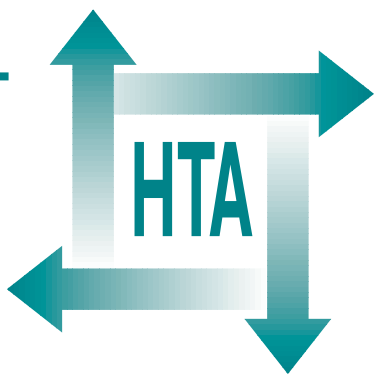
Health Technology Assessment 1999; Vol. 3: No. 20

Methodology

Factors that limit the quality, number and progress of randomised controlled trials

RJ Prescott
CE Counsell
WJ Gillespie
AM Grant
IT Russell
S Kiauka
IR Colthart
S Ross
SM Shepherd
D Russell

Health Technology Assessment
NHS R&D HTA Programme



Standing Group on Health Technology

Current members

Chair:
Professor Kent Woods
Professor of Therapeutics,
University of Leicester

Professor Martin Buxton
Director & Professor of
Health Economics,
Health Economics
Research Group,
Brunel University

Professor Shah Ebrahim
Professor of Epidemiology
of Ageing, University of Bristol

Professor Francis H Creed
Professor of
Psychological Medicine,
Manchester Royal Infirmary

Professor John Gabbay
Director, Wessex Institute
for Health Research
& Development

Professor Sir John
Grimley Evans
Professor of Clinical Geratology,
Radcliffe Infirmary,
Oxford

Dr Tony Hope
Clinical Reader in Medicine,
Nuffield Department of
Clinical Medicine,
University of Oxford

Professor Richard Lilford
Regional Director of R&D,
NHS Executive West Midlands

Dr Jeremy Metters
Deputy Chief Medical Officer,
Department of Health

Professor Maggie Pearson
Regional Director of R&D,
NHS Executive North West

Mr Hugh Ross
Chief Executive,
The United Bristol
Healthcare NHS Trust

Professor Trevor Sheldon
Joint Director, York Health
Policy Group, University of York

Professor Mike Smith
Faculty Dean of Research
for Medicine, Dentistry,
Psychology & Health,
University of Leeds

Dr John Tripp
Senior Lecturer in Child Health,
Royal Devon and Exeter
Healthcare NHS Trust

Professor Tom Walley
Director,
Prescribing Research Group,
University of Liverpool

Dr Julie Woodin
Chief Executive,
Nottingham Health Authority

Past members

Professor Sir Miles Irving*
Professor of Surgery,
University of Manchester,
Hope Hospital, Salford

Dr Sheila Adam
Department of Health

Professor Angela Coulter
Director, King's Fund,
London

Professor Anthony Culyer
Deputy Vice-Chancellor,
University of York

Dr Peter Doyle
Executive Director, Zeneca Ltd,
ACOST Committee on Medical
Research & Health

Professor John Farndon
Professor of Surgery,
University of Bristol

Professor Charles Florey
Department of Epidemiology
& Public Health, Ninewells
Hospital & Medical School,
University of Dundee

Professor Howard
Glennester
Professor of Social Science
& Administration, London
School of Economics &
Political Science

Mr John H James
Chief Executive,
Kensington, Chelsea &
Westminster Health Authority

Professor Michael Maisey
Professor of
Radiological Sciences,
Guy's, King's & St Thomas's
School of Medicine & Dentistry,
London

Mrs Gloria Oates
Chief Executive,
Oldham NHS Trust

Dr George Poste
Chief Science & Technology
Officer, SmithKline Beecham

Professor Michael Rawlins
Wolfson Unit of
Clinical Pharmacology,
University of Newcastle-
upon-Tyne

Professor Martin Roland
Professor of General Practice,
University of Manchester

Professor Ian Russell
Department of Health Sciences
& Clinical Evaluation,
University of York

Dr Charles Swan
Consultant Gastroenterologist,
North Staffordshire
Royal Infirmary

* Previous Chair

Details of the membership of the HTA panels, the NCCHTA Advisory Group and the HTA Commissioning Board are given at the end of this report.

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

Factors that limit the quality, number and progress of randomised controlled trials

RJ Prescott[1*]
CE Counsell[2]
WJ Gillespie[3]
AM Grant[4]
IT Russell[5]
S Kiauka[1]
IR Colthart[1]
S Ross[4]
SM Shepherd[6]
D Russell[5]

[1] Medical Statistics Unit, University of Edinburgh, UK
[2] Department of Clinical Neurosciences, University of Edinburgh, UK
[3] Department of Orthopaedic Surgery, University of Edinburgh, UK
[4] Health Services Research Unit, University of Aberdeen, UK
[5] Department of Health Sciences, University of York, UK
[6] Department of Public Health Sciences, University of Edinburgh, UK

[*] Corresponding author

Competing interests: none declared

Published December 1999

This report should be referenced as follows:

Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, *et al.* Factors that limit the quality, number and progress of randomised controlled trials. *Health Technol Assess* 1999;**3**(20).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA web site (see overleaf).

Related publications

Prescott RJ, Gillespie WJ, Counsell CE, Grant AM, Ross S, Russell IT, *et al.* Factors that limit the number, quality and progress of randomised trials. In: Black N, Brazier J, Fitzpatrick R, Reeves B, editors. A guide to good practice. London: BMJ;1998. p.86–97.

Ross S, Grant AM, Counsell CE, Gillespie WJ, Russell IT, Prescott RJ. Barriers to participation in randomised controlled trials: a systematic review. *J Clin Epidemiol.* In press.

NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Methodology Group and funded as project number 93/43/02.

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein
Monograph Editorial Manager: Melanie Corris

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.hta.nhsweb.nhs.uk

Contents

List of abbreviations

CI	confidence interval
CONSORT	Consolidated Standards of Reporting Trials
DMC	data monitoring committee
ECG	electrocardiogram*
GP	general practitioner
ITT	intention-to-treat
NIH	National Institutes of Health (USA)
NNT	number-needed-to-treat*
NSAID	non-steroidal anti-inflammatory drug
OR	odds ratio*
RCT	randomised controlled trial
SORT	Standards of Reporting Trials

* Used only in tables

Executive summary

Background

The randomised controlled trial (RCT) is the most powerful research tool for evaluating health technologies. However, for most therapeutic activities with the NHS, reliable information from RCTs is not available.

Objectives

- To assemble and classify a comprehensive bibliography of factors limiting the quality, number and progress of RCTs.
- To collate and report the findings, identifying areas where firm conclusions can be drawn, and identifying areas where further research is required.

Methods

A systematic review of the literature was undertaken, covering the period 1986–96. The scope of the review was too broad to be comprehensive in all of the areas covered, rather it attempted to cover the diversity of factors limiting the quality, number and progress of RCTs.

The issues considered were those of design, barriers to participation, conduct and structure, analysis, reporting and costs.

Results and recommendations for practice

Design

Following a systematic review of existing evidence, a well-formulated question should be developed, specifying participants, interventions and outcomes. Wide patient eligibility criteria are generally preferred to give representativeness and good recruitment rates. However, a more homogeneous group may be preferable when evaluating expensive or hazardous interventions. Outcome measures need to be clinically and socially relevant, well-defined, valid, reliable, sensitive to important change and measured at appropriate times. There is evidence that the use of intermediate or surrogate outcomes has been misleading.

The most frequent choice of study design is between a parallel group or a crossover design. Simultaneous investigations of two or more treatments are efficiently approached by using a factorial design. Simple parallel group designs with fixed sample sizes are most common but other designs should be considered.

Protection from selection bias is provided by secure random allocation, using telephone- or computer-based randomisation, and by analysis based on the groups as allocated, thus ensuring that groups being compared differ only by chance. Performance bias can be minimised by blinding treatments (when possible) and by employing clearly described treatment policies. Detection bias may be avoided by blind outcome assessment and attrition bias by ensuring follow-up of all patients randomised.

Pre-study sample size calculations should always be made and funding bodies, independent protocol review bodies and journal editors should all demand them. A sensitivity analysis should be considered, with indicative estimates rather than unrealistically precise numbers. Small trials should be reported as hypothesis forming.

Barriers to participation

Barriers to clinician participation include: time constraints, lack of staff and training, concern about the impact on doctor–patient relationships, concern for patients, loss of professional autonomy, difficulty with consent procedures, lack of reward and recognition, and an insufficiently interesting question.

Barriers to patient participation include: additional demands of the trial, patient preferences, concern caused by uncertainty and concerns about information and consent.

To overcome barriers to clinician recruitment, a trial should address an important research question and the protocol and data collection should be as straightforward as possible, with demands on clinicians and patients kept to a

minimum. Dedicated research staff may be required to support clinical staff and patients. The recruitment aspects of an RCT should be carefully planned and piloted.

Conduct and structure

Many trials fail to start, mainly because of lack of funding or logistical problems. Of those that start, half have recruitment difficulties, leading to abandonment or reduced size and, hence, loss of statistical power. Recruitment problems may be reduced by piloting, using multiple recruitment strategies, making contingency plans in case recruitment is slow, and using recruitment coordinators. None of these approaches has been rigorously evaluated.

Inadequate compliance with the study protocol can lead to false-negative or false-positive results. Some assessment of compliance (clinician and participant) should be made but may be difficult to measure.

Quality control is important but too much may make RCTs prohibitively expensive and hinder recruitment. Trials need good organisational and administrative bases but there is little research evaluating the optimal structure. The precise roles of steering committees and data monitoring committees have been poorly evaluated. There is concern about bias in the design, conduct, analysis and reporting of commercially sponsored trials, and independent monitoring should be considered.

Analysis

Intention-to-treat analysis is the method of choice to provide an unbiased estimate of treatment effects. In studies where the aims are more explanatory than pragmatic, consideration should be given to reporting analysis by treatment received as well as intention-to-treat.

Study protocols should identify a predetermined primary outcome supplemented by secondary outcomes and a clear statistical plan. Any subgroup analyses that are proposed as hypothesis testing should be specified in the protocol and the study must be of sufficient size to detect such an interaction. All other subgroup analyses should be considered as hypothesis-generating.

Reporting

The introduction of the Consolidation of Standards for Reporting Trials (CONSORT) guidelines should improve reporting of RCTs. Conclusions should be supported by the data presented. About 10% of trials remain unpublished while many others are only published in conference proceedings, particularly if they are small and show non-significant treatment effects: prospective registration of all RCTs is recommended. Multiple publication of a study is also a problem for studies showing significant results.

Costs

Economic evaluations are reported in few RCTs, possibly because of difficulties in conducting such evaluations and the lack of generalisability from one healthcare context to another. Some components of an economic analysis are subject to uncertainty; statistical tests and confidence intervals should, therefore, be used.

There has been little research into trial costs but costs of caring for patients in RCTs may be perceived as an unaffordable new service, delaying or preventing recruitment at some participating centres.

Conclusions

The evidence available to guide many aspects of the design, conduct and analysis of RCTs is not always being applied.

Recommendations for research

Further research is required, particularly in relation to:

- problems being experienced and solutions employed in current RCTs
- the optimum structure, staffing and organisation for the conduct of large and small trials
- the factors which influence the participation of clinicians and patients in trials.

Chapter 1
Introduction

Background

The randomised controlled trial (RCT) is widely accepted as the most powerful research method for evaluating health technologies. Its principal strength is that it minimises bias (the risk of being misled by systematic errors). Protection from selection bias is provided by random allocation to alternative technologies and analyses based on the groups as allocated, thereby ensuring that the groups being compared differ only by chance. Ascertainment bias can be avoided by arranging that the outcome is assessed in ignorance of the treatment allocated, and co-intervention bias minimised by blinding treatments (where possible) and by employing clearly described treatment policies, which should be identical for each group apart from the intervention under examination in the RCT.

These principles are now widely recognised and understood both by those responsible for decisions about policy and practice, and by clinical and health service researchers. Decision makers within the NHS are increasingly looking to evidence from RCTs, for example, through the Cochrane Collaboration and NHS Centre for Reviews and Dissemination. Many funding agencies now expect that the research methodology for comparing approaches to care should be an RCT unless there is a very good reason to justify an alternative approach. Furthermore, systematic searching for controlled trials using electronic bibliographies and handsearching of key journals is making it possible to describe the 'epidemiology' of RCTs; tens of thousands have been identified across most areas of health care.

Why then is reliable information from RCTs not currently available for most therapeutic activities within the NHS?

In part, the answers must lie with the failure, until recently, of healthcare systems to provide up-to-date systematic reviews of evidence. It will be some years before that situation changes. However, the answers also relate to the design, conduct and analysis of RCTs, and to professional and public perceptions of the place of research within the NHS.

Commonly, trials have failed to address questions which are of current importance to the health service. There are a number of reasons for this. The first is that technological developments may make the subject of a trial irrelevant. Inevitably, a major treatment advance may mean that the results of an RCT assessing a previous treatment modality are no longer seen as useful. Also, the choice of comparison in many trials has been dictated by the vested interests of both funders (commercial companies and charities) and power groups within the research community. Second, generalisability may be limited by the choice of participants, sometimes so tightly defined that the results have little relevance in practice. End-points are often intermediate or surrogate, and important to researchers rather than to patients.

Almost invariably, RCTs are too small to allow reliable identification of the sizes of effect that might plausibly be expected. Where there are no clear differences, the estimates of treatment effect are usually so imprecise that it is impossible to rule out a clinically important difference between the managements being compared. There is, therefore, a high risk of type II statistical errors, that is, the false conclusion that there is no difference when one does exist. Furthermore, true differences may also be widely distorted in small trials as only large estimates reach 'statistical significance'. Although pooling of data in a meta-analysis may sometimes resolve this problem, it needs to be carried out with great care, is labour intensive, and its findings may be distorted by publication bias, inadequate reporting of methods in published trials, insecure random allocation,[1] large losses to follow-up, and failure to base analyses on groups as allocated (intention-to-treat, ITT).

Original objectives

The original objectives of this review were specified as follows:

(i) to assemble a comprehensive bibliography reporting factors limiting the number, progress or quality of RCTs

(ii) to classify the bibliography as to the factors identified and the strength of information it provides about these factors
(iii) to undertake an in-depth analysis of factors associated with successful RCTs in three target areas: vertebral fractures and hip fractures, neurology and neurosurgery, and perinatal medicine
(iv) to collate and report the findings with particular reference to application in the NHS, identifying those areas where firm conclusions can be drawn and those where further research is required.

Before commencing the review we were unsure of the volume and quality of the literature which would be identified in meeting these aims. In preparing the protocol, it was decided that the approach would be systematic rather than necessarily comprehensive; thus, the decision on the final inclusion criteria would be reached in an iterative manner as the search progressed. Definitely included would be any RCTs comparing approaches to trial design, conduct, analysis or reporting. Other articles such as surveys or case studies of RCTs would, however, be read and categorised for relevance. If the literature were very large, then the number of included articles would be culled by focusing on those which were most relevant and most recent. The details of this process are described in chapter 2.

The search was based on a preliminary list of factors which might limit the quality, number, and progress of controlled trials in the NHS. These factors were as follows.

Barriers to patient participation
- Poor quality information for potential participants
- Suboptimal consent procedures
- Lack of awareness among those approached of the balance of risks and benefits
- Unwillingness to accept, or inability to cope with, uncertainty.

Barriers to clinician participation
- Pressures of time
- Perceived need for extra resources
- Unwillingness to admit uncertainty
- Belief that admission of uncertainty is not in the patients' best interests
- Fear of medico-legal problems
- Cumbersome ethics committee procedures
- Perceived lack of rewards and recognition for participation
- Competitive rather than collaborative view of medical research.

Other obstacles to trials
- Double standards in consent procedures: arrangements may differ inside and outside a trial
- Technology creep and the difficulty of timing trials
- Cumbersome commissioning procedures delaying a trial for so long that the window of opportunity (time of uncertainty) has closed
- Relatively high cost of RCTs
- Imbalance of funds available leading to selection of trials biased towards drugs
- Methodological issues around learning, skill and timing, particularly for technologies
- Pressure to introduce new technologies before proper evaluation
- Interference from other research projects, especially other trials (difficulties of participation in more than one trial)
- Recent NHS changes
- Service costs.

Poor design
- Unnecessary duplication (inadequate prior review of previous research)
- Over-elaborate design
- Rigidity caused by 'laboratory mentality'
- Restrictive trial entry leading to poor generalisability
- Insufficient ambition
- Unimportant questions ('me-too' trials)
- Inappropriate choice of outcomes
- Inadequate timescale
- Restricted group of participating practitioners.

Poor conduct
- No proper infrastructure
- Lack of understanding of trial organisation
- Failure to complete trial
- Failure to report trial findings, leading to publication bias
- Failure to stop a trial that is no longer necessary or which will not be useful
- Insecure randomisation
- Biased losses to follow-up
- Poor assessment of outcome
- Insufficient sample size leading to risk of type II errors or exaggeration of true differences.

Poor analysis
- Inappropriate comparison groups: failure to use ITT analyses
- Poorly conducted interim analyses
- Multiple end-points
- Multiple analyses of same end-point
- Lenient choice of level of significance

- Overemphasis on subgroup analyses
- Inappropriate statistical methods.

Poor reporting

- Failure to describe procedures in a generally accepted manner
- Failure to report confidence intervals (CIs) – type II errors
- Inappropriate conclusions
- Failure to put results in the context of other research.

