selected first (e.g. nursing schools), with successive subsampling of smaller units (e.g. nursing students).

Comparison group A group of subjects whose scores on a dependent variable are used as a basis for evaluating the scores of the experimental group or group of primary interest. The term comparison group is generally used instead of control group when the investigation does not use a true experimental design.

Confidentiality Protection of participants in a study such that their individual identities will not be linked to the information they provided and will never be publicly divulged.

Construct validity The degree to which an instrument measures the construct under investigation.

Content validity The degree to which the items in an instrument adequately represent the universe of content for the concept being measured.

Control group Subjects in an experiment who do not receive the experimental treatment and whose performance provides a baseline against which the effects of the treatment can be measured (see also *comparison group*).

Convenience sampling Selection of the most readily available persons (or units) as subjects in a study; also known as accidental sampling.

Correlation coefficient An index that summarizes the degree of relationship between two variables. Correlation coefficients typically range from +1.00 (for a perfect positive relationship), through 0.0 (for no relationship), to –1.00 (for a perfect negative relationship).

Criterion-related validity The degree to which scores on an instrument are correlated with some external criterion.

² Reproduced by permission from Polit & Hungler (11).

Dependent variable The outcome variable of interest; the variable that is hypothesized to depend on or be caused by another variable (the independent variable); sometimes referred to as the criterion variable.

Descriptive statistics Statistics used to describe and summarize the researcher's data set (e.g. mean, standard deviation).

Double-blind experiment An experiment in which neither the subjects nor those who administer the treatment know who is in the experimental or control group.

Experiment A research study in which the investigator controls (manipulates) the independent variable and randomly assigns subjects to different conditions.

Experimental group The subjects in an experimental or quasi-experimental study who receive the experimental treatment or intervention.

External validity The degree to which the results of a study can be generalized to settings or samples other than the ones studied.

Extraneous variables Variables that confound the relationship between the independent and dependent variables and that need to be controlled either in the research design or through statistical procedures.

Field notes The notes taken by researchers regarding the unstructured observations they have made in the field, and their interpretation of those observations.

Focused interview A loosely structured interview in which the interviewer guides the respondent through a set of questions using a topic guide.

Frequency distribution A systematic array of numerical values from the lowest to the highest, together with a count of the number of times each value was obtained.

Generalizability The degree to which the research procedures justify the inference that the findings represent something beyond the specific observations on which they are based; in particular, the inference that the findings can be generalized from the sample to the entire target population.

Grounded theory An approach to collecting and analysing qualitative data with the aim of developing theories and theoretical propositions "grounded" in real-world observations.

Hawthorne effect The effect on the dependent variable caused by subjects' awareness that they are "special" participants under study.

History A threat to the internal validity of a study; refers to the occurrence of events external to the treatment but concurrent with it that can affect the dependent variable of interest.

Hypothesis A statement of a predicted relationship between the variables under investigation.

Independent variable The variable that is believed to cause or influence the dependent variable; in experimental research, the independent variable is the variable that is manipulated.

Inferential statistics Statistics that permit us to infer whether relationships observed in a sample are likely to occur in a larger population of concern.

Informed consent An ethical principle that requires researchers to obtain the voluntary participation of subjects, after informing them of possible risks and benefits.

Instrument The device or technique that a researcher uses to collect data (e.g. questionnaires, tests, observation schedules).

Internal consistency A form of reliability, referring to the degree to which the subparts of an instrument are all measuring the same attribute or dimension.

Internal validity The degree to which it can be inferred that the experimental treatment or independent variable, rather than the uncontrolled extraneous factors, is responsible for observed effects on the dependent variable.

Inter-rater reliability The degree to which two raters, operating independently, assign the same ratings for an attribute being measured; such ratings normally occur in the context of observational research or in coding qualitative materials.

Interval measure A level of measurement in which an attribute of a variable is rank ordered on a scale that has equal distances between points on that scale (e.g. Fahrenheit degrees).

Intervention In experimental or quasi-experimental research, the experimental treatment or manipulation.

Judgemental sampling A type of non-probability sampling method in which the researcher selects subjects for the study on the basis of personal judgement about which ones will be most representative or productive; also referred to as purposive sampling.

Maturation A threat to the internal validity of a study that results when the outcome measure (dependent variable) changes as a result of time passing.

Mean A descriptive statistic that is a measure of central tendency, computed by summing all scores and dividing by the number of subjects.

Measurement The assignment of numbers to objects according to specified rules to characterize quantities of some attribute.

Median A descriptive statistic that is a measure of central tendency, representing the exact middle score or value in a distribution of scores; the median is the value above and below which 50 per cent of the scores lie.

Mode A descriptive statistic that is a measure of central tendency; the score or value that occurs most frequently in a distribution of scores.

Mortality A threat to the internal validity of a study, referring to the differential loss of subjects (attrition) from different groups.

Nominal measure The lowest level of measurement that involves the assignment of characteristics into categories (e.g. males, category 1; females, category 2).

Nonparametric statistics A general class of inferential statistics that does not involve rigorous assumptions about the distribution of the critical variables; most often used to test hypotheses about variables measured on the nominal or ordinal scales.

Non-probability sampling The selection of subjects or sampling units from a population using non-random procedures; examples include accidental, judgemental and quota sampling.

Normal distribution A theoretical distribution that is bell-shaped and symmetrical; also called a normal curve.

Null hypothesis The hypothesis that states there is no relationship between the variables under study; used primarily in connection with tests of statistical significance as the hypothesis to be rejected.

Observational research Studies in which the data are collected by means of observing and recording behaviours or activities of interest.

Ordinal measure A level of measurement that yields rank orders of a variable along some dimension.

Parametric statistics A class of inferential statistics that involves (a) assumptions about the distribution of variables; (b) the estimation of a parameter; and (c) the use of interval or ratio measures for the dependent variable.

Participant observation A method of collecting data through the observation of a group or organization in which the researcher participates as a member.

Pilot study A small-scale version, or trial run, done in preparation for a major study.

Population The entire set of individuals (or objects) having some common characteristic(s).

Post-test The collection of data after the introduction of an experimental intervention.

Pre-test The collection of data prior to an experimental intervention; sometimes referred to as baseline data.

Probability sampling The selection of subjects or sampling units from a population using random procedures; examples include simple random sampling, cluster sampling and systematic sampling.

Purposive sampling A type of non-probability sampling method in which the researcher selects subjects for the study on the basis of personal judgement about which ones will be most representative or productive; also referred to as judgemental sampling.

Quasi-experiment A study in which subjects cannot be randomly assigned to treatment conditions, although the researcher does manipulate the independent variable and exercises certain controls to enhance the internal validity of the results.

Quota sampling The non-random selection of subjects in which the researcher pre-specifies characteristics of the sample to increase its representativeness.

Random number table A table of digits from 0 to 9 set up in such a way that each number is equally likely to follow any other; used in randomization or random sampling.

Randomization The assignment of subjects to treatment conditions in a random manner (i.e. in a manner determined by chance alone); also known as random assignment.

Range A measure of variability, consisting of the difference between the highest and the lowest values in a distribution of scores.

Ratio measure A level of measurement in which there are equal distances between score units, and which has a true meaningful zero point; the highest level of measurement.

Reactivity A measurement distortion arising from the subject's awareness of being observed or, more generally, from the effect of the measurement procedure itself.

Reliability The degree of consistency or accuracy with which an instrument measures the attribute it is designed to measure.

Reliability coefficient A quantitative index, usually ranging in value from 0.00 to 1.00, that provides an estimate of how reliable an instrument is.

Replication The duplication of research procedures in a second investigation for the purpose of determining if earlier results can be repeated.

Respondent In a self-report study, the research subject.

Response rate The rate of participation in a survey, calculated by dividing the number of persons responding by the number of persons invited to respond.

Sampling The process of selecting a portion of the population to represent the entire population.

Sampling frame A list of all elements in the population from which the sample is drawn.

Selection bias A threat to the internal validity of the study that results from differences between groups being compared that can account for differences in the groups' performance on the dependent variable.

Significance level The probability that an observed relationship could be caused by chance (i.e. because of sampling error); significance at the 0.05 level indicates the probability that a relationship of the observed magnitude would be found by chance only 5 times out of 100.

Simple random sampling The most basic type of probability sampling, wherein a sampling frame is created by enumerating all members of a population of interest, and then selecting a sample from the sampling frame through completely random procedures.

Standard deviation The most frequently used statistic for measuring the degree of variability in a set of scores.

Statistical significance A term indicating that the results obtained in an analysis of sample data are unlikely to have been caused by chance, at some specified level of probability.

Strata Subdivisions of the population according to some characteristic (e.g. males and females); singular is stratum.

Stratified random sampling The random selection of subjects from two or more strata of the population independently.

Test statistic A statistic used to test for the statistical significance of relationships between variables; the sampling distributions of test statistics are known for circumstances in which the null hypothesis is true; examples include chi-squared, F -ratio, t , and Pearson's r .

Unstructured interview An oral self-report in which the researcher asks respondents questions without preconceived views regarding the specific content or flow of information to be gathered.

Validity The degree to which an instrument measures what it is intended to measure.

Validity coefficient A quantitative index, usually ranging in value from 0.00 to 1.00, that provides an estimate of how valid an instrument is; usually computed in conjunction with the criterion-related approach to validating an instrument.

Variable A characteristic or attribute of a person or object that varies (i.e. takes on different values) within the population under study (e.g. body temperature, age, heart rate).

6. Discussion of exercises

This discussion is designed to provide more knowledge than you have already gained from reading the main text. This is because issues can be discussed that may seem abstract and difficult to understand without reference to specific examples. Do not, therefore, be discouraged if an answer you have written down is incorrect or seems inadequate. Go back to the relevant article and try to apply the knowledge you have gained through reading this discussion.

6.1 Exercise 3. The aims of the study

Allen et al.

The aims of the study are neither easy to find nor clearly stated, although relevant information is given in the final paragraph of the Introduction. The authors could have included a statement such as “The aim of the study was to evaluate the effectiveness of a home-care preoperative teaching programme for older people as measured by the degree of self-care attained”. In actual fact, the researchers formulated two research hypotheses but these are not stated until the Results section of the article.

McBride

The aims of the study are easy to find (under the subheading Aims!). The aims are stated fairly clearly and in some detail. At the end, two of the issues under study are expressed in a format suggestive of research questions.

King & Jensen

The aims can be found fairly easily at the end of the introductory section (which has no title). The purpose of the study is clearly stated and the research question “What is it like for women to have cardiac surgery?” is posed.

6.2 Exercise 4. The literature review

Allen et al.

The literature review is easy to find as it is entitled Review of the literature. Nevertheless, many of the statements in the Introduction (the opening sentence, for instance) should have been backed up with references and might have been included in the review.

The review focuses specifically on nursing literature on preoperative patient teaching. Gaps in the literature are identified, as in the statement “There is no research that addresses the effect of preoperative teaching ...” (page 304). This particular statement relates to the aims of the study.

The review makes no attempt to evaluate the quality of documents. All the documents discussed could have serious methodological flaws weakening and/or invalidating their findings.

The ideas in documents are compared and contrasted, as in the discussion on procedural, behavioural and sensory information.

Only 2 out of 20 references in the review are to documents published in the previous 5 years. This raises concerns about its comprehensiveness.

The article uses the Harvard system of referencing.

McBride

The literature review is fairly easy to identify as it takes up the whole of the section entitled Background. Some parts of this section, including the final paragraph, however, contain no references.

Relevant literature on nursing and health promotion in the United Kingdom is reviewed but the review is quite narrow. For example, it does not discuss literature on health promotion as “victim-blaming”, a concept introduced in the Analysis (page 94).

Gaps in the literature are implied rather than stated: the suggestion that health promotion in hospitals is an under-researched subject is somewhat tangled up with suggestions that little health promotion is occurring in hospitals.

There is no attempt to evaluate the quality of documents.

Attempts to compare and contrast documents are minimal.

There are only nine references in the review (raising further queries about comprehensiveness), although the majority are recent.

The article uses the Harvard system of referencing.

King & Jensen

The literature review is fairly easy to find as it takes up the whole of the (untitled) introductory section, apart from the last sentence.