Reasons for modifying objectives

Objective (iii) was revised to a much smaller scale, mainly because it became apparent that meeting objectives (i), (ii) and (iv) would occupy the group fully if any useful outcome was to emerge from objectives (i) and (ii). At first, it was agreed that the highest priority should be given to identifying and analysing methodological studies which provided rigorous evidence about how the design, conduct, analysis and reporting of RCTs affects their number, quality and progress. It was initially estimated that this might take up about half the allocated resource, and plans were made as to how objective (iii) might be met. As searching proceeded, it became apparent that reviewing factors limiting number, quality and progress of RCTs would occupy the group fully. Furthermore, only limited good quality data appeared likely to emerge from the published literature on the list of trial characteristics implied by our list of hypotheses, making it difficult to define the components of a successful trial. Thus, although unsuccessful trials would be difficult to identify, three possible methods were considered. First, trials included in the Cochrane Collaboration specialised trial registers but excluded from systematic reviews might be examined. Second, it might be possible to examine trials approved by ethics committees but not subsequently published, although this had been done before.[2,3] Third, principal investigators of sampled successful trials might be asked to nominate the unsuccessful trial known to them that most clearly resembled their successful trial, and the nominated unsuccessful and successful trials would then be compared.

The second and third of these options, it was decided, would involve substantial new research for which the available resources would not be adequate. Two existing specialised trial registers, however, were explored to identify the reasons why trials were excluded from systematic reviews. Also, one issue which prevented trials from being included in systematic reviews was examined in detail, namely surrogate outcomes; these data have been brought together in appendix 2.

Levels of evidence

In this study, the levels of evidence are stated in the individual sections, in relation to the recommendations made. This approach, rather than classification into rigid categories, was used because of the variety of types of evidence which can occur in this particular review. Unlike reviews of clinical effectiveness, such as those published in the Cochrane Database of Systematic Reviews, where the primary studies can be assessed according to pre-determined criteria, information relevant to this systematic review might come from primary research on trial methodology, from primary research on treatment (with information on trial methodology being secondary), from secondary research on trial methodology, from systematic reviews of treatment, or from articles which are primarily educational. Although the evidence-based approach is supported whenever possible, there are areas covered by this review where good practice is based on logic rather than evidence. An example would be the use of CIs in reporting the results of RCTs. Nevertheless, in our presentation of the evidence there will be an informal hierarchy, with evidence from RCTs or trials of methodology potentially giving the strongest evidence. Evidence derived from observational studies and surveys of trials or trialists occupies the next level. The lowest level is evidence derived from individual descriptions of identified problems and how they were dealt with. Within review articles, those with an explicit systematic methodological strategy will always be considered potentially stronger than non-systematic reviews.

Structure of this report

The strategies for identification and registration of studies, and the results of the search are discussed in chapter 2 of this report. Issues relating to the design of RCTs, barriers to undertaking RCTs and problems in the conduct of RCTs are described in chapters 3–5. The analysis, reporting and costs of RCTs are described in chapters 6–8. The main recommendations arising from the review are presented separately for practice and research in chapters 9 and 10. Finally, a critique of the study is presented in chapter 11. To simplify presentation, each chapter has its own reference list.

References

1. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

2. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;**337**:867–72.

3. Easterbrook PJ, Matthews DR. Fate of research studies. *J R Soc Med* 1992;**85**:71–6.

Chapter 2

Methodology of the review

Introduction

The general methods used in the review are described here. Additional details, specific to each chapter, are given subsequently. A description of some of the general principles of searching the literature (especially electronic databases) is followed by a description of the methods and sources used to obtain and assess the literature relevant to this review. Quality control measures are outlined, as are practical problems that were encountered with the conduct of the review.

Given the broad scope of the review, a decision was made to be systematic rather than necessarily comprehensive. For example, although the aim was to identify as many factors as possible that limit the number, quality, and progress of RCTs, that did not necessarily mean identifying all articles dealing with one specific limiting factor. This had implications for the search and ascertainment of the literature, and the final inclusion criteria were reached in an iterative manner as the search progressed.

Searching the literature

It is generally recognised that literature reviews should draw on relevant material from a variety of sources. The most commonly used sources are electronic databases, bibliographies within identified papers, and handsearches of key journals. In some contexts, such as reviews of effectiveness, it is important to identify unpublished material in order to reduce publication bias. This was considered impractical for this review (systematic searching of unpublished data would have required a survey of all known trialists) and so was not attempted.

Principles of searching electronic databases

Many bibliographic databases contain a controlled indexing vocabulary with keywords to enable efficient searches. Keywords are terms that describe the content of an article although not necessarily appearing in the title or body of the article. An experienced indexer has to read an article, understand the content and then assign the appropriate keyword. This process requires indexers to have knowledge of the topic covered in the article. For databases like MEDLINE, which contains articles from more than 3500 journals, the workload for indexers is very high and mistakes during the process are inevitable. Although initiatives like the MEDLINE Enhancement Project of the Cochrane Collaboration help re-index articles, the quality of indexing in most bibliographic databases is often poor.

Index words can be sorted or structured in a hierarchy to enable easier use. The resulting thesaurus requires the user to know the index terms as well as the structure of the thesaurus. Additions or changes to the thesaurus can cause problems in the conduct of systematic reviews. The indexing of similar articles may change over time so that some index terms may not cover the whole database. For example, the publication type 'controlled clinical trials' was added to MEDLINE in January 1996. Therefore, although some attempt has been made to re-index earlier articles, only a small proportion of the articles that were sought in this study had been indexed using this term.

Another way of searching for relevant articles is to use text search facilities. A term can be searched for in fields such as the title, abstract or journal. This method is particularly useful in combination with a keyword search to overcome indexing problems. For this study, a text word search was essential in the identification of articles describing problems with trials, not necessarily trials themselves. Text word searches offer great flexibility through the use of truncations, sometimes called 'wildcards', like '*'. For example, the term 'random*' will retrieve articles with any of the following terms in them: 'random', 'randomised', 'randomized', 'random-allocation'. Different published versions of bibliographic databases (such as the SilverPlatter or Ovid versions of MEDLINE) offer additional text operators with which the distance between two text words can be specified. For example, Ovid provides an 'adj *n*' operator which only retrieves references with two text words in the proximity of '*n*' words.

Synonyms and homonyms constitute the biggest problems in text word searches. Synonyms are words with similar meaning like 'difficulty', 'issue', or 'problem'. All synonyms of a word should be included in the text search strategy to ensure a sensitive retrieval. Homonyms are words with the same spelling but different meaning. For example, 'drop-out' does not only refer to patients who leave a study prematurely but also to students dropping out of school. If possible, homonyms should be combined with other text words or keywords to narrow their meaning and avoid ambiguity.

Selection of databases and time limits

It was initially decided to search six relevant bibliographic databases. However, in a pilot search, it was found that Sociofile and PsycLIT (two databases covering psychological and sociological aspects of health) did not yield sufficient numbers of new relevant articles. In addition, there was not enough time to search one database fully (SCI SEARCH), although it was used to identify a few references that cited other key references that had been found using other methods. Details of the three electronic databases that were searched and the periods covered are summarised in *Table 1*. Database searches were limited to the period 1986–April 1996, following the iterative piloting described below.

TABLE 1 *Electronic reference sources*

Bibliographic software	Period covered	Retrieval database
MEDLINE	1986–March 1996	Ovid v.3 rel. 6.1 (networked)
EMBASE	1986–April 1996	BIDS
CINAHL	1986–March 1996	SilverPlatter

Development of search strategies

For the electronic sources, specific search strategies were developed by one reseacher (SK) as each used different index terms. The database-specific searches are presented in appendix 1.

The total number of references retrieved electronically was 9732.

MEDLINE

Because of the very large number of references retrieved in an initial search on MEDLINE, it was decided to split the 10-year search period into two blocks of 5 years. For the period 1992–96 inclusive, a complex and comprehensive search strategy was used. Initially, the optimal MEDLINE search strategy developed for the Cochrane Collaboration to identify controlled clinical trials was used (see appendix 1). However, when this was combined with the 40 MeSH keywords (medical subject headings) relevant to this review, 66,000 references were obtained for the period in question. This number was clearly unmanageable and so the Cochrane strategy was modified to have a narrower focus on controlled trials by restricting it to the first two sections (down to step 19 in the strategy shown in appendix 1) and removing the MeSH term 'research design'. The new search consisted of three parts: a modified Cochrane search strategy to identify controlled trials; a keyword search combined with appropriate textwords to identify trials of study methodology; and an extensive textword search to retrieve references describing problems in the design, conduct, analysis or costs of a trial. This strategy yielded 7747 references.

For the period 1986–91, a limited search focused on papers describing trials or comparisons of trial designs and excluded many textwords describing problems. This search yielded 417 references.

It was noted that the continuous updates with new or re-indexed articles in MEDLINE made search strategies almost impossible to reproduce. The final version of our strategy in MEDLINE was used in April 1996. In May 1996 the database was upgraded, the 5-year blocks changed, and some articles were re-indexed. Within a month it was not possible to reproduce the data set retrieved for this review.

EMBASE

Since there is a large overlap in journals between EMBASE and MEDLINE, it was decided to limit the search to focus on review articles, trials of methodology and reviews of problems with RCTs. The number of keywords was limited and text words were used to focus the search. A total of 863 references were retrieved from EMBASE.

Cumulative Index of Nursing and Allied Health Literature (CINAHL)

For CINAHL a comprehensive search strategy, similar to that used for MEDLINE, was used. CINAHL contains keywords similar to MeSH headings which made it easy to use the previously developed search. Nevertheless, the search had to be slightly modified to allow the incorporation of terms that are unavailable on MEDLINE. The CINAHL search yielded a total of 705 references.

Handsearching

The intention, in the original grant proposal, was to handsearch three major journals: *Controlled Clinical Trials, Statistics in Medicine,* and *Clinical Trials and Meta-analysis* from 1986 to 1996; however, this was not possible because of logistic and financial difficulties. Two journals (*Controlled Clinical Trials* and *Clinical Trials and Meta-analysis*) were not available in Scotland. One possible way of overcoming this problem would have been to hire an experienced handsearcher to undertake this work at the British Library in Boston Spa, but resources were insufficient. However, the tables of contents of all volumes of *Controlled Clinical Trials* (kindly provided by Professor Stuart Pocock, London School of Hygiene and Tropical Medicine) were scanned for the years 1989–96. *Statistics in Medicine* proved to be less relevant to the project, as expected; most articles described statistical theory rather than practical problems in the analysis of trials.

Other methods of identifying articles

When each chapter of this review was written, further relevant articles were found by reviewing the bibliographies of included articles and from the personal knowledge of the authors.

Handling of references

All retrieved articles were downloaded from the electronic databases into the bibliographic management software, Reference Manager©. The package assigned a unique identifier to each reference which was later used to retrieve relevant articles. The source of the reference was also added to the bibliographic information.

Assessment of references

Following downloading into Reference Manager, an initial screening of references was undertaken by one research assistant (IRC). Titles, keywords and abstracts of all references were assessed and classified into 'possibly relevant' or 'irrelevant'. A sample of 500 references was assessed by a second research assistant (SK) and there was total agreement on the exclusion of articles initially classed as irrelevant. Irrelevant references were deleted. All 1690 'possibly relevant' references were downloaded into the program Access (© Microsoft Corporation) and copies of the original articles sought so that they could be read in full.

Some articles were unavailable within the time frame of the project. In addition, because of the costs of inter-library loans, it was decided not to order any single-page letters or comments that were not available locally. Other articles were not requested if the abstract indicated that it would provide, at best, low-level evidence about a topic already extensively covered. At this point, it was also decided to exclude 44 non-English language articles as translation would not have been affordable. A total of 480 'possibly relevant' articles were not obtained.

Full text articles were read and, if the articles were relevant, the information from them was extracted into a coding form (see appendix 1); this summarised the problem area, study methodology (primary research, reviews, other articles), subject area, intervention type, clinical activity (prevention, treatment, screening, or diagnosis), and setting (hospital, general practice, or community). Owing to time constraints, letters or short news items were not coded in full and only their titles were entered into the database.

Each full text article was assessed for its relevance in the eight different categories of the review by one of two reviewers (IRC, SK). The categories and the number of articles in each are listed in *Table 2*. Multiple entries for an article were possible because articles could cover several aspects relevant to the review. A numerical scoring system was developed that classified each article into highly relevant (3), relevant (2), borderline relevant (1), and irrelevant (0) for each area of the review. Some articles, despite having promising titles and abstracts, turned out to be of poor content and/or quality and were discarded as irrelevant. Both reviewers read and assessed those questionable articles. This procedure ensured that no possibly relevant article was discarded without double-checking. In addition, certain articles thought to be particularly important were tagged 'key article'. Once the review process was complete, one author (RP) went through the 185 'key articles' to ensure that they had been included in the report and any that were missing were reviewed once again to check whether they should be included.

A total of 638 articles were given a relevance score of 2 or 3 in at least one category. Not surprisingly, there was a significant overlap between categories such as design and conduct.

Within each of the eight problem areas, additional keywords were assigned to allow more specific retrieval at the time of synthesis (see appendix 1).

TABLE 2 *Number of articles assessed with relevance level by category (possible multiple responses)*

	Relevance level	Number in category
Design	0	422
	1	70
	2	132
	3	154
Analysis	0	657
	1	37
	2	49
	3	35
Costing	0	732
	1	22
	2	17
	3	7
Other obstacles	0	711
	1	18
	2	30
	3	19
Clinician participation	0	646
	1	22
	2	51
	3	59
Patient participation	0	554
	1	35
	2	85
	3	104
Reporting	0	666
	1	14
	2	34
	3	64
Conduct	0	498
	1	53
	2	102
	3	125

Synthesis

The retrieval and synthesis strategy for each problem area is described at the beginning of the relevant chapter.

Problems

A number of logistic and technical problems were encountered during the course of the project.

Problems with electronic databases

In this review a different syntax had to be used for each database because different publishers of bibliographic databases use different syntaxes. This requires additional time and knowledge of the different systems to 'translate' search strategy. A common basic syntax across the medical information industry would facilitate systematic review.

The network version of MEDLINE available was not able to save search strategies and this significantly increased the amount of time spent developing and implementing searches. Searchers proposing to use a network should ascertain the availability of a search-saving facility before committing to the network.

Reference Manager is a bibliographic package and had limitations in handling complex data relationships. The version used here had difficulty in identifying duplicate downloads if they came from different databases.

Inter-library loans

Unexpected costs were encountered for inter-library loans. Many higher education institutions subsidise their real costs but the advent of the systematic review has caused some to reconsider that policy. In designing a systematic review, researchers should not assume that any usual library subsidy will be available.

Chapter 3

Design issues

Methods

For most of the sections within this chapter, a set of relevant study keywords was identified, and a list of articles prepared, subdivided by the relevance score (0–3) for design. The keywords used for the areas covered in this chapter were as specified below:

- Randomisation – randomisation
- Choice of patient and clinical population – characteristics, eligibility criteria, result generalisability, experience
- Relevance of the research question – inappropriate comparison (plus articles obtained from reading undertaken under other headings)
- Type of trial design – rigid design, over-elaborate, washout, trial type (plus many articles obtained from reading undertaken under other headings)
- Sample size – sample size
- Use of a run-in period – run-in
- Trial outcome – outcome assessment, inappropriate outcomes, inappropriate end-points
- Timescale – timescale
- Practical aspects of trial design – protocol.

In most of the areas covered, all positively scoring articles were read in full. For the keywords 'eligibility criteria', 'randomisation' and 'sample size', with 29, 38 and 40 articles, respectively, scoring three, only these high-scoring articles were routinely read in full. For articles scoring one or two in these areas, abstracts were scanned and any appearing to cover aspects not previously encountered were also read in full.

Randomisation

Randomisation is generally regarded as the *sine qua non* for the comparison of treatments in a clinical trial to be potentially free from bias. It is the cornerstone of this review that only RCTs are considered. Nevertheless, there are often difficulties in conducting RCTs, and these are described later in chapters 4 and 5. This has led some authors to propose alternatives to conventional RCTs. Some of these alternative designs themselves involve the use of randomisation, such as the randomised consent design,[1] which has recently been reviewed by Altman and colleagues.[2] Such designs are not pursued here but it has been noted that several other papers identified by this search strategy consider alternatives to conventional RCTs.[3–9]

The process of randomisation has several components in which design considerations can apply. The first of these components is the technical one by which a randomisation sequence is generated. This may be based on simple randomisation, or may be considerably more complicated. The second component is the concealment of allocation before the next patient is entered (that is, the method used to ensure that the allocation is only known once formal 'irretrievable' trial entry has occurred). A third component is the timing at which the randomisation takes place (e.g. at initial enrolment or at the end of a run-in phase), and the fourth element relates to whether or not to conceal the allocation after randomisation (i.e. blinding of treatments), if this is possible.

Returning to the first of these components, a widely used alternative to simple randomisation (in which every participant has the same chance of receiving each of the treatment options, irrespective of the previous allocation in the trial) is the use of randomised permuted blocks. A simple example of these would be where the allocation scheme is arranged so that in every successive, say, eight patients, four would receive treatment A and four would receive B. Block lengths can be any multiple of the number of treatments. To avoid predictability of allocation, the block lengths can be varied randomly. In addition, stratification can be employed to ensure balance in respect of key prognostic factors. For example, within every eight male patients, four receive treatment A and four receive B, with a similar blocking arrangement for females. If the trial designer requires stratification over many factors, this approach becomes self-defeating, as the numbers become sparse in each stratum and the benefits of blocking are lost. This has led to an alternative system of allocation for ensuring balance in key prognostic factors, known as minimisation, which uses a dynamic system of allocation, so that at any stage in the trial the allocation of treatment to the next patient depends on the balance between the treatment

groups which will be achieved with each possible allocation.[10] This may or may not incorporate an element of 'random' allocation.