Nursing and medical literature relevant to the aims of the study is discussed. A gap in the literature is identified regarding “what women experience when having cardiac surgery”. The research aims to fill this gap.

Individual documents are not critically evaluated in terms of methodology, but the author makes a blanket criticism that most research has been done by men on men and the findings then applied to women.

Documents are compared and contrasted, in particular those based on “the prevailing male model” and those that examine feminist or gender issues.

There are only 10 references in the review but they are mostly recent.

The article uses the numerical system of referencing.

6.3 Exercise 5. The choice of research methods

Allen et al.

This article describes an experiment. The study meets the three criteria for a true experiment because the researchers manipulated the independent variable by giving “home care preoperative teaching” to some subjects and withholding it from others. In addition, a control group was used and subjects were assigned to control and experimental groups on a random basis. Unfortunately, the researchers do not make most of these matters explicit: the words experiment and random are not used. The researchers are rather vague on the matter of randomization, merely stating that group assignment was performed “according to a computer-generated list”. We are left to assume that assignment was random.

McBride

This describes a survey. Again, the researcher could have been more helpful: the word survey does not appear in the abstract or in the early part of the section entitled The research project.

Note: Turn back to page 15 before reading further.

Allen et al.

The dependent variables were the subjects’ knowledge, skill and anxiety and also the number of postoperative nursing visits and amount of nurses’ time required by subjects.

There is a possibility that subjects (particularly those in the control group) might have sought and gained helpful information from other sources preoperatively and postoperatively. Some of this information might have been similar to the content of the teaching programme (this problem is acknowledged in the Discussion). The authors do not discuss whether the control and experimental groups differed significantly in terms of visual impairment, either preoperatively or postoperatively. This factor could affect people’s knowledge, skills and anxiety.

All the participants should have been fully informed about the experimental/control group format of the study and the nature of the intervention (see Ethical considerations). It should therefore have been obvious to them which group they were in. We are not informed whether the researcher who performed pre-test data collection was “blind to subject group” at that time: if not, he/she could have manipulated pre-test scores. The study was, therefore, not carried out under double-blind conditions. Double-blind conditions are, in fact, impossible to achieve except where experimental and control measures appear to be identical to all researchers and participants (e.g. the active medication and the placebo can be made to look identical).

McBride

The author states, under the subheading Choice of study population, that the population is “hospital-based nurses on acute adult wards”. We are not told whether this refers to all such nurses in the United Kingdom, in a defined geographical area or a particular hospital. Some of the statements used to justify her choice of population are highly debatable. These include ideas that adult patients are “able to make their own lifestyle choices”, are “autonomous”, and are (by implication) “wholly responsible for themselves”. These ideas may be particularly inappropriate to economically and educationally disadvantaged people, to people with disabilities and to those with mental health problems (among others), all of whom may be admitted to acute adult wards for medical and surgical care. In addition, elderly people are frequently nursed on acute adult wards.

The author defines “hospital-based physiotherapists, dietitians, occupational therapists, and surgical and medical consultants who have an active clinical practice” as a second survey population. Again, she fails to give any idea of the geographical extent of the population.

6.4 Exercise 6. The choice of research methods

King & Jensen

The authors do not describe the study as ethnographic, phenomenological or even qualitative. This is fairly unusual but is not necessarily a flaw. The authors do, however, state that they take a grounded theory approach. The study resembles a phenomenological study rather than an ethnographic study, as it focuses on the experiences of individuals rather than social interaction.

6.5 Exercise 7. Pilot work

Allen et al.

It is stated that “a pilot study was conducted” and that changes were made “in the teaching programme and the administration of the study” (page 306), but there are no details of what modifications were made as a result. No details are given about the size or composition of the pilot sample or how researchers gained access to the sample.

McBride

“The questionnaire was piloted in a different hospital ... No changes were made to the format” (page 93–94). It is quite unusual for no modifications to instrument design to be made following a pilot study: this raises possible concerns about the rigour of piloting procedures. It would appear that the questionnaire used in the pilot and main studies was exactly the same as that used in the “previous survey of general practitioners”, but this is not made clear. Again, no details about size, composition or access are given.

King & Jensen

No pilot work was conducted. In a qualitative study, particularly one involving a grounded theory approach, this is not a flaw. In quantitative studies instruments and sampling procedures have to be finalized before any data is collected, but in qualitative research instruments (and sometimes sampling decisions) evolve as the main study progresses.

6.6 Exercise 8. Sampling procedures

Allen et al.

Sample selection is described under the subheading Subjects. The researchers selected all people who met their eligibility criteria and who were due to have cataract surgery within a specific two-week period. Without the random assignment of people to experimental and control groups such a sample would be described as a convenience sample.

The sample was composed of only 36 people. This is probably too small to allow modest differences between the groups to achieve statistical significance. The inadequate size of the sample is acknowledged in the Discussion. In addition, we are not told anywhere in the article

whether there were equal numbers of subjects in the experimental and control groups (we have to assume that this was the case).

There is no mention of the researchers taking statistical advice. A statistician might have conducted a “power analysis” to predict the size of sample appropriate to the study design.

McBride

We have already discussed the problem of defining the study population. To obtain their survey sample, the researchers took the population of eligible nurses from one “large teaching hospital” and added in “a random selection of 50 qualified nurses from another district general hospital in a different district health authority ... to improve the generalizability of the findings” (page 93). We have to assume from this statement that the researchers wished the findings to be generalized to nurses working on acute adult wards throughout a wide geographical area, such as England or the United Kingdom. The population sample plus random sample described above, however, amounts to a convenience sample.

The survey sample of “other health care professionals” comprised the population of all such professionals in one district health authority. If we again assume that the target population was similar health professionals throughout England or the United Kingdom, however, this represents a convenience sample. The researcher may have been able to select representative samples of nurses and other health professionals by using existing databases (such as lists of relevant staff employed by trusts and health authorities) and employing non-probability sampling. We do not know whether this possibility was explored.

The sample comprised 296 nurses. We are given no figures for the study population. If the researchers had looked at official figures for nurses employed in England or the United Kingdom they would have found that the population of nurses working on acute adult wards runs to tens of thousands. The sample would, therefore, appear to be inadequate in size.

We are not told how large the sample (or samples) of other health professionals were (nor are we given estimates of population figures). We have to try to work out sample sizes from Table 1, which gives numbers of respondents and the response rates. We can calculate that the samples of occupational therapists, physiotherapists and dietitians were very small indeed (20, 20 and 24, respectively). It is difficult to argue that these samples are representative.

There is no mention of the researchers taking any statistical advice on sampling.

6.7 Exercise 9. Sampling procedures

King & Jensen

Quite a clear account of theoretical sampling is given in the second paragraph of the Method section. This is the sampling process appropriate to a grounded theory approach.

The researchers claim that “sampling continued until the categories identified ... were saturated”. It is, perhaps, surprising that no new concepts were emerging after only 10 interviews, given the diversity of the interviewees in terms of the demographic data presented in Table 1. It is not possible to state that the small sample size represents an outright flaw, as we have no firm evidence to suggest that saturation did not occur. We might, however, retain some reservations about the comprehensiveness of the data.

Exercise 10. Data collection methods

Allen et al.

The data collection instruments were the State-Trait Anxiety Inventory (STAI) and the Knowledge and Skill Test (KST). These instruments were administered as a questionnaire, except where subjects were unable to read them, in which case they were administered in a structured interview format.

Neither instrument is provided. This means that we have no clear details of their content and cannot assess their quality. Although the STAI is described as “one of the best standardized and most highly used of anxiety measures” (page 305), the reference to support this dates from 1978. A more recent reference might have helped to assure us that the STAI had not been discredited or superseded by a superior measure since 1978. The KST is something of an unknown quantity as it was “developed by the investigators” and had never been used before. Sometimes it is not possible to present lengthy instruments in an article. This does not apply in this particular case, as the STAI consisted of just 20 items and the KST of 10 items plus a skills check list. These issues are considered further in the discussion on rigour (page 42).

Exercise 11. Data collection methods

McBride

The survey questionnaire is not included. We cannot assess its quality and, although we know that it had been used before with general practitioners, we cannot assess its appropriateness to nurses and other health professionals. We are given no idea of its length, a factor that may affect the quality of responses.

We do know that respondents were asked to indicate their level of agreement or disagreement with a number of statements on health promotion (page 94). The examples of statements given on pages 95 and 98 suggest that the questionnaire may be poor in quality. All five statements lead respondents towards giving positive answers about the role of nurses in health promotion. Unless these were balanced with similar statements that led in a negative direction, this may have caused “response bias” in the data (i.e. respondents may have answered in the way they felt was expected of them).

In addition, the third question presented on page 95 is worded so as to be complex or (in this case) “triple-barrelled”: respondents might think it appropriate for nurses to tell people to stop smoking and lose weight but inappropriate to tell them to take more exercise (or vice versa). In other words, they may agree with part of the statement and disagree with part of it. This makes it impossible for them to give a response that reflects their opinions with any accuracy.

It is also interesting to note that the questionnaire is described in the abstract as “semi-structured”. This is an unusual term to apply to a quantitative data collection instrument, particularly a questionnaire. Quantitative instruments are usually designed to be highly structured and standardized, with a high proportion of closed questions. This approach facilitates consistency in the numerical coding of data.

Exercise 12. Data collection methods

Allen et al.

It is unclear whether the subjects’ skills in instilling medications were assessed through observation or self-reporting (questionnaire/interview). The former would provide the clearest evidence (although reactivity might be a problem with nervous patients).

Exercise 13. Data collection methods

King & Jensen

The data collection methods are described towards the end of the Method section. Data collection was based solely on tape-recorded interviews. Participants were interviewed twice (at 4–8 weeks and at 2–9 months) in their own homes. The interviews are described as unstructured and interactive.

We are not told how many interviewers performed the interviews. As 21 interviews were conducted with 10 participants, it would appear that one woman was interviewed three times, but we are given no reason for this. No data collection instrument is presented. We are not told whether the interviewer(s) worked from an interview schedule or guide that was progressively modified as they pursued theories emerging from the data. It would have been helpful to have been given a version of this schedule or guide. It is possible, however, as the interviews are described as unstructured, that there was no such schedule or guide. Some researchers feel that the term “unstructured” is misleading when applied to qualitative interviews: interviewers always approach interviews with some sort of agenda, particularly in the later stages of grounded theory research. In addition, interviewees’ concerns may bring structure to interviews. There is no discussion of ethical issues or problems of managing interviewer–interviewee relationships.

Exercise 14. Data analysis

Allen et al.

The level of data is not stated: this is by no means unusual in research documents, particularly articles. If we read the subsection entitled Data analysis, however, we read that “group” was the independent variable. To put this more fully: receipt of the “home care preoperative teaching programme” (by the experimental group) and its non-receipt (by the control group) was the independent variable. This variable produces data that can only be measured at the nominal level, e.g. 1. experimental group 2. control group (where 1 and 2 are the numerical codes). The absence of a data collection instrument makes it difficult to determine the level of measurement appropriate to the dependent variables (knowledge, skill and anxiety). The STAI may have been based on scales producing ordinal- or interval-level data, and the KST probably involved calculation of a test score that would be subject to measurement at the interval level.

McBride

On page 94 it is stated, “Respondents were asked to agree or disagree with statements relating to health promotion. This was recorded as ‘strong agreement’, ‘slight agreement’, ‘slight disagreement’ or ‘strong disagreement’”. This statement suggests that the ordinal level of measurement was appropriate to the data collected.

Exercise 15. Data analysis

Allen et al.

This study uses both descriptive and inferential statistics.

The descriptive statistics presented are the mean and the standard deviation, used to describe the age of subjects (page 306), KST test scores and levels of anxiety (page 307). We are not told whether the data were normally distributed, which makes it difficult to decide whether the correct measures of central tendency and dispersion were used.

The inferential tests used in the study are analysis of covariance and the *t*-test. The *t*-test is a parametric test that comes in two forms, one for data from independent groups (*t*-test for independent samples) and one for data gained through repeated measures and pre-test/post-test

designs (*t*-test for paired or dependent samples). The *t*-test mentioned on page 306 should have been a *t*-test for independent samples, since it was used to compare post-test data from the two groups. This is not, however, clearly stated. In the Results (page 307) it is stated that a “paired *t*-test” was performed to compare pre-test and post-test scores within the two groups. We do not know whether the data were normally distributed and, thus, whether parametric tests were appropriate.