No empirical evidence has been identified that recommends any specific method but the educational articles express views similar to those found in Pocock's text-book on clinical trials.[11] For example, in describing the design of trials for treating angina, Ford[12] refers to the method of permuted blocks being a standard approach, with stratification to achieve balance across treatment groups with respect to key predictor variables. The advent of modern computer technology has made the more complex systems of allocation much easier to implement. Nevertheless, chance differences become less likely the larger the trial. Stratification and minimisation may therefore be particularly useful in small studies.

The extent to which various methods have been used is not clear from the literature. This arises from the poor quality of reporting of this aspect of RCTs (see chapter 7). The limited evidence which is available about blocking or stratification is variable. Talley[13] reported that stratification was not used in any of 16 trials identified in a systematic review of trials for the treatment of *Helicobacter pylori* positive functional dyspepsia. Altman and Doré,[14] in reviewing 80 RCTs published in major medical journals, found that in 32 trials in which the type of randomisation was specified, 31 employed stratification and 23 used blocking.

Although there is little information on the type of randomisation scheme used, there is evidence that the numbers in the arms of trials are more similar than would be expected without blocking.[14,15] This may be because some form of blocking has been used but not described, or because other selection processes have been involved, or because alternation rather than true random allocation has been applied.

The physical form of the randomisation process is important if entry bias is to be avoided. Bias may occur if the allocation of the next person to enter the trial is known in advance, as it may affect the decision to enter a participant into the trial. It generates exactly the same problem as systematic methods of allocation of treatments, as described by Chalmers and colleagues.[16] There is empirical evidence for the importance of concealment of allocation. In a systematic review of 250 controlled trials extracted from 33 meta-analyses, Schulz and colleagues found that trials in which the allocation sequence was inadequately concealed gave odds ratios of treatment effects which were exaggerated, on average, by 30–40%.[17]

The minimum level of adequate concealment of the randomisation is often taken to be a sequential set of sealed opaque envelopes but it has been reported that this system is prone to being circumvented unless supervised by an independent 'third party' (see chapter 5). More secure methods are the use of pre-packaged blinded treatments, where this is feasible, or central telephone- or computer-based randomisation. These methods are described more fully in Pocock's book.[11] Ford[12] is among those who advocate telephone randomisation, not only for ensuring secure randomisation but also for reducing the number of patients who violate the entry criteria specified in the protocol.

It appears to be universally accepted that the timing of the randomisation should ideally take place as close as possible to the time of starting the randomised treatment. This minimises the chances of the participants' circumstances changing so as to make the treatment allocated inappropriate after formal trial entry. Sometimes, this will be at the end of a run-in phase (see page 18); this minimises the risk of patient withdrawal after formal trial entry.

While the fourth element, of maintaining blinding of treatment, where feasible, does have potential implications for the process of randomisation (e.g. allocation to a numbered treatment pack), it is considered in this report under the conduct of the trial (see chapter 5).

Randomisation is usually at the patient level in most RCTs but, as Balas and colleagues[18] point out, in some areas of health technology assessment randomisation may be at a different level (e.g. at general practice level). Such studies raise additional problems of trial design, with sample size determination becoming more complicated, and they also result in the need for more sophisticated methods of analysis. The topic of 'cluster' RCTs is important but is outside the remit of this review and is not considered further.

Recommendations

1. There should always be a clear protocol for the preparation of (a) the sequence generation and (b) the concealment up to irrevocable trial entry of the randomisation; the operation of the system should not include any staff involved in determining entry of patients to the trial. (Basis: logical argument and anecdotal evidence.)

2. A telephone- or computer-based randomisation scheme provides secure treatment allocation and allows systematic checking of entry criteria. (Basis: logical argument and one major systematic review.)

Recommendation for research

- Further systematic reviews are required to investigate any effect of the method of randomisation on treatment effect sizes.

Choice of patient and clinical population

The only randomised studies of trial methodology identified that related to the selection of patients for inclusion in an RCT concerned the question of obtaining informed consent; this topic is considered separately in chapter 4. Empirical evidence on the representativeness of RCT study populations is available from several systematic reviews, case-studies and surveys, but the arguments presented for the viewpoints described below are based more firmly on logical reasoning than on empirical evidence.

Begg[19] neatly summarised the factors that influence which patients become trial participants as patient factors (inclusion and exclusion criteria), institutional factors (which centres participate) and physician preference, together with the related issue of patient consent. He found that published views on the advisability of wide or restrictive inclusion/exclusion criteria differed considerably. Arguments in favour of restrictive criteria are usually focused on the concept of forming homogeneous groups of patients and on possible gains in power if between-patient variation decreased.

An original observation by Begg,[19] which may partially explain restrictive exclusion criteria, is that the trial protocol can be seen as a document outlining state-of-the-art patient management. The protocol may thus specify the use of strict procedures at specified times but, by using departures from them as exclusion categories, the net effect is a reduction in the generalisability of the results of the trial. Wide eligibility criteria, which Begg prefers, are also claimed to increase power by increasing the number of available patients. Other reasons given by Yusuf and colleagues[20] are reduced costs per patient enrolled, the greater applicability of the trial and the greater likelihood of detecting subgroup effects. These authors found wide eligibility criteria generally preferable, although with exceptions: for example, clear evidence of benefit from treatment in some subgroups; expensive or toxic treatments or hazardous investigations which might only be justified in high-risk groups.

Collins and colleagues[21] also advocated wide eligibility criteria which, they argued, should be based simply on the 'uncertainty principle' that "the fundamental eligibility criterion is that both patient and doctor should be substantially uncertain about the appropriateness of each of the trial treatments for the particular patient". Adopting this principle, they stated that the degree of informed consent for such a trial should not differ greatly from that which is applied outside trials. They saw this as giving an approximate parallel between good science and good ethics.

The evidence that restrictive entry criteria can have a profound impact on recruitment to RCTs is overwhelming but, on logical grounds, unsurprising. Case-studies in the treatment of burns,[22] left ventricular dysfunction,[23,24] the effect of hormones on cardiovascular disease,[25] stroke trials,[26] lymphoma[27] benign prostatic hypertrophy[28] and schizophrenia[29] provide examples.

Despite the importance of the entry criteria for an RCT, there are wide variations between trials in the same area and many published reports do not describe the inclusion/exclusion criteria. Thus, Rosetti and colleagues,[30] in a systematic review of 102 RCTs in glaucoma, commented that patient selection and eligibility differed widely across studies. Begg,[31] reviewing current trials of adjuvant therapy for breast cancer in the USA found a wide variety of entry criteria, with a median of 23 exclusion categories, the rationale for which were often unclear. In a systematic review of therapeutic trials in *H. pylori*-positive functional dyspepsia, Talley[13] found that most trials had vague symptom definitions for entry and that sufficient details of ineligible patients were not presented. Marsoni and colleagues[32] found that of the 37 trials identified in advanced ovarian cancer, six gave an inadequate description of patient characteristics and only four summarised patients who were eligible but not randomised; three of these gave the reasons for non-randomisation. In 50 published trials of acute otitis media, 40 gave 'lucid inclusion criteria' and, in 43, exclusions were clearly stated.[33] Of 90 RCTs in a general practice journal, Silagy and Jewell[34] found no mention of how the study population had been obtained in 22%. Garcia-

Cases and colleagues[35] found that 26 of 50 clinical trial protocols reported selection criteria. Gurwitz and colleagues[36] focused on the exclusion of females and the elderly from trials of acute myocardial infarction. They found that ten of the 214 RCTs reviewed excluded women and a further 16 excluded women of child-bearing age. The effect of having a maximum age in 60% of RCTs was found to be the exclusion of a higher proportion of women, because of the age structure of the population. Thus there were 23% women in trials with no age limit, compared with 18% when an age limit was set.

Trial participants may differ from general patient populations in other ways. For example, DeLuca and colleagues[37] found from a survey of cardiovascular trial patients that the better-educated were more likely to consent to participate. The use of a run-in to exclude non-adhering patients will also tend to exclude less well-educated and Afro-American patients.[23] From a combination of RCTs and registries, Maynard and colleagues[38] estimated that women were 25% less likely to be included in trials of thrombolytic therapy. Csernansky and colleagues[39] observed that, in the USA, the practicalities of conducting trials in psychiatric conditions can cause restriction of patients. Studies in Department of Veteran Affairs hospitals will be based on men only, with adult onset of disease. Ansseau[40] highlighted the differences between participants in depression trials in the USA, where participants may be solicited through advertisements and receive payments, and those in Europe, who will be attending a doctor. European patients are usually more severely ill and in a higher social class; hence, results of trials in the USA and Europe are sometimes contradictory. Some of these issues are discussed in more detail elsewhere.[41]

In cancer trials, concern has been expressed at the low proportion of patients entered into RCTs. Hunter and colleagues[42] surveyed 44,156 newly diagnosed patients in the USA and found that 17,773 were initially eligible for a National Cancer Institute trial protocol. Of these 56% were found to be clinically eligible and 19% were treated in an RCT. The percentage on a trial protocol was found to decrease with age. The main reasons for non-entry were found to be physician refusal and patient refusal (see chapter 4).

The eventual inclusion of patients into RCTs is, of course, determined by the choice of clinicians taking part in the trial. Traditionally, RCTs were more likely to have participants from teaching hospitals rather than from district general hospitals or the community, with an inevitable bias in the representativeness of the patient population. Attempts to widen the research base by approaches such as the Community Clinical Oncology Program in the USA are claimed to have wider benefits than more speedy and more representative recruitment into RCTs. It was also envisaged that there would be an increase in information dissemination and diffusion of protocol-like management to patients not on protocol.[42]

In a survey of investigators in the Eastern Co-operative Oncology Group (ECOG), Taylor and colleagues[43] found enormous variation between respondents in the numbers they entered into RCTs. They concluded that "ECOG physicians appear to impose routinely additional inclusion/exclusion criteria based on their personal assessment of each eligible patient". In the year following the survey, only 38% of 1001 physicians entered any patients into one of the group's trial, although 91% expected to do so, and the total entry was only one-sixth of the expected entry; however, physicians who were more research-orientated showed higher accrual.

In stroke, competing trial protocols were seen as a reason for non-entry into trials by Saver.[44] He proposed randomisation for choice of trial to which the patient would be asked to consent. Sacco[45] suggested that stroke data banks can be used to determine the likely throughput to trials with specified inclusion/exclusion criteria, as well as using the data banks to assess the representativeness of trial populations.

Recommendations

1. Permissive inclusion/exclusion criteria are generally preferred, to allow both faster trial accrual and a more representative trial population. Exceptions where there may be an advantage from using a more homogeneous population are with expensive or toxic treatments or hazardous investigations which might only be justified in high-risk groups. (Basis: strength of the logical argument presented for this approach.)
2. The clinicians entering patients into multicentre trials should be chosen to give representative patient populations, subject to their having the relevant skills and resources to administer the trial treatments and procedures, and having an adequate throughput of appropriate patients. (Basis: logical argument.)

Relevance of the research question

One aspect of the relevance of an RCT concerns the definition of an appropriate patient population, as discussed above. Another relates to the end-points of an RCT, which is covered in the later section on trial outcomes. Consideration is given here to whether RCTs are being designed with appropriate treatment groups and relevant comparator groups. First, the question of ways in which the dose of a drug can be regulated in an RCT is considered. In doing so, it is assumed that this is not a Phase II RCT to determine appropriate dosing but a Phase III evaluative trial.

The usual approach would be to use a predetermined dose of the drug in one treatment arm and compare the results with those from placebo or active treatment arm(s). The potential problem with this design is that "some percentage of patients will not achieve the effect presumed to be a marker of the drug's mechanism of action".[46] This may be overcome by using what has been termed a 'post-dosed design'. This allows for modification of the dose on the basis of serum levels or response to treatment. As such, the design reflects the reality of post-trial administration of treatment. Friedman and Schron[46] reported that this design has been used successfully in many hypertension trials but, in a pilot study for the Cardiac Arrhythmia Suppression Trial, it was not adopted for the main study. Instead a 'pre-dosed design' was used. This involves a run-in phase which is unblinded and in which doses are modified to give a final dose which is tolerated and produces a response, in this trial the suppression of arrhythmias. Non-responders and those with adverse effects are then withdrawn from the trial prior to randomisation to the successful dose of the drug or to matching placebo. The authors[46] noted a potential difficulty with this design is that it does not provide a direct assessment of the treatment effect during the titration phase.

A design of essentially similar nature was described by Greenhouse and Meyer[47] as the classic design for a maintenance therapy trial. After an initial acute phase on one specified treatment, the responders are randomised either to the treatment used in the acute phase or to a comparator, which may be another active treatment or placebo, for the maintenance phase. Greenhouse and Meyer illustrated the bias that the classical maintenance therapy design has in favour of the run-in drug against the comparator. They advocate randomisation prior to the titration phase to both the acute treatment and the maintenance treatment, with evaluation of the response to the combined acute and maintenance policy. Those failing to stabilise in the acute phase would be regarded as failures for the combined policy. They argued convincingly that such a design would determine an optimal clinical strategy and would allow for conclusions to be drawn to a more general population. It has been noted that the absence of bias from the 'pre-dosed design' is dependent on the strong assumption that non-responders to the trial drug would also be non-responders to placebo; an assumption which appears reasonable but which may not be true.

Another closely related design, advocated by Klein[48] for use in psychotropic drug trials, is the drug discontinuation design. In this design, the first phase using the drug of interest may or may not involve titrating the dose. In the second phase, as with the 'pre-dose design', responders are randomised to receive drug or placebo, the purpose being to observe whether there is a withdrawal syndrome associated with the drug.

The other important design element involving treatment is the comparator to be used. The use of a placebo treatment arm in drugs trials allows determination of whether the drug is active in the condition tested, while achieving double-blindness, which is an important safeguard against bias. However, active comparators may be preferred, either for ethical reasons, so that treatment is not withheld or to answer questions of clinical relevance. In some RCTs with a pragmatic objective, an untreated control arm may be preferable despite potential bias due to lack of blinding. The appropriate choice for any RCT will depend on the clinical area and the objectives of the trial. While noting the importance of the topic, this is not pursued further in this review because of the lack of any evidence of a methodological nature.

The choice of treatments to be evaluated in an RCT and the comparator groups may also be affected by the use of concomitant medication. The treatments will often be intended to be used singly and concomitant use of treatments likely to interact with trial medication may be an exclusion criterion. However, in trials where continuing active medication is necessary, a comparison of interest may be between standard therapy plus trial drug versus standard therapy plus placebo. This is described by Wilensky[49] as an 'add-on design', and is equally applicable to parallel group and crossover studies. This may correspond to future practice, in which case the design is to be recommended, but, as a means of assessing

whether the trial drug is effective, this conclusion may be affected by possible interactions with the standard therapy.

Recommendation

- The only general recommendation that can be made in this area is that the choice of trial treatments should take into account the findings from previous trials, together with a critical assessment of whether the choice will allow the main objectives of the trial to be achieved. (Basis: logical argument.)

The type of trial design

The most fundamental technical aspect of trial design is the choice between a parallel group or a crossover design. In many clinical areas there is no possibility of considering a crossover design. This will be the case for acute conditions and when the treatment may be a disease modifier. The latter will cause an extreme form of carry-over, in which the outcome in later stages of the crossover is modified by treatments given in earlier stages. It is concern about carry-over which often causes a parallel group design to be preferred. The potential advantage of a crossover design is its efficiency. As within-participant comparisons are made, fewer participants are required for the same power. In this review, only one systematic review of treatment was found that allowed detailed examination of the use of crossover designs in relation to other features of the trial. Other systematic reviews of treatment have reported the percentage of trials which used a crossover design but the details provided have not allowed crossover trials and parallel group trials to be compared. Thus, in a systematic review of the treatment of glaucoma, 33 of 102 RCTs analysed used a crossover design.[50] In a systematic review of 90 RCTs, which appeared in the *British Journal of General Practice*, 21% used a crossover design.[34] The corresponding literature from the USA found that 11% of trials were crossover in design.[51] None of these reviews commented on the appropriateness of the use of the crossover. In a systematic review of 16 therapeutic trials in *H. pylori*-positive functional dyspepsia, a crossover design was used in one. The use was inappropriate because eradication of *H. pylori* in an active treatment arm meant that a carry-over effect was inevitable.[13] In the evaluation of antiepileptic drugs, Wilensky[49] stated, without referencing the evidence, that crossover trials require only a quarter of the number of participants needed for a similarly powerful parallel study but, because of the danger of carry-over effects, most studies now use the parallel design. The most informative systematic review was by Klein,[52] who produced a detailed table on 43 randomised double-blind, placebo-controlled treatment trials for irritable bowel syndrome. There were 21 crossover trials in the review. The median number of patients in the crossover trials was 26 compared with 63.5 in parallel group trials, with the median durations of treatment being 4 weeks and 8 weeks, respectively. Klein believed that 8 weeks would be the minimum treatment length to be clinically relevant. Only one (5%) crossover trial met this requirement compared with 12 (55%) of the parallel group studies. The mean percentage of patients evaluable was 89% in both types of design. The outcome of the trials, in terms of the statistical significance of a global efficacy measure, was almost identical for both classes of design: 9/19 for crossover; 10/19 for parallel groups. Interestingly, the statistical methods in the published reports were assessed as appropriate for 14 of 18 parallel group studies (78%) compared with 8 of 19 (42%) of the cross-over trials (the smaller numbers result from some studies being unevaluable).

The most extreme example of a crossover design is the '*n*-of-one' design where a single patient receives two or more treatments over several treatment periods. As Simon[53] commented, in a major review of statistical methodology for clinical trials, inferences are then specific to that patient. He commented that, as for the crossover design, the validity of results will depend on assumptions which often cannot be adequately tested by the data. Within the class of crossover designs there are many alternative designs available as well as the *n*-of-one design; however, this is beyond the scope of the present review.

Decisions on parallel or crossover designs are directed at the number and duration of treatment periods.

Another design element compatible with both of these designs is the factorial design. This allows the simultaneous investigation of the effect of two or more treatment factors, by the way the allocated treatments are arranged. The simplest factorial design is a 2×2 design, in which, if treatments A and B are to be compared against control, four treatment groups (A; B; A + B; control) would be formed. Such designs are efficient and offer the prospect of answering two therapeutic questions for the price of one. However, as Simon[53] commented, this assumes a lack of interaction between the treatments. The situations which

he identified as making the factorial design attractive are: if the end-points for evaluating the two factors are distinct and there is little chance of an interaction; if the magnitude of the interaction can be assumed to be small relative to the size of the main effects; or if the likelihood of both main effects being clinically significant is small.

It had been the subjective view of the authors that the use of factorial designs is uncommon, which appears to be substantiated by this review. Only one of the systematic reviews identified reported the use of factorial designs. This was in the review of RCTs published in the *British Journal of General Practice*, referred to earlier,[34] of which 4% used a factorial design.

A further category of design which can be employed in conjunction with the designs considered earlier is the sequential design. In the non-sequential design, a pre-determined number of participants are recruited and analysis of data takes place only at the completion of the trial when all patients have been recruited and there is adequate follow-up. In contrast, the sequential design uses interim analyses of the data collected to date to decide whether or not recruitment to the trial should continue. This design is considered more fully in chapter 5, under 'early stopping of trials', but a brief introduction is given here.

The essence of a sequential design is that conventional analyses conducted at several stages during the progress of a trial would give a falsely inflated chance of a 'statistically significant' treatment difference, because of a form of multiple testing. Instead, each of the interim analyses may be thought of as 'spending' some of the (conventionally) 5% of the probability available. Alternative types of sequential design differ in the way the design 'spends' this probability over the repeated tests. The potential advantage of the sequential design is that it may be possible to conclude that the treatments differ significantly from each other at an earlier point in time than with a fixed sample size design. This could have important cost and ethical implications, as in principle the better treatment could universally be applied sooner. There is, however, a danger with early stopping of a trial that the results may be statistically significant but not convincing enough to ensure changes in treatment policy. This occurred in a trial reported by Pignon and colleagues.[54] Two treatments for small-cell lung cancer were compared in the trial and, at the first planned interim analysis, a statistically significant difference was found. The data monitoring committee (DMC) recommended continuation, however, because of fears of imbalance in the treatment groups, short follow-up and some unexpected findings. The second interim analysis was judged to confirm the difference and trial entry was stopped. After a further delay, during which additional deaths occurred, a manuscript was prepared but rejected by two well-known journals. Their major concerns were that the difference in treatments was unlikely to induce such a large difference in survival and that the follow-up period was short. After a further year of follow-up, which reinforced the findings, the paper was accepted for publication.

The authors draw a number of conclusions from this case study, principally directed against the possibility of very early stopping. Some sequential designs have a relatively high chance of early stopping but stopping rules, such as the O'Brien and Fleming design,[55] require stronger evidence in the early phases of the trial in order for the trial to be stopped. Other approaches which are discussed include delaying the first analysis until a minimum number of events and/or when a minimum follow-up was reached. They considered that meeting a statistical early stopping rule should be only one of several factors to be considered in a decision whether to stop a trial. One such factor is the impact of early stopping on the likely wide acceptability of the trial results.

A similar attitude to the use of stopping rules is presented by Fleming and DeMets,[56] who urged flexibility in implementing sequential designs.

The above types of clinical trial design are by no means exhaustive but represent the designs which are most common. Other designs, which will be met elsewhere in this review, are the randomised consent design (see chapter 4) and the comprehensive cohort design (see chapter 6).

Recommendation

- The parallel group design, with fixed sample size and fixed treatment schedules, remains the most commonly used RCT design. Inexperienced trialists should not, however, automatically adopt this design, as other designs can sometimes produce major benefits. The views of experienced trialists should be sought whenever possible. (Basis: logical reasoning.)

Sample size

There is wide acceptance in the literature that RCTs should ideally be sufficiently large for there to be a high probability of obtaining a statistically significant difference between treatment groups if there is a clinically significant difference in their effects. Indeed, a systematic review on the ethics of clinical trials has recently reported that 'underpowered' trials are generally regarded as unethical,[57] although the authors did not share this view. The four elements from which sample size determinations arise are: the level of statistical significance which is sought (α-level); the risk of obtaining a non-significant result when a specified difference exists (β-level); the difference which it is important for the trial to be able to detect as statistically significant; and the expected variability of the response variable.

The convention is very strong that an α-level of 0.05 should be set when designing the trial. This level is rarely commented upon in systematic reviews. An exception is the systematic review of trials of artificial surfactant in neonatal respiratory distress by Raju and colleagues.[58] Of the 21 RCTs included, 16 provided sample size calculations. The only trial not to use the 5% α-level used 1%.

There is very strong evidence that the β-level or equivalent power (1–β) have not been considered in the planning of the majority of RCTs. In chapter 7 (see *Table 25*), the median percentage of trials with adequate reporting of sample size consideration was 9% in the systematic reviews considered. This does not necessarily imply that only 9% made appropriate sample size calculations, as in some instances it may only be the reporting which is deficient. However, most reviews also commented that the actual sample sizes used were, in most cases, inadequate to give sufficient power to detect clinically relevant differences. In a frequently cited review from 1978, 71 'negative' trials, with binary outcomes, in major medical journals were considered.[59] *Post-hoc* β-levels were calculated for these trials in relation to the detection of a relative improvement of 25% and 50%, using one-tailed tests of significance at the 5% level. The median values for β were 0.74 for a 25% improvement and 0.40 for a 50% improvement. These compare with generally recommended values for β of 0.2 or 0.1, when using two-tailed tests. A level of 80% power (β = 0.2) was only reached for 7% of trials in relation to a 25% improvement, and 31% in relation to a 50% improvement. Moher and colleagues[60] also looked at the *post-hoc* power of negative trials (using two-tailed tests) to detect 25% or 50% relative improvements. They examined three major medical journals for 1975, 1980, 1985 and 1990. Overall, 16% had 80% power to detect a 25% improvement and 36% had 80% power to detect a 50% improvement. They did not find any consistent improvement over time. Talley,[13] in a systematic review of RCTs in the treatment of *H. pylori* functional dyspepsia, commented that for 80% power to detect a clinically significant difference, 100 patients per group were required. Only one of 16 trials exceeded 100 patients in total. Commenting in a systematic review of controlled treatment trials in irritable bowel syndrome, Klein[52] estimated that about 650 patients would be needed for 80% power to detect a clinically significant difference. Of the 43 trials, 35% recruited less than 30 patients and only 7% exceeded 100. Vandekerckhove and colleagues[61] considered that 700 patients per group were necessary to detect plausible differences in success rates for most treatments of infertility problems but found that the average size in 501 RCTs was 96, with a range from 5 to 933. Only 18 of these papers discussed power, 12 gave details of their sample size calculation and only six reached that number. The most common convention was to set the power at 80%, although some authors have argued that equal values for α and β would be preferable.[62]

The third element in the sample size determination is the difference which is specified. Raju and colleagues[58] view this factor "as a bridge between statistical procedures and clinical decision making. For this reason, there are no specific guidelines regarding the choice of this factor, nor can any be expected." In Raju and colleagues' systematic review, the median expected treatment effect used in sample size calculations was a 50% change, which is a large effect. Peto and colleagues[63] argued that most available treatments have only moderate treatment effects but that these are still worthwhile. Although the authors generally agreed with the comments of Raju and colleagues and Peto and colleagues, Torgerson and colleagues[64] demonstrated with case studies how the economic cost of treatments can be used to influence the difference which would be of clinical relevance. Naylor and colleagues[65] discussed the use of a patient perspective in setting this choice difference but this appears to be a theoretical approach which lacks practical application. In public health terms, the difference should also depend on how common the disease in question is and the frequency of poor outcomes. With a common disease and a frequently poor outcome, even a small relative risk reduction will produce a major

absolute benefit. Decision-analytical techniques may be helpful in deciding the size of difference which is clinically significant.

The fourth element in sample size determination is the variability of the observations. For outcome variables that are continuous, this may be estimated from previous similar studies (with systematic reviews being preferred when available) or a pilot study may be required to provide such an estimate. With binary response data, estimates of the probabilities of 'success' provide the required information. Outcomes which are times-to-events-occurring are commonly referred to as survival data, whether or not it is time to death being considered. Sample sizes are then often calculated on the basis of, say, 5-year survival rates (i.e. a binary outcome). However, as Fayers and Machin[66] pointed out, the power is effectively decided by the number of deaths at the time of analysis. They provided a nomogram, based on analysis using the log-rank test, which links the significance level, the power, the hazard ratio and the number of deaths. The number of deaths itself needs to be estimated from the rate of input to the trial, the duration of input, the death rates and the timing of the analysis.

Fayers and Machin[66] claimed that the use of a nomogram for determining sample size is to be preferred to the use of tables. It forces one to be aware of the inherent imprecision of the numerical estimates. They also recommended a sensitivity analysis, particularly of the difference that it is desired to detect, in order to explore the impact of varying assumptions.

Nearly all RCTs are designed with equal numbers of patients per treatment arm. This is usually satisfactory, although there is little impact on power if the randomisation ratio is 3:2. Medical Research Council trials have been designed with a 2:1 ratio.[66] This design will be attractive if:

(a) access to one trial treatment is limited
(b) there is a wish to accumulate experience with a new treatment more rapidly
(c) there is a wish to estimate an effect, such as an uncommon adverse effect, more precisely in one group
(d) there is concern about withholding a new therapy from half the participants.

Imbalance in sample sizes arising for reasons other than design has been noted by Altman and Doré.[14] In their systematic review of RCTs in which blocking was not used in randomisation they noted that 26 trials had more patients in the control groups while only ten trials had more in the experimental group. They attribute the difference to differential withdrawal rates.

Sample size considerations can often lead to the need for extremely large trials, sometimes referred to as mega-trials. Peto and colleagues[63,67] gave examples to support their view that important results in the treatment of vascular and neoplastic disease could only be established by large-scale randomised evidence. This arises because treatment effects that are only moderate can be extremely important, especially in common conditions. They further argued that such large trials must be kept simple if they are to be feasible.

At the other extreme, there is debate on the value of small trials. Some authors[68,69] have argued that studies with a too small sample size may be scientifically useless and hence unethical. Others have given examples showing that the small clinical trial can have value.[70] Fayers and Machin[66] took a compromise view, which the authors consider to have merit and which is similar to views expressed by Edwards and colleagues.[57] They accepted that any trial is better than no trial provided that publications make clear that the power was low and that the results can, at best, be regarded as hypothesis forming, and that the trial, like all trials, should be registered before it is commenced, so that, even if unpublished, the results will be available for use in an overview or meta-analysis.

Other factors can also affect the planned sample size for an RCT. It was noted earlier in this chapter that the numbers available for evaluation can often be less than the numbers randomised and reasonable allowance should be made for this. It has also been seen, in the section on patient populations, that a homogeneous population may be recommended to increase power, although there are strong contrary arguments. If adherence to treatment is expected to be a problem, sample size requirements can be inflated several-fold.[71,72] The use of a run-in period to mitigate this effect is described below.

Recommendations

The views of Fayers and Machin,[66] described earlier, are endorsed enthusiastically; the conclusions of their paper are summarised below in a modified form.

1. Sample size calculations should consider a sensitivity analysis, and should give 'ballpark' estimates rather than unrealistically precise numbers.

2. When small trials are necessary, they should be reported as hypothesis forming.
3. Clinical trials should be pre-registered, to allow unpublished results to be traced.
4. Full details of sample size calculations should always be reported.
5. Funding bodies, independent protocol review committees and journals should all demand provision of sample size considerations.

(Basis: logical argument.)

Use of a run-in period

Our search strategy revealed three papers which dealt specifically with the design issue of whether a run-in period should be used.[23,73,74] A run-in period in an RCT is one in which, prior to the use of the randomised treatments, the patient receives a non-randomised medication (often placebo) and is usually subject to most of the other trial procedures which pertain in the main phase of the trial. An important potential advantage of a run-in period, which is the main focus of each of the papers cited, is the identification of adherers and non-adherers to treatment. After a run-in period it would be common practice to undertake a returned pill count and then only those with a pre-determined compliance rate (e.g. greater than 80% of the prescribed number of pills having been used) would enter the randomised phase of the trial. Davis and colleagues[23] presented empirical evidence from a large pilot study which was part of the Cholesterol Reduction in Seniors Program (or CRISP). After two screening visits, a placebo run-in lasting 3 weeks was commenced at the third visit. Poor adherers were those with a pill count of less than 80%, or who failed to return their unused pills. Irrespective of the adherence category, participants were randomised to one of two doses of lovastatin or to placebo, and results at 3 months and 6 months were reported. In the 85% of participants who were good adherers at baseline, 89% and 81% were good adherers at 3 and 6 months, respectively. The rates were lower in poor baseline adherers, at 71% and 64%, respectively. The overall mean adherence rates (percentage of pills taken) at 3 months were reduced from 90.9% in good baseline adherers to 89.3% by the inclusion of all participants. Corresponding 6-month figures were 85.5% and 83.4%. In terms of the treatment effect, the difference in decrease in low-density lipoprotein cholesterol levels between the high-dose group and placebo was 59.2 mg/dl in good baseline adherers and 56.3 mg/dl in the total group. The authors concluded that a placebo run-in to exclude poor adherers would not have been useful in the main study.

Several points of criticism can be levelled at this analysis. The most important is that the impact of the run-in period should not be assessed just on the size of the resulting treatment effect but also on the size of its standard error, and this is not given. The study also failed to report the number of participants attending each of the first four visits. Another potential advantage of a run-in period or even repeated screening visits is that non-attenders can be excluded prior to randomisation, when they will not influence estimates of treatment effects.

Lang[73] presented two examples of the use of a run-in. The Physicians' Health Study was a randomised, placebo-controlled, double-blind trial of aspirin in the reduction of cardiovascular mortality and beta-carotene in decreasing cancer incidence. This case study illustrated the importance of the choice of using active treatments or placebo during the run-in. Active aspirin was used so that the participants with aspirin intolerance could be identified prior to randomisation and it was postulated that its effect on cardiovascular disease would be short-lasting. In contrast, beta-carotene was postulated to have cumulative effects, so a placebo beta-carotene was used during the run-in. Lang argued that the choice of active or placebo agents will depend on the expected incidence of side-effects which would stop treatment and considerations of likely carry-over effects from the run-in to the post-randomisation period. She also drew attention to the need for the placebo to be a good one if a run-in is used, as participants may notice the change, with consequent loss of treatment 'blindness'. The importance of the run-in period in the Physicians' Health Study, which used mail-based questionnaires, was evidenced by non-response at the end of the run-in from 8000 of the 33,000 original participants, with a further 3000 respondents unwilling to participate further.

The second case study presented was the Coronary Primary Prevention Trial to test the effect of long-term lowering of cholesterol on the risk of coronary heart disease. This study was criticised by Lang[73] for failing to use a 3-month pre-randomisation phase, during which there were five contacts with the participants, to test some elements of compliance.

Schechtman and Gordon[74] took a modelling approach to determine when a run-in strategy will be effective. A key construct in their approach was categorisation of participants into compliant

participants, zero compliers who are treatment intolerant or comply so badly as to obtain no treatment benefit, and partial compliers who would normally be excluded by the run-in but would have a partial response to treatment. They found that the presence of many zero compliers means that a run-in is associated not only with a reduction in the number of randomised participants to achieve a specified level of power but also with a reduction in the total number of participants who enter the run-in phase. If, however, there are few zero compliers and the size of the treatment effect in partial-compliers is at least 60% of the size of the treatment effect in compliant participants, then the conditions under which a run-in produces sample size benefits are limited.

Although the emphasis in these papers was on the effect of compliance on the value of a run-in period, there is empirical evidence that this is not the only reason in practice for using a run-in period. In a brief report in her paper, Lang[73] also described a small systematic review she had undertaken using MEDLINE on papers published between January and June 1988 with the keyword 'run-in'. Within the 26 trials identified, the central goal was to establish a population with stable disease and, although all trials involved pill-taking during the run-in, only two reported that compliance was measured. The length of run-ins in these trials ranged from 1 week to 4 weeks.