McBride

This study uses descriptive statistics only.

The only statistics presented are percentages.

Exercise 16. Data analysis

Allen et al.

A null hypothesis was not formulated. The authors do not state their research hypotheses until this point; these should have been stated in the Introduction and/or at the beginning of the section entitled The study.

There appears to be a typographical error on page 307, where it is stated that the “level of significance ... was established at $\alpha = 0.05$ ”. This should read $P = 0.05$.

None of the tests (analysis of covariance and *t*-tests) associated with the hypotheses produced statistically significant results. These tests measured differences between the experimental and the control group in terms of postoperative knowledge, skill and anxiety and the need for nurse visits and time. The results of these tests are not presented. As an afterthought, the researchers performed paired *t*-tests to determine whether there was a difference in pre-test and post-test KST and A-State scores within the two groups.

Some of the results were statistically significant and are presented in Tables 1 and 2. These are:

- a significant increase in post-test knowledge in the experimental group ($P = 0.025$);
- a significant increase in post-test skill in the experimental group ($P = 0.019$) and the control group ($P = 0.000$); and
- a significant decrease in post-test anxiety in the control group ($P = 0.05$).

Statistically significant results are often marked with an asterisk (*), as in Tables 1 and 2.

The practice of trawling through results to find any that are statistically significant and then devising a theory to fit them is not generally condoned in experimental research. This is what the researchers involved in this study may have done.

Exercise 17. Data analysis

King & Jensen

There is no description of how the researchers organized the data. We do not know, for example, whether they assigned data to coded categories and subcategories or whether they used index cards or a computer package.

The links between theoretical sampling and data analysis are made fairly clear in the second paragraph of the Method section. The links between data collection and data analysis are

less clear: no data collection instrument is presented (or even mentioned) and there is no discussion of how interview content evolved as theories emerged through data analysis.

When taking a grounded theory approach to data analysis, researchers often try to identify a core category that is capable of linking the other categories together and explaining variation in the data. The researchers in this study identified “preserving the self” as the core category. The women in the study struggled to preserve the self before, during and after their cardiac surgery. The researchers developed five categories relating to phases in this process: “waiting, getting there, surviving, being there and moving on” (page 100). They also identified four categories that reflected preservation strategies: “relating with others”, “managing the experience”, “normalizing” and “making sense” (pages 100–101).

Exercise 18. Presentation of data

Allen et al.

As stated above, data relating to the hypotheses were not presented. Even though there were no statistically significant differences between the groups, the researchers should have presented these results. The control group may have had higher post-test scores for skill and knowledge and lower post-test scores for anxiety than the experimental group, and may have used visiting nurses less. This would be an interesting trend in the data.

McBride

It is difficult to assess whether all the data are presented as we are not given a copy of the research questionnaire, nor are we told how many questions were asked. The researcher may have included questions on topics such as nurses’ ability to undertake health promotion given their current workloads, and then have omitted the data because it did not fit her theories. It would appear that some of the data obtained from hospital consultants is missing (see Tables 3 and 5).

Exercise 19. Presentation of data

Allen et al.

Two tables are presented (page 307). These are not frequency tables as no figures relating to number of patients or individual test scores are given. The tables are used to present the results of statistical tests on the pre-test and post-test KST and A-State scores of the experimental and control groups. This seems an appropriate way to present this data.

The tables are quite clear, although the title for Table 1 might have been better placed (above the table, for example). The labelling is clear, the significance level is given and significant *P* values are marked with an asterisk. The tables add to the information given in the final paragraph under Results.

McBride

Data are presented in frequency tables (Tables 1–3 and 5–8) and histograms (Fig. 1 and 2). Bar charts would have been more appropriate than histograms, as both figures present data on discrete variables such as occupational group and work content.

There are a number of presentational flaws, particularly in the tables. In Table 1, totals should have been given at the bottom of the table for enrolled nurses, registered general nurses and total nurses. It might also have been useful to know the sample size and the response rate for each of these groups.

The title of Table 2 would have been more appropriate if the word “other” had been omitted. In addition, sample sizes should have been given as well as the number of respondents (it is not immediately obvious that the column labelled n refers to respondents rather than sample).

In Tables 3 and 5–8, the numbers of respondents should have been presented along with percentages. In Tables 3 and 5 we are not told the number of responses on which the percentages were based (5 of the 21 dietitians may have failed to answer and percentages may have been based on the answers given by the remaining 16). It is unclear why most of the data relating to hospital consultants has been omitted from Tables 3 and 5: the only item presented in Table 3 (42.8%) does not tally with what is written in the first paragraph of Professional aspects (54.9%).

Table 4 is a mystery! It is unclear whether the classification system presented here was derived from the survey data or from another source. In the text (pages 95–96) the author claims that uncertainty and inconsistency regarding issues of control and empowerment “leads to” a variety of practice styles and relationships. No statistical evidence is presented to support this claim, nor does the author state that it is based on a qualitative evaluation of respondents’ comments.

Exercise 20. Discussion of data

Allen et al.

The authors do acknowledge several limitations of their study: these include the possibility that patients obtained information from other sources and that the sample size was too small. They do not specifically state that the statistical tests associated with their hypotheses might have proved significant with a larger sample size. It might have been better to present the two limitations above under the subheading Limitations. The researchers acknowledge that the KST may have lacked validity: we cannot judge this for ourselves without a copy of the document. They also acknowledge that extraneous variables such as concurrent medical conditions, difficult personal situations and having English as a second language affected the length and frequency of postoperative visiting. The researchers do not acknowledge that degree of visual impairment might have affected test results both preoperatively and postoperatively. Some of the details included under Limitations do not seem to belong there, such as discussion of the relative importance of knowledge and skill.