Subsequent to the period of the systematic review, Senn has questioned whether placebo run-ins are justified.[75] He is concerned on ethical grounds that deceit is often used in informing the patient about this phase of the trial. On statistical grounds he noted the disagreement on the efficiency of a placebo run-in period and observed that in any case an active treatment would serve as well as a placebo. Commenting on Senn's paper, Ramsay disagreed about the ethical issue, provided that the informed consent is appropriate.[76] As a form of words he suggested: "During this study there will be one or more periods during which you will have inactive (placebo) treatment. It is important for the success of the study that you are unaware which study periods these are". He also argued against active treatment during the run-in, except when needed for ethical reasons, because of variation in the withdrawal effect. Ramsay saw the major advantage of the run-in in familiarising all concerned (including patients) with the procedures that are to be used in the study.

Thus the use of a run-in is a complicated decision and Davis and colleagues[23] commented that "the literature on the evaluation of a placebo run-in is sparse".

Recommendations

1. A run-in period before the post-randomisation phase of an RCT should always be given consideration. Run-in periods have logistic and resource implications and address more explanatory than pragmatic considerations, and these issues should be taken into account. (Basis: judgement of the authors.)
2. The use of a run-in period to exclude non-adherers is particularly likely to be helpful when an appreciable proportion of participants are expected to be treatment intolerant or to fail to comply well enough to achieve appreciable treatment benefit. (Basis of recommendation: theoretical considerations and limited case studies.)
3. A run-in period may more often be of value to ensure stability of disease in participants rather than to detect non-adherence. (Basis: one small systematic review.)
4. When run-ins are employed to ensure stability, compliance should be assessed to provide data for possible enhancement of treatment estimates using compliance rate as a baseline covariate. (Basis: judgement of the authors.)

Recommendations for research

1. There is need for a systematic review of the use of run-in periods to assess the benefit of this design aspect of RCTs over a range of clinical conditions and various modes of intervention.
2. There should be research, both theoretical and empirical, directed at methods of using information on compliance during the run-in phase (and potentially during the post-randomisation phase) to improve estimates of treatment effect.

Trial outcomes

The selection of appropriate measures to assess the effect of treatment is clearly of paramount importance. As well as being an issue in the design of RCTs, the question of multiple end-points is a serious issue in the analysis of RCTs, and this topic will be referred to later in chapter 6.

General discussion of this subject is handicapped by the vastly different sets of variables which are appropriate in varying clinical conditions and when addressing different types of research question. This review has not revealed any papers which have identified and evaluated

methodological issues concerning the number and selection of outcome variables. Most of the relevant papers have been educational, based on the author(s)' experience and knowledge of a specialist area, with limited case studies or systematic reviews of practice.

The features of the outcome variables regarded by the authors as being of major importance differ across clinical areas. In situations where death is possible during an RCT, hard, easily defined end-points such as mortality are often recommended. Thus, in RCTs of severe sepsis and septic shock, Balk and Bone[77] recommended mortality as the primary end-point. Organ failure was also important but difficult to define. They recommended using definitions that have been verified in previous trials and are generally well accepted. In other clinical areas, there may be debate as to whether to use mortality or cause-specific mortality as the primary outcome measure. This is debated by Friedman and Schron[46] for the design of anti-arrhythmic drugs trials. Accurate classification of cause of death is a problem and could be subject to bias, and total mortality may be preferred. In other situations, such as trials of screening for breast cancer, the effect being sought would be lost in all-cause mortality; hence, cause-specific mortality is essential.[78]

In trials in which mortality will not be the major end-point, if an end-point at all, many authors recognised the need for standard definitions of outcome. Thus, in a systematic review of 141 trials of single agent chemotherapy in small-cell lung cancer, Grant and colleagues[79] stated that "few reports defined their response criteria precisely or stated that WHO response criteria were used; several gave no information about the criteria used". In considering malignant gliomas, Fine[80] also commented that many studies do not even describe the definition of response being used in a given trial. He commented on the resulting difficulty of comparing treatment regimens from one study to another. Bigby and Gadenne[81] called for clear definitions of outcomes which have clinical and biological significance. Rosetti and colleagues[50] carried out a systematic review of 102 published RCTs on the medical treatment of primary open-angle glaucoma. They found that only 31 (30%) used clinically relevant outcomes, while 70% used intra-ocular pressure lowering effect as a surrogate for visual function. They noted that none had established a link between intra-ocular pressure and visual damage. Bigby and Gadenne[81] also criticised the use of surrogate end-points, referring to earlier studies on the use of anti-arrhythmic drugs after myocardial infarction; they produced benefit in the substitute end-point of the occurrence of abnormal ventricular depolarisation but, in RCTs against placebo, their use was associated with an excess of mortality. Fleming and DeMets[82] gave several convincing examples of the dangers of surrogate end-points. Examples favouring surrogate end-points are presented in the section below on timescales . The issue is considered again in appendix 2, where summary data from two Cochrane Collaboration trial registers are presented.

Papers in the area of rheumatology have expressed unease with outcomes which have been used to date. Felson and colleagues[83] criticised the use of outcome measures, such as radiographic findings to assess the healing of bone erosions, which are insensitive to change. They also reported that nearly all rheumatoid arthritis trials use numerous outcome measures. Commenting that this strategy relied more on the quantity of measures than the quality, they considered that the design of such trials could be improved by reducing the quantity of the measures employed. They also regretted the lack of standardised outcome measures in this area but accepted that its realisation may be problematical in ensuring that standardisation is truly uniform. Ratain and Hochberg[84] also commented that most trials in rheumatoid arthritis select too many, rather than too few, end-points. Blair and Silman[85] stated that funding and ethical approval for any future trials in rheumatoid arthritis should require the adoption of a central core of sensitive measures. They also made the potentially important point for trial design that patient familiarity with the proposed outcome assessment can affect the response. They cited a randomised trial showing that preliminary experience with a treadmill had a direct effect on reducing the magnitude of cardiorespiratory response.[86]

In contrast to the above calls for a reduced number of outcome measures, Klein,[48] in dealing with psychotropic trials, recommended that standard measures should be supplemented by new tailored measures, which "must be shown to be valid and speak directly to the drug benefit". In trials of irritable bowel syndrome, Maxton and colleagues[87] argued for more than the standard measures, as non-colonic symptoms are common. They observed that, in the trial situation, data recorded on diary cards probably reflect real changes better than the data recorded at patient visits, while acknowledging the limitations of what can be recorded on a diary card. Publishing in a psychiatric journal, Leon and colleagues[88] stated that outcome measures should

be selected on the basis of their psychometric properties and urged the development and dissemination of psychometrically sound scales.

Recommendation

- Outcome measures should ideally be objective, valid, reliable, sensitive to change and clinically relevant. In clinical areas with 'standard' outcome measures, these should be included so that the results from different RCTs can be combined meaningfully. (Basis: logical argument.)

Recommendations for research

1. A consensus on core variables which should regularly be recorded needs to be established in individual clinical areas, to allow combined analyses of trials in the area.
2. Systematic reviews are required in additional areas to document trials using surrogates, unreliable or invalid outcome measures.

Timescale

There are many elements of the design of an RCT where timing can be an important consideration. This review yielded little hard evidence related to this but the experience of trialists, often presented via case studies, can provide some guidance.

Easterbrook and Matthews[89] reviewed the fate of research protocols approved by Central Oxford Research Ethics Committee for the period 1984–87, of which 45% were clinical trials. Of these studies, 21% were never started and a further 13% of clinical trials were abandoned. Pilot studies were conducted in 15% of initiated studies and 12% of these studies were subsequently abandoned, although there is no separate information on studies which were clinical trials. The authors concluded that "one clear message for future investigators is the importance of a thorough preliminary exploration of the pragmatic as well as the scientific aspects of a planned research project". They went on to argue that "many problems cannot be anticipated and only come to light during the course of a study, hence the value of a preliminary pilot or feasibility study".

The duration of the treatment phase of an RCT is critical, both in terms of the trial's capacity to demonstrate statistically significant differences and in terms of how the trial can relate to clinical practice. Clearly a trial which is too short for a particular treatment's mode of action will have little chance of success. Nevertheless, there is evidence from reviews in particular areas of application that many trials have too short a treatment period. Thus, Klein,[52] in a systematic review of randomised, double-blind, placebo-controlled trials of the treatment of irritable bowel syndrome dating back to 1966, found that 30 of 43 trials had a treatment period of 6 weeks or less, when he considers that 8–12 weeks would be the minimum to be truly clinically relevant. Klein also found a similar problem in psychotropic drug trials.[48] He considered that 4–6 weeks are needed for clear-cut distinction in outpatient antidepressant trials but stated that many trials last for only 3 weeks, making detection of drug benefit unlikely. As the drugs would be used in clinical practice for prolonged periods, he considered that, even in Phase II studies, data on maintenance effects should be obtained. If side-effects or relapse rates were sufficiently bad, expensive Phase III studies could be avoided and drugs with only short-term benefits would not be marketed. Awad[90] was also concerned that clinical trials of new neuroleptics have an average treatment duration of 4–6 weeks in acutely psychotic patients who need long-term medication. Although he saw these trials as a necessary first step, they do not tell us about long-term use.

Kerr[91] highlighted the difficulty of finding a balance in terms of study duration in cancer pain trials. In common with the previous authors, he agreed that "too short a period may not be adequate to demonstrate the efficacy or safety of the medi-cation or delivery system tested". In contrast, he described a crossover trial with 2-week treatment periods where only 21% of 699 patients who entered the trial achieved the crossover. In order to limit dropout in this class of RCTs, he suggested studying patients earlier in the natural history of their disease. Nixon[92] commented that cancer prevention trials usually need to be of much longer duration than cancer treatment trials and identified ways in which they can be shortened. One is to use intermediate surrogate end-points. A trial involving patients with adenomatous polyps is described which uses the size and number of recurrent polyps as a surrogate for colon cancer. This will reduce the duration of the trial from 10 to 5 years. Another strategy is to seek more appropriate participants for the prevention trial and he cited an evaluation of fat reduction for prevention of breast cancer, by recruiting patients already treated for the disease. This is justified by the supposed mechanism of action and the RCT should be completed in 5 years with 2000 patients, compared with 10 years and 16,000 patients for a primary prevention trial in high-risk participants.

In many of the above examples, the periods of treatment and follow-up would be the same. While this is often the case in prevention trials, or in RCTs of a chronic disease, it is less likely to be so with acute conditions. In these situations, decisions on the timing of starting treatment relative to an acute event may also be important. It is often important to start treatment within a few hours of the event, which can cause logistical problems that may require extreme efforts to overcome. In such trials the treatment period may be short but a long follow-up is often needed. This carries the danger noted by Nixon[92] that "over time, an original hypothesis may become less appealing or pertinent as new strategies emerge".

Intimately linked with treatment and follow-up duration is the timing of assessments and their number. At least one pre-treatment assessment will usually be desirable, and Frison and Pocock[93] gave examples of the benefit of multiple pre-treatment assessments. Post-treatment assessments will depend on the expected action of the treatments and the inconvenience/hazard of the assessments, and little useful general guidance can be given. As well as giving thought to the ethical implications of multiple assessments, which may involve some risk, albeit slight, or put extra burden on participants thus increasing drop-out, the statistical analysis needs consideration at the design stage. The problems which arise from the multiple testing that can be generated by multiple assessments will be discussed in chapter 6. For now, a need for a balance between the desire for additional information, the importance of not overburdening participants and the analytical problems which this may create is noted.

The total duration of an RCT will depend not just on the length of treatment and follow-up but often more crucially on the time necessary to recruit the required number of patients. This is an area on which many trials can founder and is considered in detail in chapter 5.

Recommendation

Most of the issues relating to timescale will be specific to the clinical area in which an RCT is performed; hence, the recommendation is limited to one which is general.

- Duration of treatment and follow-up should be sufficiently long to identify effects on clinically important outcomes. (Basis: systematic reviews indicating that treatment periods are commonly too short.)

Practical aspects of trial design

This area has appreciable overlap with the conduct of the trial (chapter 5). Fundamental to an RCT is the construction of the study protocol. In modern practice, this is essential, if only to gain ethical approval. Melink and Whitacre[94] tabulated the following components of a protocol which they considered essential:

- objectives of the study
- scientific background and study rationale
- patient selection criteria
- treatment information
- study design
- treatment plan
- adverse effect evaluation criteria
- dose adjustment plan
- required study parameters for patient monitoring
- response evaluation criteria
- statistical considerations
- criteria for study termination
- special comparison studies
- data submission requirement and forms
- references
- consent form.

These authors stressed that several cancer cooperative groups in the USA, who are in the process of developing the protocol, consulted widely on the clarity and feasibility of the protocol and on the responsibilities of nurses and other study investigators.

Pharmaceutical companies and major trialists will have a standard structure to the protocol. For inexperienced trialists expert assistance is required and an interesting attempt to provide this is described by Wyatt and colleagues.[95] They reported the development of a knowledge-based critiquing system for the authors of clinical trial protocols. Preliminary evaluation showed that passable draft protocols could be completed much more quickly but the authors were disappointed by the difficulties which those using the system experienced. Such approaches may be useful for the inexperienced, in view of the scarcity of the expertise which is required, but these approaches require considerably more development.

Allied to the development of the trial protocol is the determination of the resources to support it. An infrastructure is needed comprising experienced field workers, data managers, statisticians, a randomisation service, and experts in other areas such as health economics. Evidence and recommendations concerning the infrastructure will be developed in chapter 5.

References

1. Zelen M. Randomized consent designs for clinical trials: an update. *Stat Med* 1990;**9**:645–56.
2. Altman DG, Whitehead J, Parmar MKB, Stenning SP, Fayers PM, Machin D. Randomised consent designs in cancer clinical trials. *Eur J Cancer* 1995;**31A**:1934–44.
3. Bradley C. Designing medical and educational intervention studies: a review of some alternatives to conventional randomized controlled trials. *Diabetes Care* 1993;**16**:509–18.
4. Brewin CR, Bradley C. Patient preferences and randomised clinical trials. *BMJ* 1989;**299**:313–15.
5. Chang RW, Falconer J, Stulberg SD, Arnold WJ, Dyer AR. Prerandomization: an alternative to classic randomization. The effects on recruitment in a controlled trial of arthroscopy for osteoarthrosis of the knee. *J Bone Joint Surg Am* 1990;**72**:1451–5.
6. Grunkemeier GL, Starr A. Alternatives to randomization in surgical studies. *J Heart Valve Dis* 1992;**1**:142–51.
7. Hansson L, Hedner T, Dahlof B. Prospective randomized open blinded end-point (PROBE) study. A novel design for intervention trials. *Blood Press* 1992;**1**:113–19.
8. Henshaw RC, Naji SA, Russell IT, Templeton AA. Comparison of medical abortion with surgical vacuum aspiration: women's preferences and acceptability of treatment. *BMJ* 1993;**307**:714–17.
9. Kopelman L. Consent and randomized clinical trials: are there moral or design problems? *J Med Philos* 1986;**11**:317–45.
10. Taves RR. Minimization: a new method of assigning patients to treatment and control groups. *Clin Pharmacol Ther* 1974;**15**:443–53.
11. Pocock SJ. Clinical trials – a practical approach. Chichester: John Wiley; 1983.
12. Ford I. Angina clinical trial methodology [review]. *Postgrad Med J* 1991;**67** (suppl 3):S2–8.
13. Talley NJ. A critique of therapeutic trials in Helicobacter pylori-positive functional dyspepsia [review]. *Gastroenterology* 1994;**106**:1174–83.
14. Altman DG, Doré CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990;**335**: 149–53.
15. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994;**272**:125–8.
16. Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983;**309**:1358–61.
17. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.
18. Balas EA, Austin SM, Ewigman BG, Brown GD, Mitchell JA. Methods of randomized controlled clinical trials in health services research. *Med Care* 1995;**33**:687–99.
19. Begg CB. Selection of patients for clinical trials. *Semin Oncol* 1988;**15**:434–40.
20. Yusuf S, Held P, Teo KK, Toretsky ER. Selection of patients for randomized controlled trials: implications of wide or narrow eligibility criteria. *Stat Med* 1990;**9**:73–86.
21. Collins R, Peto R, Gray R, Parish S. Large-scale randomized evidence: trials and overviews. In: Wetherall DJ, Ledingham JGG, Warrell DA, editors. Oxford textbook of medicine, vol. 1. 3rd ed. Oxford; Oxford University Press; 1996. p.21–32.
22. Coombs DW, Dimick A, Bronstein JM, Potts LH, Bowens B. Conceptual and methodologic problems in the evaluation of a new burn treatment modality. *J Burn Care Rehabil* 1993;**14**:568–71.
23. Davis CE, Applegate WB, Gordon DJ, Curtis RC, McCormick M. An empirical evaluation of the placebo run-in. *Control Clin Trials* 1995;**16**:41–50.
24. Carew BD, Ahn SA, Boichot HD, Dierenfeldt BJ, Dolan NA, Edens TR, *et al.* Recruitment strategies in the studies of left ventricular dysfunction (SOLVD): strategies for screening and enrolment in two concurrent but separate trials. *Control Clin Trials* 1992;**13**:325–38.
25. Johnson S, Mebane-Sims I, Hogan PE, Stoy DB. Recruitment of postmenopausal women in the PEPI Trial (Postmenopausal estrogen/progestin interventions). *Control Clin Trials* 1995;**16** (4 suppl):20–35S.
26. LaRue LJ, Alter M, Traven ND, Sterman AB, Sobel E, Kleiner J. Acute stroke therapy trials: problems in patient accrual. *Stroke* 1988;**19**:950–4.
27. Stone JM, Page FJ, Laidlaw CR, Cooper I. Selection of patients for randomised trials: a study based on the MACOP-B vs CHOP in NHL study. *Aust N Z J Med* 1994;**24**:536–40.
28. Rosenbaum TP, Shaw PJ. Too stringent patient selection criteria in a clinical trial of alpha-blockers reduce the value of the results. *Eur Urol* 1991;**20**:175–8.
29. Bowen J, Hirsch S. Recruitment rates and factors affecting recruitment for a clinical trial of a putative anti-psychotic agent in the treatment of acute schizophrenia. *Hum Psychopharmacol* 1992;**7**:337–41.

30. Rossetti L, Marchetti I, Orzalesi N, Scorpiglione N, Torri V, Liberati A. Randomized clinical trials on medical treatment of glaucoma. Are they appropriate to guide clinical practice? *Arch Ophthalmol* 1993;**111**:96–103.

31. Begg CB, Engstrom PF. Eligibility and extrapolation in cancer clinical trials. *J Clin Oncol* 1987;**5**:962–8.

32. Marsoni S, Torri W, Taiana A, Gambino A, Grilli R, Liati P, et al. Critical review of the quality and development of randomized clinical trials (RCTs) and their influence on the treatment of advanced epithelial ovarian cancer. *Ann Oncol* 1990;**1**:343–50.

33. Claessen JQPJ, Appelman CLM, Touw-Otten FWMM, de Melker RA, Hordijk GJ. A review of clinical trials regarding treatment of acute otitis media. *Clin Otolaryngol* 1992;**17**:251–7.

34. Silagy CA, Jewell D. Review of 39 years of randomized controlled trials in the *British Journal of General Practice* [review]. *Br J Gen Pract* 1994;**44**:359–63.

35. Garcia-Cases C, Duque A, Borja J, Izquierdo I, de la Fuente V, Torrent J, *et al.* Evaluation of the methodological quality of clinical trial protocols. A preliminary experience in Spain. *Eur J Clin Pharmacol* 1993;**44**:401–2.

36. Gurwitz JH, Col NF, Avorn J. The exclusion of the elderly and women from clinical trials in acute myocardial infarction. *JAMA* 1992;**268**:1417–22.

37. DeLuca SA, Korcuska LA, Oberstar BH, Rosenthal ML, Welsh PA, Topol EJ. Are we promoting true informed consent in cardiovascular clinical trials? *J Cardiovasc Nurs* 1995;**9**:54–61.

38. Maynard C, Selker HP, Beshansky JR, Griffith JL, Schmid CH, Califf, RM, *et al.* The exclusion of women from clinical trials of thrombolytic therapy: implications for developing the thrombolytic predictive instrument database. *Med Decis Making* 1995;**15**:38–43.

39. Csernansky JG, Newcomer JW, Miller LS, Faustman WO. Clinical factors that may confound the assessment of drug efficacy [review]. *Psychopharmacol Bull* 1991;**27**:231–5.

40. Ansseau M. The Atlantic gap: clinical trials in Europe and the United States. *Biol Psychiat* 1992;**31**:109–11.

41. Edwards SJL, Lilford RJ, Braunholtz DA, Jackson JC, Hewison J, Thornton J. Ethical issues in the design and conduct of randomised controlled trials. *Health Technol Assess* 1998;**2**(15).

42. Hunter CP, Frelick RW, Feldman AR, Bavier AR, Dunlap WH, Ford L, *et al.* Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log. *Cancer Treat Rep* 1987;**71**:559–65.

43. Taylor KM, Feldstein ML, Skeel RT, Pandya KJ, Ng P, Carbone PP. Fundamental dilemmas of the randomized clinical trial process: results of a survey of the 1737 Eastern Cooperative Oncology Group investigators. *J Clin Oncol* 1994;**12**:1796–805.

44. Saver JL. Coping with an embarrassment of riches. How stroke centers may participate in multiple, concurrent clinical stroke trials. *Stroke* 1995; **26**:1289–92.

45. Sacco RL. Interactions between stroke data banks and clinical trials. *Neuroepidemiology* 1994;**13**:275–82.

46. Friedman LM, Schron EB. Statistical problems in the design of antiarrhythmic drug trials. *J Cardiovasc Pharmacol* 1992;**20**(suppl 2):S114–18.

47. Greenhouse JB, Meyer MM. A note on randomization and selection bias in maintenance therapy clinical trials [review]. *Psychopharmacol Bull* 1991;**27**:225–9.

48. Klein DF. Improvement of Phase III psychotropic drug trials by intensive Phase II work [review]. *Neuropsychopharmacology* 1991;**4**:251–71.

49. Wilensky AJ. Protocol design [review]. *Epilepsy Res Suppl* 1993;**10**:107–13.

50. Rossetti L, Marchetti I, Orzalesi N, Scorpiglione N, Liberati A. Is proper methodology associated with the use of a clinically relevant outcome measure? The case of randomized clinical trials on medical treatment of open-angle glaucoma. *Online J Current Clin Trials* [serial online] 1993 Nov 11; Doc No 100: [6512 words; 199 paragraphs].

51. Silagy CA, Jewell D, Mant D. An analysis of randomized controlled trials published in the US family medicine literature, 1987–1991 [see comments]. *J Family Pract* 1994;**39**:236–42.

52. Klein KB. Controlled treatment trials in the irritable bowel syndrome: a critique [review]. *Gastroenterology* 1988;**95**:232–41.

53. Simon R. A decade of progress in statistical methodology for clinical trials [review]. *Stat Med* 1991;**10**:1789–817.

52. Pignon JP, Tarayre M, Auquier A, Arriagada R, le Chevalier T, Ruffie P, *et al.* Triangular test and randomized trials: practical problems in a small cell lung cancer trial. *Stat Med* 1994;**13**:1415–21.

54. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;**35**:549–56.

55. Fleming TR, DeMets DL. Monitoring of clinical trials: issues and recommendations [review]. *Control Clin Trials* 1993;**14**:183–97.

56. Edwards SJL, Lilford RJ, Braunholtz D, Jackson J. Why "underpowered" trials are not necessarily unethical. *Lancet* 1997;**350**:804–7.

58. Raju TN, Langenberg P, Sen A, Aldana O. How much 'better' is good enough? The magnitude of treatment effect in clinical trials. *Am J Dis Child* 1992;**146**:407–11.

59. Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 'negative' trials. *N Engl J Med* 1978;**299**:690–4.

60. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994;**272**:122–4.

61. Vandekerckhove P, O'Donovan PA, Lilford RJ, Harada TW. Infertility treatment: from cookery to science. The epidemiology of randomised controlled trials [review]. *Br J Obstet Gynaecol* 1993;**100**:1005–36.

62. Lilford RJ, Johnson N. The alpha and beta errors in randomised trials [letter]. *N Engl J Med* 1990; **322**:780–1.

63. Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials [review]. *J Clin Epidemiol* 1995;**48**:23–40.

64. Torgerson DJ, Ryan M, Ratcliffe J. Economics in sample size determination for clinical trials. *QJM* 1995;**88**:517–21.

65. Naylor CD, Llewellyn-Thomas HA. Can there be a more patient-centred approach to determining clinically important effect sizes for randomized treatment trials? *J Clin Epidemiol* 1994;**47**:787–95.

66. Fayers PM, Machin D. Sample size: how many patients are necessary [editorial review]? *Br J Cancer* 1995;**72**:1–9.

67. Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials [review]. *Ann N Y Acad Sci* 1993;**703**:314–40.

68. Newell DJ. Type II errors and ethics. *BMJ* 1978;**ii**:1789.

69. Altman DG. Statistics and ethics in medical research. III. How large a sample? *BMJ* 1980; **281**:1336–8.

70. Powell-Tuck J, MacRae KD, Healy MJR, Lennard-Jones JE, Parkins RA. A defence of the small clinical trial: evaluation of three gastroenterological studies. *BMJ* 1986;**292**:599–602.

71. Melnikow J, Kiefe C. Patient compliance and medical research: issues in methodology. *J Gen Intern Med* 1994;**9**:96–105.

72. Schechtman KB, Gordon MO. The effect of poor compliance and treatment side effects on sample size requirements in randomized clinical trials. *J Biopharm Stat* 1994;**4**:223–32.

73. Lang JM. The use of a run-in to enhance compliance. *Stat Med* 1990;**9**:87–5.

74. Schechtman KB, Gordon ME. A comprehensive algorithm for determining whether a run-in strategy will be a cost-effective design modification in a randomized clinical trial [review]. *Stat Med* 1993;**12**:111–28.

75. Senn S. Are placebo run-ins justified? *BMJ* 1997;**314**:1191–3.

76. Ramsay L. Commentary: placebo run-ins have some value. *BMJ* 1997;**314**:1193.

77. Balk RA, Bone RC. Methodological issues in the design and implementation of controlled clinical trials in severe sepsis and septic shock. Discussion about the trial protocol of supplemental immunoglobulin therapy in sepsis and septic shock. *Theor Surg* 1994;**9**:32–7.

78. Alexander FE, Anderson TJ, Brown HK, Forrest APM, Hepburn W, Kirkpatrick AE, *et al.* The Edinburgh randomised trial of breast cancer screening: results after 10 years of follow-up. *Br J Cancer* 1994;**70**:542–8.

79. Grant SC, Gralla RJ, Kris MG, Orazem J, Kitsis EA. Single-agent chemotherapy trials in small-cell lung cancer, 1970 to 1990: the case for studies in previously treated patients. *J Clin Oncol* 1992;**10**:484–98.

80. Fine HA. The basis for current treatment recommendations for malignant gliomas [review]. *J Neurooncol* 1994;**20**:111–20.

81. Bigby M, Gadenne AS. Understanding and evaluating clinical trials. *J Am Acad Dermatol* 1996;**34**:555–94.

82. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;**125**:605–13.

83. Felson DT, Anderson JJ, Meenan RF. Time for changes in the design, analysis, and reporting of rheumatoid arthritis clinical trials [see comments] [review]. *Arthritis Rheum* 1990;**33**:140–9.

84. Ratain JS, Hochberg MC. Clinical trials. A guide to understanding methodology and interpreting results [review]. *Arthritis Rheum* 1990;**33**:131–9.

85. Blair PS, Silman AJ. Can clinical trials in rheumatology be improved [review]? *Curr Opin Rheumatol* 1991;**3**:272–9.

86. Bartz J, Purves S. Improving the validity of clinical exercise testing: the relationship between practice and performance. *Arch Phys Med Rehabil* 1989;**70**:599–604.

87. Maxton DG, Morris J, Whorwell PJ. Improving clinical trials in irritable bowel syndrome: some practical aspects. *Eur J Gastroenterol Hepatol* 1992;**4**:337–41.

88. Leon AC, Marzuk PM, Portera L. More reliable outcome measures can reduce sample size requirements. *Arch Gen Psychiat* 1995;**52**:867–71.

89. Easterbrook PJ, Matthews DR. Fate of research studies. *J R Soc Med* 1992;**85**:71–6.

90. Awad AG. Methodological and design issues in clinical trials of new neuroleptics: an overview. *Br J Psychiatry* 1993;**163**:51–7.

91. Kerr IG. Clinical trials to study pain in patients with advanced cancer: practical difficulties. *Anticancer Drugs* 1995;**6**:18–28.

92. Nixon DW. Special aspects of cancer prevention trials [review]. *Cancer* 1994;**74**(9 suppl):2683–6.

93. Frison L, Pocock SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design [see comments]. *Stat Med* 1992;**11**:1685–704.

94. Melink TJ, Whitacre MY. Planning and implementing clinical trials. *Semin Oncol Nurs* 1991;**7**:243–51.

95. Wyatt JC, Altman DG, Heathfield HA, Pantin CF. Development of design-a-trial, a knowledge-based critiquing system for authors of clinical trial protocols. *Comput Methods Programs Biomed* 1994;**43**:283–91.

Chapter 4

Barriers to participation in clinical trials

Methods

Clinician participation or patient participation was identified by 265 papers in the main database as being an important issue (scoring 3 for relevance). From this subset of the main database, all primary research papers were selected which reported findings relating to recruitment of clinicians or patients to clinical trials ($n = 84$); only papers reporting new data (quantitative or qualitative) on aspects of trial conduct or design were included in the analysis for this chapter. Papers relating solely to Phase I or Phase II trials were excluded, as were papers commenting on supposed barriers without supporting evidence.

Each of the papers was reviewed by two readers (SR and one other) and summarised using subject areas determined by the study group prior to the review. The subject areas of interest were:

- barriers to clinician participation
- barriers to patient participation
- difficulties around the clinician/participant interface
- difficulties specific to trials among certain types of people
- difficulties related to specific types of comparison
- barriers to participation within protocols
- factors related to the healthcare system
- barriers to RCTs (as opposed to other study designs)
- other factors beyond the control of investigators.

A brief description of the study design and relevant findings were abstracted from each paper and entered into the Microsoft Access database. Data were checked for accuracy of abstraction.

Having identified specific issues within the subject areas described above, the data extracted from the research papers were transformed into matrices or tables to assist in critical analysis.[1] Once this sorting was complete, each of the tables was reviewed to produce an interpretation which was consistent with the material. These interpretations were subjected to a process of critical, iterative analysis,[2] in which an interpretation was evaluated against the data and discarded or modified until a consistent interpretation was reached.[3]

Following the iterative analytical process, two main categories of barrier emerged from the data:

- barriers to clinician participation
- barriers to patient participation.

For the purpose of this review, the term 'clinician' includes all clinical staff, for example, nurses, physiotherapists, physicians, radiotherapists and surgeons. Concerns about additional demands of RCTs and the consent process were major themes within both of these categories.

The tables resulting from the analysis are presented in a truncated form, with interpretation of the findings, and recommendations for overcoming barriers to recruitment based on the findings.

To access findings from research before 1985, systematic reviews of the literature were also identified. Non-systematic reviews were excluded as these were found to be of variable quality.[4] Over the period of study, the concept of systematic reviewing was not prevalent and only five papers describing formal systematic reviews were found. These related to the inclusion of women in clinical trials of antihypertensive medications,[5] the recruitment of older people to clinical trials of arthritis,[6] recruitment of ethnic or racial minorities,[7] the problems associated with RCTs in surgery,[8] and a general review of recruitment experience.[9] Because so few systematic reviews were identified, the main findings of these papers are presented individually.

Barriers to clinician participation

The success of RCTs depends on the participation of clinicians. To achieve this, clinicians must agree to participate when invited, must recruit patients who are eligible (including offering participation to eligible patients) and must follow the trial protocol. Each of these stages represents a potential barrier to participation which varies from trial to trial. For the purpose of this review, the differing levels of participation have been considered as

one, so that barriers to participation include barriers to absolute protocol adherence as well as barriers to taking part in particular trials.

Time constraints

Eight papers (seven surveys of clinical trial participants) reported that time was a barrier to clinician participation (*Table 3*). Two aspects were identified: firstly, the time pressures from usual clinical practice and management duties precluded their commitment to clinical trials;[10–13] secondly, the time demands of recruitment and follow-up in clinical trials were felt to be a barrier.[12,14–17] Two 1994 surveys of NHS consultants involved in research indicated that the pressures had worsened following recent NHS reforms.[11,13]

Recommendations

1. Participation in clinical trials should be encouraged as a component of the core activity of clinicians. (Basis: judgement of the authors.)

TABLE 3 *Time constraints*

Study (country)	Title	Subject area	Setting	Method	Results
Aaronson, *et al.,* 1996[10] (The Netherlands)	Telephone-based nursing intervention improves the effectiveness of the informed consent process in cancer clinical trials	Cancer	Hospital	Survey of participating clinicians to investigate reasons for not recruiting patients	Lack of time was physicians' most common reason for not referring patients (44/103).
Benson, *et al.,* 1991[14] (USA)	Oncologists' reluctance to accrue patients onto clinical trials: an Illinois Cancer Center study	Cancer	Hospital	Survey of 224 clinicians who participated in research	73% stated that excessive time needed for patient follow-up. Community physicians in particular reported that excessive time required for patients on trials.
Dickinson, 1994[11] (UK)	Clinical research in the NHS today	Not applicable	Medical Research Society	Survey of 294 clinicians who participated in research	Major problems encountered in undertaking research included diminished research time for NHS-funded clinicians to pursue clinical research (83%), pressure put on clinicians to undertake management roles (73%), diminished calibre and numbers of keen researchers due to reduced career opportunities (61%).
Fisher, *et al.,* 1991[15] (USA)	Clinical trials in cancer therapy: efforts to improve patient enrolment by community oncologists	Cancer	Community	Survey of 75 clinicians who participated in studies	Problems of undertaking research included time demands on self and staff (35%), and obtaining on-study tests (14%).
Foley & Moertel, 1991[12] (USA)	Improving accrual into cancer clinical trials	Cancer	Hospital	Survey of 209 participating clinicians	28% considered that too many tests were impediment to recruitment, 27% that complex studies were a barrier or required too much physician time (21%). 23% were concerned that physician was too busy.
Smyth, *et al.,* 1994[13] (UK)	Conducting clinical research in the new NHS: the model of cancer. UK Coordinating Committee on Cancer Research	Cancer	Hospital	Survey of 287 participating clinicians	Physicians under new trust regime find little time to talk to patients. Main disincentive for potential trialists was lack of time.
Taylor, *et al.,* 1984[17] (USA)	Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer	Cancer	Hospital	Survey of 91 participating clinicians	Additional time to follow patients and to explain procedures mentioned.
Taylor, 1985[16] (USA)	The doctor's dilemma: physician participation in randomized clinical trials	Cancer	Hospital	Survey of 91 participating clinicians	9% of those who did not enrol all patients argued that time needed to explain to patients, rigid rules governing eligibility, inflexibility of prescribed treatment formulas, were insurmountable barriers to accrual.