The first two sentences of the Discussion make claims that cannot be supported. The experimental home care programme was not tested against “the normal teaching” since all subjects received the latter regardless of whether they were in the experimental or the control group (page 306). The authors discuss the relative importance of knowledge and skill acquisition in the light of results that were not significant (pre-test and post-test knowledge scores in the control group and amounts of nursing time required by experimental and control groups). This discussion fails to take account of the fact that the results might have achieved statistical significance if a larger sample size had been used.

The article is, for the most part, written in a tentative style. The authors do not tend to make sweeping generalizations about all cataract patients or claim that they have proved anything. The one exception to this is the statement beginning “It is evident that ...” at the end of the second paragraph of Limitations. It is not wise to make such a claim, particularly about a study with so many uncontrolled extraneous variables.

McBride

The author does not acknowledge any study limitations, in spite of the flaws in sampling and instrument design noted above. She does note, however, that the data relating to C and G grade nurses are “unreliable” because of the small number of respondents (page 97).

As noted above, presentation and discussion of the data are combined. In general, there is little discussion of the data presented (see Practical policy changes, page 95, for instance). Where discussion does occur, it is sometimes unclear how it relates to the data collected (see Linking primary and secondary care, page 96). A few of the author’s ideas are well developed and credible (as in Involvement of hospital nurses in health promotion, page 97), although some of the discussion should have been supported with references. For example, the statement “It has been suggested that hospital admission is a stressful time for patients so health promotion is inappropriate” might have been supported (or challenged) with references on learning theory.

The author largely attributes the attitudes studied to survey respondents only, as in “Of the nurses, 73.3% felt that organized preventive care needs good systematic records” (Systematic recording, page 96). In places, however, she generalizes the findings to wider populations of nurses and health professionals, as in the final two paragraphs of Effectiveness of health promotion by hospital nurses, (page 97). In addition, the statements in these paragraphs are not tentative in that they are not couched in terms such as “It would appear ...” or “The findings suggest ...”.

Exercise 21. Discussion of data

King & Jensen

The data are presented and discussed in the section entitled Results. There is further discussion (and reference to relevant literature) in the Discussion but, as this is more like a conclusions and recommendations section, it will be considered later.

The authors’ theory seems credible. They focus on the phases that women go through and the strategies that they engage in while struggling to preserve the self before, during and after cardiac surgery. The various parts of the theory fit well with the core theory of preserving the self. The authors present a number of enlightening and substantial interview extracts and discuss deviant cases, particularly in the discussion of Moving on (page 103).

It is not easy to think of competing theories that would have been as coherent and well integrated as the one presented. Having said this, it should be acknowledged that the researchers have, of necessity, selected and presented data that are congruent with their theory. Despite the small sample size, the authors develop discussion relating to specific groups of interviewees, such as the age-related discussion in Being there (page 103). It seems possible that (despite claims that categories were saturated) a greater number of interviewees might have yielded a wider range of views, which may not have been so easily categorized by age (for instance, not all older people readily accept help).

Fig. 1 (a flow diagram) is presented to aid comprehension of how process and strategy relate to each other within the theory. The title places them within the framework of the core category. The diagram is confusing and is, therefore, unlikely to aid understanding of the theories.

Exercise 22. Conclusions and recommendations

Allen et al.

The authors recommend that “the effectiveness of formal information sharing warrants further investigation” and that their own study “needs to be repeated with a larger sample size which includes patients having day surgery”. They also suggest, somewhat obliquely, that the optimal content of preoperative education (in terms of information needs and skills practice) requires investigation. The authors discuss concurrent medical conditions, difficult personal situations and language problems and conclude that “future studies would need a qualitative dimension to ascertain the presence and influence of these multiple confounding factors”. All these recommendations are made in the Discussion: it might have been better to state them, or to repeat them, in the Conclusion. In the Conclusion the authors recommend research into “cost-effectiveness and appropriate target population” for in-home preoperative teaching.

All these recommendations are for further research.

The recommendations seem reasonable, even though the one about skill and knowledge may have been prompted by statistical results based on an inadequate sample size. In addition, the recommendation made in the Conclusion is mainly based on patients’ comments and researchers’ impressions rather than on numerical data. None the less, it seems a reasonable suggestion to make. Suggestions for further research may be legitimate and welcome whatever their source, whereas recommendations for changing practice need to be based on a body of credible evidence.

McBride

It is quite difficult to identify specific recommendations for practice and research in this article. The last paragraph, however, suggests that “a co-ordinated comprehensive research-based strategy” for the planning of hospital-based health education and health promotion needs to be devised, and that plans should also take account of preventive work in primary care.

It is not clear whether “research-based” means based on existing research or based on further research. The suggestion that a strategy be devised appears to be a policy recommendation designed to change practice.

The author might have emphasized the importance of further research rather than making policy recommendations to change practice. This is particularly the case since she suggests that there is little existing literature on health promotion in hospitals (see discussion of The literature review above). It would be unwise to base policy on evidence presented in just one article (particularly given the flaws discussed above).

King & Jensen

The authors make comments in the Discussion that amount to a recommendation that nurses counsel women, their families and friends about postoperative activity (including housework) and the need for social support (page 104). They also state, in the closing sentence, that “Further nursing research acknowledging women’s perception of self will help develop an understanding necessary to effectively collaborate with women having cardiac surgery”.

The first recommendation relates to clinical practice and the second to research.

These recommendations arise from the data. In particular, the first recommendation relates mainly to the phase named Being there. It would seem quite reasonable to implement them,

although the former might have implications for nurses' workloads and cost-effectiveness that might require investigation.

Exercise 23. Ethical considerations

Allen et al.

The researchers state in the first sentence under Procedure (page 306) that they obtained informed consent. No further details are given, so we have to assume that this included assurances of confidentiality, as is usual.

No other ethical issues were identified by the researchers.

Since members of the control group did not receive the home care preoperative education package, they might have had unmet information needs. They might, therefore, have asked the researchers for relevant information about cataract surgery and the postoperative period. The giving of such information would compromise the study design, but the withholding of information could lead to patient anxiety and a poor outcome from surgery. Researchers may have had to decide how to deal with this ethical dilemma, but this is not discussed.

McBride

Formal informed consent is not mentioned: this is because it describes a questionnaire survey. The researchers might, however, have mentioned whether they distributed explanatory letters along with the questionnaires. This is usual practice as, at the very least, the purpose of the research and issues of confidentiality need to be explained to potential respondents.

No other ethical issues were identified.

Apart from the issue of confidentiality mentioned above, no unidentified or unaddressed ethical issues are apparent.

King & Jensen

Informed consent is mentioned at the end of the penultimate paragraph in the Method section. Again, we have to assume that assurances of confidentiality were given.

No other ethical issues were identified.

Issues raised under the Results and Discussion sections suggest that the researchers explored some potentially distressing subject areas related to disability, self-concept and relationships. The researchers do not mention that they were aware of the possibility of causing emotional upset and what steps they intended to take (or took) to avoid or minimize this. In addition, they might have emphasized the importance of confidentiality in this study, as linkage of data with specific women might have caused difficulties in relationships at home.

Exercise 24. Rigour

Allen et al.

The internal validity of this study is undermined by all the uncontrolled extraneous variables detailed in the discussion of exercises under Discussion of data. The fact that the researchers changed from a between groups to a within groups design at the data analysis stage complicates any assessment of internal validity. The internal validity of the study is subject to the following five threats.

1. *History.* The cataract operation is an event that could affect the dependent variables, particularly anxiety and skill. There is no reason, however, to suspect that the operation had differential effects on the experimental and control groups. Members of the control group may have sought and gained information, similar to that in the teaching programme, from other sources.
2. *Reactivity.* The participants may have been in favour of preoperative home care teaching. The control group, in particular, may have tried to emphasize their anxiety and lack of knowledge and skill postoperatively.
3. *Selection bias.* Although the sample appears to have been randomly allocated to experimental and control groups, it is too small to ensure that the groups had no pre-existing differences. For example, we do not know whether the groups were similar in terms of preoperative visual impairment.
4. *Mortality.* Since we are not given figures for the sizes of the experimental and control groups, and are not told whether any subjects dropped out, concerns about mortality exist.
5. *Lack of reliability.* There is no mention of reliability testing of the KST. We are not provided with a copy of the instruments so that we can evaluate them.

The validity of the instruments is also in doubt because no copies are provided. This is perhaps less of a problem with the STAI, which had been validated in a number of previous studies (page 305). The KST had never been used before, although the researchers do state that they assessed its content validity, using a panel of experts (pages 305–306). The researchers acknowledge some of the above flaws in their Discussion.

The study is also low in external validity; it seems quite likely that different results would be obtained if the study were replicated using an adequate sample (as noted in the discussion of exercises on Sampling procedures and Discussion of data). The authors acknowledge that the sample was too small.

The reliability of the data collection instruments (particularly the KST) must also be subject to doubt, since we are not given the opportunity to evaluate them. The STAI had been evaluated in terms of “stability and internal consistency reliability” in previous studies (page 305) but no mention is made of tests to ensure the reliability of the KST.

McBride

We cannot fully assess internal validity in this study as we are not provided with a copy of the research instrument. We are told that the questionnaire had been used previously in a survey of general practitioners (page 93) but we are not told whether its validity was assessed. In this study, pilot work was carried out but no modifications to the questionnaire were deemed necessary. This is quite unusual and may cast doubt on the rigour of the piloting procedures.

The questionnaire items presented in the article (pages 95 and 98) suggest that this study is low in internal validity. All the items lead respondents towards expressing positive attitudes to hospital nurses’ involvement in health promotion. We do not know if these were balanced by items that led in a negative direction, but there must be strong suspicions that this was not the

case. This is a problem because some survey respondents tend to agree with statements whatever their content (automatic agreement is more common than automatic disagreement) and some tend to answer in a way that they feel is acceptable to the researchers or to society as a whole (a form of reactivity). The results of this study may reflect these tendencies rather than genuine attitudes. This is not acknowledged by the researcher.

The external validity of this study is also low because of the use of inadequately sized convenience samples. On page 97 the author describes some of the data as “unreliable” because it is based on responses from small numbers of nurses. In fact, all the data lack external validity and thus generalizability because of sampling flaws.

We cannot fully evaluate the reliability of the questionnaire as no copy is provided. Two of the items presented on pages 95 and 98 are flawed as they involve more than one question. The triple-barrelled question on page 95 was highlighted in the discussion of exercises on Data collection methods. The statement “Nurses can change people’s lifestyles despite cultural and environmental influences” (page 98) puts respondents in a similar quandary, as they may think that the statement is true of cultural influences but not of environmental influences (or vice versa). These two questions are unreliable as respondents are unable to answer them with any accuracy. There may have been several other unreliable items in the questionnaire. Reliability testing is not mentioned in the article.

Exercise 25. Rigour

King & Jensen

The researchers do not go to great lengths to establish the rigour of their work. They tend to use terms drawn from quantitative research such as “validation” and “generalizability” (see the first paragraph of Method) rather than the alternative terms suggested by Sandelowski (18) and Lincoln & Guba (19). They state that “Concepts and constructs that emerged from the data were tested as they emerged by going back to the informants and to current research for validation”. This statement refers to attempts to establish both internal validity (through a form of respondent validation) and external validity (through backfitting the data to existing literature). The authors suggest that studies based on grounded theory have “limited generalizability” (external validity) but do not expand on the reasons for this.

The researchers do not take a reflexive approach to establishing internal validity. There is, therefore, little or no discussion of how interview environment and interviewer/interviewee interaction influenced the data collected or acknowledgement that data is subject to researchers’ interpretations. The authors do enhance internal validity by providing substantial extracts from the interviews and discussing deviant cases.

Although reliability is not discussed, this is by no means unusual in qualitative research. However, the authors might have given rather more details about decision-making processes, particularly during data collection, in order to establish the rigour of their work.