2. Clinicians should be asked to do the minimum required from them, with other activities such as follow-up, performed by specifically funded and employed staff. (Basis: logical argument.)

Staffing and training

RCTs are commonly run in everyday clinical settings, often without additional support. Problems associated with lack of trained staff were highlighted in 12 studies: eight surveys of participating clinicians, one survey of patients, one survey of centres who refused to take part in a trial, one report of a failed trial and one description of trial recruitment at two sites (*Table 4*). Clinicians participating in clinical trials were often ill-prepared for such a role.[11,18–21] Lack of research experience[11,22,23] and training[24] were found to be barriers to patient recruitment. Lack of available support staff, for example, clinical trial nurses, was also blamed for poor recruitment.[12,13,24–26] A stable research team was preferable in conducting RCTs.[27]

Recommendation

- Clinicians need preparation and appropriate support staff if they are to participate in RCTs. (Basis: anecdotal evidence.)

Rewards and recognition

The impact of rewards for clinicians taking part in RCTs is difficult to identify. It is unclear whether lack of rewards acts as a deterrent to participation. Five papers (four studies: three surveys of clinicians participating in trials and one report of a failed trial) mentioned rewards as incentives for participating in trials (*Table 5*). These included benefit to the individual's reputation and that of their institution,[20,28] and economic incentives.[12,21] One study suggested that personal encouragement and support are required to achieve successful participation.[19]

Recommendations

1. Clinicians should be rewarded appropriately and adequately for taking part in RCTs. The rewards need not be financial but should include positive feedback and support. (Basis: anecdotal evidence.)
2. Contribution should be credited, such as in all publications, during career progression. (Basis: judgement of the authors.)

Recommendation for research

- Research to identify the most appropriate form of reward for clinicians is needed.

Impact on the doctor–patient relationship

The possibility that RCTs may adversely affect the doctor–patient relationship (*Table 6*) and that this may act as a barrier to clinician recruitment was emphasised in 12 papers (11 studies: seven surveys of participating clinicians, two reports of failing trials, one survey of non-participating centres and one description of an RCT of different information for patients). The main issues highlighted were clinicians' difficulties in admitting that they did not know which treatment was best[14,16,17,29] and the perceived conflict between their roles as clinicians and researchers.[18,26,28,30] As a result, clinicians have considered their rapport with patients to have been damaged if they entered patients into a trial.[16,17,21,29,31–33]

Recommendation

- RCTs should be framed and organised in ways which minimise differences between research and clinical practice, and address questions of sufficient importance for clinicians to be comfortable with the need for the research role. (Basis: logical argument.)

Recommendation for research

- Research is needed to identify trial designs which interfere least with the clinician–patient relationship.

Concern for patients

Concern for the patient was raised in nine papers (eight surveys of participating clinicians and one description of accrual to cancer trial) (*Table 7*). Included were concern about treatment toxicity or side-effects,[12,22] the burden of the trial for patients,[10,12,30] including travel distance[12] and cost,[14,15] lack of patient transport,[25] and a reluctance to recruit more severely ill patients.[10,34] Taylor and colleagues[17] mentioned the fear of feeling responsible if a patient did not receive the treatment which turned out to be best.

Recommendations

1. Trial design should seek to minimise the burden for patients and reassure clinicians about the demands on patients. (Basis: logical argument.)
2. The value of RCTs as a 'risk minimising' strategy, when there is uncertainty, should be emphasised to clinicians. (Basis: judgement of the authors.)

Perception of importance of RCT

Only two studies mentioned the importance of the research question. Foley and Moertel[12] found that scientifically uninteresting trials were felt to be an impediment to recruitment by 17% of

TABLE 4 *Staff and training*

Study (country)	Title	Subject area	Setting	Method	Results
Dickinson, 1994[11] (UK)	Clinical research in the NHS today	Not applicable	Medical Research Society	Survey of 294 clinicians who participated in research	Major problems: diminished research time for NHS-funded clinicians to pursue clinical research (83%); pressure put on clinicians to undertake management roles (73%); increasing clinical demands on university-funded medical staff (71%); unwillingness of management to cover minor research costs (63%); diminished calibre and numbers of keen researchers due to reduced career opportunities (61%).
Foley & Moertel, 1991[12] (USA)	Improving accrual into cancer clinical trials	Cancer	Hospital	Survey of 209 participating clinicians	Inadequate support from personnel identified by 22% as problem for RCTs. Inadequate healthcare coverage identified as impediment to patient entry by 31%. Interdisciplinary coordination was problem for 16%.
Henzlova, *et al.,* 1994[27] (USA)	Patient perception of a long-term clinical trial: experience using a close-out questionnaire in the Studies of Left Ventricular Dysfunction (SOLVD) trial	Cardio-vascular	Hospital	Survey of 3522 patients who participated in long-term heart failure trial	Stability of investigating team advantageous for patient satisfaction (7% of patients mentioned changes of staffing as a negative experience).
Morse, *et al.,* 1995[25] (USA)	Issues of recruitment, retention and compliance in community-based clinical trials with traditionally under-served populations	AIDS/HIV	Community	Telephone survey of 14 sites participating in trial	Structural barriers cited included: lack of financial resources (50%); lack of staff (31%); lack of space (39%); lack of medical staff (30%).
Penn, *et al.,* 1990[26] (UK)	Reasons for declining participation in a prospective randomised trial to determine the optimum mode of delivery of the preterm breech	Obstetrics	Hospital	Survey of 11 centres who refused to participate in a trial	Extra work and staffing identified as major barriers to participation in RCTs: 5/11 centres who refused to participate considered that they had insufficient staff to obtain properly informed consent.
Shea, *et al.,* 1992[24] (USA)	Enrollment in clinical trials: institutional factors affecting enrollment in the Cardiac Arrhythmia Suppression Trial (CAST)	Cardio-vascular	Hospital	Survey of 112 participating sites	Higher 'rank of physician responsible for CAST' significantly associated with higher enrolment rate. Higher enrolment correlated with being affiliated with medical school, having staff training programmes, having higher proportion of eligible patients cared for by trained staff, having CAST staff available more days per week, having more nurse–clinicians available.
Taylor, *et al.,* 1987[18] (various)	Interpreting physician participation in random-ized clinical trials: the physician orientation profile	Cancer	Hospital and private practice	Survey of 484 participating clinicians	67% agreed with statement: 'When published data and my clinical judgement conflict, I am more likely to rely on personal clinical experience'.
Taylor, 1992[19] (various)	Physician participation in a randomized clinical trial for ocular melanoma	Cancer/ ophthal-mology	Private practice and hospital	Survey of 101 participating physicians	Physicians must be educated on the rationale of RCTs to improve participation. 'Academic concerns' given by 75% as disincentive for participating.
Taylor, *et al.,* 1994[20] (USA)	Fundamental dilemmas of the randomized clinical trial process: results of a survey of the 1737 Eastern Cooperative Oncology Group investigators	Cancer	Hospital and private practice	Survey of 1737 participating clinicians	Suggests that participants in RCTs are ill-prepared regarding the process of such investigations. Most frequent obstacle to successful completion of clinical trial is physician's reluctance to participate (55%).

continued

TABLE 4 contd *Staff and training*

Study (country)	Title	Subject area	Setting	Method	Results
Tognoni, *et al.,* 1995[21] (Italy)	Randomised clinical trials in general practice: lessons from a failure	Cardio-vascular	General practice	Report of a failed trial	Authors comment on problems with GPs' 'unsatisfactory attitude...[to] controlled research'.
Wadland, *et al.,* 1990[23] (USA)	Recruitment to a primary care trial on smoking cessation	Smoking cessation	Primary care	Study to describe and compare rates of recruit-ment in RCT of smoking cessation in two primary care practices	Comparison of recruitment by study personnel and receptionist: study personnel appeared to enhance recruitment and limit bias.
Winn *et al.,* 1984[22] (USA)	An evaluation of physician determinants in the referral of patients for cancer clinical trials in the community setting	Cancer	Hospital	Vignette study of 82 clinicians	Physicians with previous experience of RCTs had greater referral rate (77% vs. 39%). Physician less likely to refer patient when side-effects considered serious.

TABLE 5 *Rewards and recognition*

Study (country)	Title	Subject area	Setting	Method	Results
Foley & Moertel, 1991[12] (USA)	Improving accrual into cancer clinical trials	Cancer	Hospital	Survey of 209 participating clinicians	24% regarded excess costs as an entry impediment (varied between professionals). To improve protocol accruals, 71% agreed that non-essential research testing should be subsidised and 41% considered that there should be increased funding for physician participation time.
Taylor, 1992[19] (various)	Physician participation in a randomized clinical for ocular melanoma	Opthal-mology	Private practice and hospital	Survey of 101 participating physicians	Incentives and disincentives to participate must be addressed and explained to potential participants. Support required for investigators, continual feedback, personal encouragement, intensive education.
Taylor, 1992[28] (USA)	Integrating conflicting professional roles: physician participation in randomized clinical trials	Cancer	Hospital and private practice	Survey of 101 participating clinicians	Rewards for scientists are on a 'macro' level, i.e. status in professional community, rather than financial reward. Physicians expect both. Status conferred on physicians by patients, on scientists by peers.
Taylor, *et al.,* 1994[20] (USA)	Fundamental dilemmas of the randomized clinical trial process: results of a survey of the 1737 Eastern Cooperative Oncology Group investigators	Cancer	Hospital and private practice	Survey of 1737 participating clinicians	Main rewards are for clinical practice (67%) rather than contribution to scientific knowledge. Major reason for participation is that it benefits institution (54%); 97% considered that 'participation in RCTs is an asset to my reputation'.
Tognoni, *et al.,* 1991[21] (Italy)	Randomised clinical trials in general practice: lessons from a failure	Cardio-vascular	General practice	Report of a failed trial	Authors comment that clinician compliance may have been improved through economic incentives.

respondents in their survey of participating clinicians. Tognoni and colleagues,[21] in a report of a failed trial, stated that the questions to be tested must be of definite interest to participating clinicians.

Recommendation

- The questions addressed by RCTs should be of sufficient importance to clinicians for them to be willing to take part and comply with protocol requirements. (Basis: anecdotal evidence.)

TABLE 6 *Impact on the doctor–patient relationship*

Study (country)	Title	Subject area	Setting	Method	Results
Benson, *et al.*, 1991[14] (USA)	Oncologists' reluctance to accrue patients onto clinical trials: an Illinois Cancer Center study	Cancer	Hospital	Survey of 224 clinicians who participated in research	Physicians consider that research protocols undercut patients' beliefs in physicians' knowledge and decision-making power. Most physicians were not at ease when talking to patients about choice of treatment.
Chang, *et al.*, 1990[31] (USA)	Prerandomization: an alternative to classic randomization. The effects on recruitment in a controlled trial of arthroscopy for osteo-arthrosis of the knee	Ortho-paedic surgery	Hospital	Report of a trial which was apparently failing but which was a success following a change of design (to pre-randomisation)	Concern that participation in RCT would adversely affect patient–physician relationship.
Langley, *et al.*, 1987[32] (Canada)	Why are (or are not) patients given the option to enter clinical trials?	Cancer	Hospital and general practice	Survey of 87 participating and non-participating clinicians	Family physicians rated concern about scientific design as main reason for not offering patients entry to clinical trials, followed by rapport with patients.
Penn, *et al.*, 1990[26] (UK)	Reasons for declining participation in a prospective randomised trial to determine the optimum mode of delivery of the preterm breech	Obstetrics	Hospital	Survey of 11 centres who refused to participate in a trial	Paper suggests that clinicians problems over inferred consent may be a mask over the conflict the doctor finds himself in between physician and experimenter.
Simes, *et al.*, 1986[33] (Australia)	Randomised comparison of procedures for obtaining informed consent in clinical trials of treatment for cancer	Cancer	Hospital	RCT of two types of information provision for patients; discussion of impact on doctors	Suggests that when doctors are aware that trial will involve total disclosure they are reluctant to include patients. Major factor in reluctance of doctors to achieve fully informed consent for their patients is perceived impact on doctor–patient relationship.
Siminoff, *et al.*, 1989[30] (USA)	Doctor–patient communication about breast cancer adjuvant therapy	Cancer	Hospital	Observation and audio-recording of 100 patient–doctor interactions and surveys of participants	Physician has unavoidable ethical dilemma between duty of care to current patients, to the clinical trials themselves and to future patients who may benefit.
Taylor, *et al.*, 1984[17] (USA)	Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer	Cancer	Hospital	Survey of 91 participating clinicians	Of those who did not enter all patients on trial, 73% mentioned their relationship with patient, e.g. 15 considered that their rapport with patient would be jeopardised (major impact on enrolment in trials?). Difficulty in telling patients that they did not know which treatment was better was expressed by 15 (23%).
Taylor, 1985[16] (USA)	The doctor's dilemma: physician participation in randomized clinical trials	Cancer	Hospital	Survey of 91 participating clinicians	Of those who did not enrol all patients: 73% made some reference to their relationship with patient, principally that ultimate decision-making power was removed from either physician or patient; 23% expressed difficulty in telling patients that they did not know which procedure was better; 8% of those not enrolling all patients were frightened of their personal responsibility if one treatment proved to be superior.

continued

TABLE 6 contd *Impact on the doctor–patient relationship*

Study (country)	Title	Subject area	Setting	Method	Results
Taylor, *et al.*, 1987[29] (Canada)	Physician response to informed consent regulations for randomized clinical trials	Cancer	Hospital	Survey of 170 participating clinicians	81% considered that telling patients that physicians do not know which treatment is better may have negative effect on patients. Informed consent affected relationship with trial patients for 24%, and obtaining informed consent always highlighted dual roles as investigator and primary caregiver for 52% (plus 39% sometimes). 48% considered patient awareness of dual role helpful; 55% (plus 40% sometimes) said dual role always made them uncomfortable.
Taylor, *et al.*, 1987[18] (various)	Interpreting physician participation in randomized clinical trials: the physician orientation profile	Cancer	Hospital and private practice	Survey of 484 participating clinicians	76% wanted to be measured as a successful physician on basis of help to individual patients and 24% on basis of their research contributions; 67% agreed that 'when published data and my clinical judgement conflict, I am most likely to rely on personal clinical experience'.
Taylor, 1992[28] (USA)	Integrating conflicting professional roles: physician participation in randomized clinical trials	Cancer/ ophthal-mology	Hospital and private practice	Survey of 101 participating clinicians	Two potentially conflicting roles described, as physician and scientist: clinician's role to reduce patient uncertainty; research role to provide evidence. In life-threatening disease, research physicians may present uncertainty in a positive light.
Tognoni, *et al.*, 1991[21] (Italy)	Randomised clinical trials in general practice: lessons from a failure	Cardio-vascular	General practice	Report of a failed trial	Change of role for GP from confident and reassuring; seen as 'shifting their image'.

Loss of autonomy

Loss of clinical autonomy, including loss of decision-making power and independence, being accountable to a third party, and restriction of the ability to individualise patient care, was reported in seven papers (covering five surveys of clinician trial participants, all by the same first author using a similar survey methodology) as a reason for not recruiting all patients to clinical trials (*Table 8*).[16–20,28,29]

Recommendation

- Pragmatic RCTs are likely to be more acceptable to clinicians, since this type of design permits more clinician freedom. (Basis: judgement of the authors.)

Incompatibility of protocol with normal practice

The authors of eight papers (seven studies: six surveys of participating clinicians and one report of a failed trial) identified incompatibilities of the protocol with usual clinical practice as a barrier to recruitment of clinicians[26] or patients[15–17,21,30,35,36] (*Table 9*). Such incompatibilities included personal preference for one of the treatments,[17,30,36] unwillingness to recruit patients to a trial including a 'no treatment' arm,[35] lack of choice of treatment,[15,16] a requirement for patients to be withdrawn from maintenance treatment prior to recruitment[21] or an incompatibility with departmental policy, such that staff lacked the clinical skills to undertake both arms of the study.[26]

Recommendations

1. Clinicians are likely to find it easier to use pragmatic RCT protocols because they test interventions as used in 'everyday' care. (Basis: logical argument.)
2. Protocols should be designed to minimise incompatibility with normal practice. (Basis: anecdotal evidence.)
3. Clinicians should be recruited who understand and are in agreement with the proposed research protocol. (Basis: logical argument.)

Recommendation for research

- Research is required to identify and modify aspects of trial design which are not consistent with normal practice.