1. References

1. BURNARD, P. & MORRISON, P. *Nursing research in action: developing basic skills*. London, Macmillan, 1990.
2. DREW, C.J. *Introduction to designing and conducting research*. 2nd ed. Missouri, C.B. Mosby Company, 1980.
3. PERFILJEV, G. & EGOENKOVA, Y. *Management and leadership: glossary for Russian-speaking nursing leaders*. Copenhagen, WHO Regional Office for Europe, 1993 (document).
4. HOCKEY, L. The nature and purpose of research. In: Cormack, D.F.S., ed. *The research process in nursing*. Oxford, Blackwell Science, 1996.
5. BRIGGS, A. *Report of the Committee on Nursing*. London, H.M. Stationery Office, 1972.
6. *Guidelines for nursing research development*. Geneva, International Council of Nurses, 1985.
7. CLARK, E. & KEEBLE, S. *Health care research in professional practice*. London, Distance Learning Centre, South Bank University, 1995 (Research Awareness Module 1).
8. SADLER, C. Sleeping like a baby. *Nursing times*, **88**(7): 18–19 (1992).
9. RUNDELL, S. Research has shown ... *Nursing times*, **88**(7): 24 (1992).
10. *Research-mindedness and nurse education*. London, Royal College of Nursing, 1982.
11. POLIT, D.T. & HUNGLER, B.P. *Essentials of nursing research: methods, appraisal and utilization*. Philadelphia, J.B. Lippincott Co., 1989.
12. LEININGER, M. *Qualitative research methods in nursing*. Orlando, Grune and Stratton, 1985.
13. HAMMERSLEY, M. & ATKINSON, P. *Ethnography: principles in practice*. London, Routledge, 1983.
14. GLASER, B.G & STRAUSS, A.L. *The discovery of grounded theory*. Chicago, Aldine, 1967.
15. MAY, K.A. Interview techniques in qualitative research: concerns and challenges. In: Morse, J.M., ed. *Qualitative nursing research: a contemporary dialogue*. London, Sage, 1989.
16. *Ethics related to research in nursing*. London, Royal College of Nursing, 1993.
17. *International code of ethics for midwives*. London, International Confederation of Midwives, 1993.
18. SANDELOWSKI, M. The problem of rigor in qualitative research. *Advances in nursing science*, **8**(3): 27–37 (1986).
19. GUBA, E.G. & LINCOLN, Y.S. *Effective evaluation*. San Francisco, Jossey-Bass, 1981.

Annex 1

FURTHER RELEVANT MATERIALS FROM THE LEMON PROJECT

Introduction

The relevance of research to nursing practice

The LEMON package³ provides educational material for nurses, much of it based on research findings. It contains specific chapters on clinical skills, communication skills and other relevant subjects.

Carrying out a literature search

How to do a literature search

Chapter 12 of the LEMON package, entitled Bibliography and useful information, also provides material relevant to this subject.

In some countries or in rural areas, it may be necessary to write to and/or telephone a national or international library to do the search on your behalf. Some may not provide a search service but may be able to provide copies of articles that you have already identified. Your LEMON group may be able to give you more information about libraries in your local area/country.

Country LEMON groups will insert names of major libraries in their countries that can be used by nurses and (if possible) whether they provide search facilities.

It is useful to make photocopies of selected journal articles and keep them. However, your country may have laws about the photocopying of documents and there are also international laws relating to copyright (the protection of authors and publishers), a subject dealt with in LEMON Chapter 12. In addition, your chosen library may not have photocopying facilities. The librarian should be able to tell you about this. You may, therefore, have to read journal articles in the library and make notes on them.

Country LEMON groups to state which libraries can access an inter-library lending service.

In addition, the library of the Royal College of Nursing in the United Kingdom, the largest nursing library in Europe, can provide you with copies of journal articles via your country LEMON group (see page 74).

LEMON Chapter 12 gives useful information on purchasing books and subscribing to journals, as well as a bibliography of WHO publications and other material available in your country.

³ The WHO Learning Materials on Nursing (LEMON) package was developed to ensure universal availability of appropriate, high-quality learning materials on nursing in all official languages in central and eastern European countries and in the newly independent states of the former USSR in order to assist the process of nursing development for better health care. The project has been devolved to country level, and the focal point for nursing and the Ministry of Health in each country will be able to provide you with further information.

Reading and evaluating research reports

Data collection methods

Questionnaires should be designed with a high proportion of closed rather than open questions. You can read more about open and closed questions in LEMON Chapter 3 entitled Communication. Questions should be arranged in a logical order and should avoid complexity and ambiguity. Pilot work (as discussed above) can help researchers to identify and eliminate questionnaire design flaws.

Ethical considerations

LEMON Chapter 9 entitled Professional and moral issues deals with ethical aspects of nursing but does not place a specific emphasis on research. You may, however, wish to read or reread this chapter, as it gives a much more detailed discussion of philosophical and legal aspects of ethics than can be provided here.

Where to go from here

Ideally, you should design and conduct your own research before making changes but you need further training in research methods to enable you to do this. Contact your country LEMON group for information about research courses.

Country LEMON Groups will insert the contact address of their group here.

You could also contact the address above if you want to know more about LEMON Chapters 1–12.

Annex 2

RESEARCH ARTICLES

ARTICLE 1

Journal of advanced nursing, 17: 303–309 (1992)

Effectiveness of a preoperative teaching
programme for cataract patients

Marion Allen, PhD RN

Professor, Faculty of Nursing, University of Alberta, Edmonton, Canada

Cheryl Knight MN RN

Clinical Nurse Specialist, Good Samaritan Auxiliary Hospital, Edmonton, Canada

Coleen Falk BScN RN

Former Executive Director, Victorian Order of Nurses, Edmonton, Canada

Vicki Strang MN RN

Assistant Professor, Faculty of Nursing, University of Alberta, Edmonton, Canada

Reproduced by kind permission of Blackwell Science Ltd,
Oxford, United Kingdom

ARTICLE 2

Journal of advanced nursing, **20**: 92–100 (1994)

Health promotion in hospitals: the attitudes,
beliefs and practices of hospital nurses

Anita McBride, BH (Hons) MA RGN
Associate Director, Centre for Health Promotion Evaluation,
Oxford, United Kingdom

Reproduced by kind permission of Blackwell Science Ltd,
Oxford, United Kingdom

ARTICLE 3

Heart & lung – the journal of critical care, **23**: 99–105 (1994)

Issues in cardiovascular care.

Preserving the self: women having cardiac surgery

Kathryn M. King, RN, MN & Louise Jensen, RN, PhD
Edmonton, Alberta, Canada

Reproduced by kind permission of Mosby-Year Book Inc.,
St Louis, MO, USA