TABLE 7 *Concern for patients*

Study (country)	Title	Subject area	Setting	Method	Results
Aronson, *et al.*, 1991[10] (The Netherlands)	Telephone-based nursing intervention improves the effectiveness of the informed consent process in cancer clinical trials	Cancer	Hospital	Survey of participating clinicians to investigate reasons for not recruiting patients	Worry about burden of trial for patients (n = 6). Patients not randomised more likely to have advanced stage disease or been recruited to a Phase II trial.
Antman, *et al.*, 1985[34] (USA)	Selection bias in clinical trials	Cancer	Hospital	Accrual, protocol compliance and toxicity measured at 3-month intervals during RCT	Clinicians less likely to recruit patients with more advanced stages of disease, visceral primaries and large tumours, and older patients.
Benson, *et al.*, 1991[14] (USA)	Oncologists' reluctance to accrue patients onto clinical trials: an Illinois Cancer Center study	Cancer	Hospital	Survey of 224 clinicians who participate in research	Some 79% believed that research is costly to participants. A high percentage of oncologists feel uncomfortable when patients are on RCTs.
Fisher, *et al.*, 1991[15] (USA)	Clinical trials in cancer therapy: efforts to improve patient enrolment by community oncologists	Cancer	Community	Survey of 75 clinicians who participate in studies	Costs to patients were identified by 22% as a barrier to enrolling patients.
Foley & Moertel, 1991[12] (USA)	Improving accrual into cancer clinical trials	Cancer	Hospital	Survey of 209 participating clinicians	Concern about: too much personal burden on patients (25%); too much toxicity (19%); extra procedures (28%); excessive travel distance (29%); inadequate healthcare coverage (31%); were identified as impediments to patient recruitment.
Morse, *et al.*, 1995[25] (USA)	Issues of recruitment, retention, and compliance in community-based clinical trials with traditionally underserved populations	HIV/AIDS	Community	Telephone survey of 14 sites participating in trials	Respondents suggested that patient transport (21%) may affect protocol participation.
Siminoff, *et al.*, 1989[30] (USA)	Doctor–patient communication about breast cancer adjuvant therapy	Cancer	Hospital	Observation and audio-recording of 100 patient–doctor interactions and surveys of participants	Physicians uncomfortable with not treating younger patients or recommending a trial with a long course of therapy.
Taylor, *et al.*, 1984[17] (USA)	Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer	Cancer	Hospital	Survey of 91 participating clinicians	Fear of feeling responsible if patient did not receive treatment which turned out to be best mentioned by five.
Winn, *et al.*, 1984[22] (USA)	An evaluation of physician determinants in the referral of patients for cancer clinical trials in the community setting	Cancer	Hospital	Vignette study of 82 clinicians	Single most important factor in determining whether a patient will be referred is toxicity of experimental regimen (90% referral if toxicity low, 52% if toxicity high). No difference based on physician age, speciality or length of time in practice.

TABLE 8 *Loss of autonomy*

Study (country)	Title	Subject area	Setting	Method	Results
Taylor, *et al.*, 1984[17] (USA)	Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer	Cancer	Hospital	Survey of 91 participating clinicians	12/66 argues that they were pragmatists and did not want always to stick to the protocol.
Taylor, 1985[16] (USA)	The doctor's dilemma: physician participation in randomized clinical trials	Cancer	Hospital	Survey of 91 participating clinicians	Of those who did not enrol all patients: 73% made some reference to their relationship with patient, principally that ultimate decision-making power was removed from either physician or patient; 18% agreed they were pragmatists, preferring to act on clinical judgement even when it conflicted with published data.
Taylor, *et al.*, 1987[29] (Canada)	Physician response to informed consent regulations for randomized clinical trials	Cancer	Hospital	Survey of 170 participating clinicians	Some 54% disliked being accountable to a third party.
Taylor & Kelner, 1987[18] (various)	Interpreting physician participation in randomized clinical trials: the physician orientation profile	Cancer	Hospital and general practice	Survey of 484 participating clinicians	RCT restricts ability to individualise patient care (77%).
Taylor, 1992[19] (various)	Physician participation in a randomized clinical trial for ocular melanoma	Cancer/ ophthal-mology	Hospital and private practice	Survey of 101 participating physicians	Restriction of flexibility in trials. Doctor discomfort with need to randomise patients rather than select individual treatments.
Taylor, 1992[28] (USA)	Integrating conflicting professional roles: physician participation in randomized clinical trials	Cancer/ ophthal-mology	Hospital and private practice	Survey of 101 participating physicians	Clinically orientated doctor loses his customary independence of action when taking part in trials.
Taylor, *et al.*, 1994[20] (USA)	Fundamental dilemmas of the randomized clinical trial process: results of a survey of the 1737 Eastern Cooperative Oncology Group investigators	Cancer	Hospital and private practice	Survey of 1737 participating clinicians	In all, 82% of clinicians were reluctant to relinquish individual decision making. 68% had no difficulty with randomisation if they considered that a patient would not be compromised by being placed in any arm of trial.

Problems in complying with the protocol

The protocol itself was blamed for restricting recruitment in eight studies (five reports of patient accrual to trials and three surveys of participating clinicians) (*Table 10*).[12,14,15,37–41] In addition, excessive data collection[14] and poorly designed data collection[38] were criticised.

Recommendation

- The entry criteria should be as simple and clear as possible so that the study will accommodate all relevant patients. Data collection should be kept to the minimum consistent with the scientific purpose of the study. (Basis: anecdotal evidence.)

Consent procedure

Obtaining consent (see *Table 11*) was acknowledged as being an important barrier to patient recruitment,[32] although some clinicians believed that consent is not primarily to protect the interests of patients.[14] The consent process was identified as a reason for failing to recruit patients[14,16,17] and making patient recruitment difficult;[29] if not required, physicians reported that they would enter more patients.[18] Significant problems were the lack of time[42] and trained staff[26] needed to explain the trial and obtain consent. Plaisier and colleagues suggested that it was harder to gain informed consent if the treatment characteristics were divergent.[41] Chang and colleagues[31] and Williams and Zwitter[43] found that clinicians believed that informed consent may be easier to obtain using a pre-randomisation design; however, Gallo and colleagues found that patients pre-randomised to standard treatment were less likely to consent than those randomised to 'experimental'

TABLE 9 *Incompatibility of protocol with normal practice*

Study (country)	Title	Subject area	Setting	Method	Results
Deber & Thompson, 1990[35] (Canada)	Variations in breast cancer treatment decisions and their impact in mounting trials	Cancer	Hospital	Survey (with vignettes) of 234/662 clinical participants	Fewer clinicians would recruit patients to a trial with a no-treatment or chemo-therapy arm (37%) than a trial with treatments they agreed were acceptable (68%). Clinicians would need to agree that both treatment options are valid and on patient categories, prognostic and treatment variables.
Fisher, *et al.,* 1991[15] (USA)	Clinical trials in cancer therapy: efforts to improve patient enrolment by community oncologists	Cancer	Community	Survey of 75 clinicians who participated in studies	Choice of therapies were perceived as an obstacle to patient recruitment for 14%.
Klein, *et al.,* 1995[36] (Canada)	Physicians' beliefs and behavior during a randomized controlled trial of episiotomy: consequences for women in their care	Obstetrics	Hospital	Survey of 43 participating clinicians	Physicians who favoured one technique over the other had more difficulty following the protocol, citing concerns about patients' well-being to justify their failure to comply with the protocol.
Penn & Sheer, 1990[26] (UK)	Reasons for declining participation in a prospective randomised trial to determine the optimum mode of delivery of the preterm breech	Obstetrics	Hospital	Survey of 11 centres who refused to participate in a trial	One department said that the draft patient information sheet was incompatible with departmental policy and that the wording of the information would 'frighten many mothers'.
Siminoff, *et al.,* 1989[30] (USA)	Doctor–patient communication about breast cancer adjuvant therapy	Cancer	Hospital	Observation and audio-recording of 100 patient–doctor inter-actions and surveys of participants	Physicians reluctant to recruit patients with poor prognosis to a clinical trial in which they may be randomised to standard treatment; they would rather go for non-standard treatment.
Taylor, *et al.,* 1984[17] (USA)	Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer	Cancer	Hospital	Survey of 91 participating clinicians	Five stated a preference for segmental mastectomy and did not enter any patients.
Taylor, 1985[16] (USA)	The doctor's dilemma: physician participation in randomized clinical trials	Cancer	Hospital	Survey of 91 participating clinicians	Of those who did not enrol all patients: 18% agreed they were pragmatists, preferring to act on clinical judgement even when it conflicted with published data; 9% argued that time needed to explain to patients, rigid rules governing eligibility and inflexibility of prescribed treatment formulas were insurmountable barriers to accrual.
Tognoni, *et al.,* 1991[21] (Italy)	Randomised clinical trials in general practice: lessons from a failure	Cardio-vascular	General practice	Report of a failed trial	Patients had to be withdrawn from treatment to be assessed for suitability: apparent conflict with usual practice. GPs were said to have an 'unsatisfactory attitude ... [to] controlled research'.

TABLE 10 *Complying with protocol*

Study (country)	Title	Subject area	Setting	Method	Results
Benson, *et al.*, 1991[14] (USA)	Oncologists' reluctance to accrue patients onto clinical trials: an Illinois Cancer Center study	Cancer	Hospital	Survey of 224 clinicians who participated in trials	In all, 82% believed some or all protocols are too rigidly designed, with excessive data collection and follow-up.
Bowen & Hirsch, 1992[37] (UK)	Recruitment rates and factors affecting recruitment for a clinical trial of a putative anti-psychotic agent in the treatment of acute schizophrenia	Mental health	Hospital	Prospective data collection on 166 patients to establish rate of recruitment; only 1/7 patients were entered	Patients ineligible because of medical issues (21/66) or psychiatric issues (78/166).
Coombs, *et al.*, 1993[38] (USA)	Conceptual and methodologic problems in the evaluation of a new burn treatment modality	Burns treatment	Community	Case study of failed trial: 95 recruited (200 target)	Data collection was poorly designed. Inclusion criteria narrowed after retest data were collected, so fewer eligible patients were available than expected. Later, entry criteria were relaxed.
Fisher, *et al.*, 1991[15] (USA)	Clinical trials in cancer therapy: efforts to improve patient enrolment by community oncologists	Cancer	Community	Survey of 75 clinicians who participated in trials	Completing flow-sheets (22%), remembering active protocols (19%), mailing follow-up data time (17%), recording follow-up data (17%); all mentioned as problems.
Foley & Moertel, 1991[12] (USA)	Improving accrual into cancer clinical trials	Cancer	Hospital	Survey of 209 participating clinicians	Respondents identified complex protocol as restrictive to RCTs (27%).
Hunter, *et al.*, 1987[39] (USA)	Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log	Cancer	Community	Report on 9508 newly diagnosed patients eligible for cancer trials, of whom 3242 were on protocol	Main reasons for patients being ineligible for trials were: protocol design, 70%; clinically eligible patients excluded by physician decision, 51%; patient refusal, 32%; follow-up problems or concomitant medical problems, 10%; other reasons, 7%.
Jack, *et al.*, 1990[40] (UK)	Recruitment to a prospective breast conservation trial: why are so few patients randomised?	Cancer	Hospital	Of 3054 patients referred to one Edinburgh Breast Unit, 63 were eventually found to be eligible for a trial; 23 refused to participate and requested specific adjuvant treatment	Over 57% of original patients excluded through failure to meet eligibility criteria. Of 84 patients treated for conservation, 23 did not fulfil the criteria for trials (e.g. tumour more extensive than appeared clinically).
Plaisier, *et al.*, 1994[41] (The Netherlands)	Unexpected difficulties in randomizing patients in a surgical trial: a prospective study comparing extra-corporeal shock wave lithotripsy with open cholecystectomy	Surgery	Hospital	Case study of trial of lithotripsy vs. cholecystectomy; only 8.3% of patients were recruited	Eligibility criteria for the experimental arm greatly reduced potential accrual.

TABLE 11 *Consent procedure*

Study (country)	Title	Subject area	Setting	Method	Results
Benson, *et al.*, 1991[14] (USA)	Oncologists' reluctance to accrue patients onto clinical trials: an Illinois Cancer Center study	Cancer	Hospital	Survey of 224 clinicians who participated in research	Necessity for physician to obtain informed consent identified as reason for failure to place patients on trials; 40% believed that consent is not obtained to protect patient but for other reasons; 59% medical oncologists and 37% surgeons stated that they sometimes used protocol guidelines without formally enrolling patients.
Chang, *et al.*, 1990[31] (USA)	Prerandomization: an alternative to classic randomization. The effects on recruitment in a controlled trial of arthroscopy for osteo-arthrosis of the knee	Ortho-paedic surgery	Hospital	Report of a trial which appeared to be failing but which was a success following change of design (to pre-randomisation)	Authors propose a modification to Zelen's design[44] based on patient preference. Operative processes lend themselves less well to classic randomisation procedures than medical processes. They imply that informed consent is easier to obtain under this system.
Gallo, *et al.*, 1995[45] (Italy)	Informed versus randomised consent to clinical trials	N/A	Community	A total of 2035 healthy visitors to scientific exhibition were enrolled in hypothetical trial and randomly assigned to groups: one-sided informed consent (n = 622); two-sided informed consent (n = 376); randomised consent to experimental treatment (n = 730); randomised consent to standard treatment (n = 307)	Agreement to participate was different, depending on type of consent procedure. One-sided informed consent (16% refused to participate), two-sided informed consent (20% refused), randomised consent experimental (12% refused), randomised consent standard (49% refused). More subjects prerandomised to standard treatment refused consent than those prerandomised to experimental (52% vs. 13%) (authors note that labelling treatment as A and B may have given different results).
Langley, *et al.*, 1997[32] (Canada)	Why are (or are not) patients given the option to enter clinical trials?	Cancer	Hospital and general practice	Survey of 87 participating and non-participating clinicians	Obtaining consent was rated third in importance among reasons for not offering patients entry to trials (after scientific design and doctor–patient rapport).
Penn & Steer, 1990[26] (UK)	Reasons for declining participation in a prospective randomised trial to determine the optimum mode of delivery of the preterm breech	Obstetrics	Hospital	Survey of 11 centres who refused to collaborate in an RCT of Caesarean section for preterm breech deliveries; reasons for non-participation	Physicians highlight problem of acquiring consent; 5/11 centres refused because of insufficient staff to gain truly informed consent.
Plaisier, *et al.*, 1994[41] (The Netherlands)	Unexpected difficulties in randomizing patients in a surgical trial: a prospective study comparing extracorporeal shock wave lithotripsy with open cholecystectomy	Surgery	Hospital	Over a 3-year period, only 8.3% of patients could be entered into a trial of lithotripsy vs. open cholecystectomy; main reasons for poor recruitment investigated	Acquiring informed consent is difficult if treatment characteristics are divergent.

continued

TABLE 11 contd *Consent procedure*

Study (country)	Title	Subject area	Setting	Method	Results
Schaeffer, *et al.,* 1996[42] (USA)	The impact of disease severity on the informed consent process in clinical research	Various (cancer and HIV/AIDS)	Clinical research centre	Patients from various trials (n = 73) and healthy volunteers (n = 52); questionnaires to assess knowledge of purpose and conduct of research, reading and comprehension levels	The longer the time spent explaining the consent form, the greater the level of immediate and long-term retention. Suggests written consent forms are limiting and research needed on differing ways of providing patient with information and not just on simplifying consent form, i.e. oral information.
Taylor, *et al.,* 1984[17] (USA)	Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer	Cancer	Hospital	Survey of 91 participating clinicians	None of the 25 respondents who cited trouble with informed consent entered all their patients into the trial.
Taylor, 1985[16] (USA)	The doctor's dilemma: physician participation in randomized clinical trials	Cancer	Hospital	Survey of 91 participating clinicians	Of those who did not enrol all patients, 38% cited trouble with informed consent, e.g. disclosing information to patients. Appeared that surgeons may be preferentially selecting patients they consider would have no difficulty with informed consent document.
Taylor, *et al.,* 1987[29] (Canada)	Physician response to informed consent regulations for randomized clinical trials	Cancer	Hospital	Survey of 170 participating clinicians	Informed consent: 65% said obtaining it makes entering patients into RCTs difficult; 70% said it is not helpful in some cases; 41% said it leads to patient non-compliance. Only 30% believed that all patients should always be required to sign a consent form before being placed in a clinical trial and 51% disliked certifying that they had obtained consent. Having to obtain informed consent is considered an intrusion into the privacy of the doctor–patient relationship by 90%; 24% said it affected their relationship with trial patients; 67% found informed consent difficult to obtain immediately following disclosure of news of serious illness.
Taylor & Kelnes, 1987[18] (various)	Interpreting physician participation in randomized clinical trials: the physician orientation profile	Cancer	Hospital and general practice	Survey of 484 participating clinicians	'If written informed consent was not required, I would enter more patients into clinical trials' (87%).
Williams & Zwitter, 1994[43] (various)	Informed consent in European multicentre randomised clinical trials – are patients really informed?	Cancer	Hospital	60/88 clinicians replied to a survey questionnaire about practice of obtaining informed consent	Written consent sought by 32% of clinicians; 21% used written information with no obligatory signing; 42% adopted verbal consent; 5% did not seek consent. Level of informed consent often fell short of that required by trial protocol. Some 70% of clinicians favoured prerandomised consent.

treatment.[45] Williams and Zwitter found that the level of informed consent may fall short of that required by the protocol.[43]

Recommendation

- The purpose and requirement for obtaining informed consent from the patient should be stressed in the study protocol. The procedures for giving information and obtaining consent should be designed to ensure that potential participants are given appropriate information, in a way that does not interrupt clinical care. This may require trained and dedicated staff to be available to provide information and obtain consent from patients. (Basis: logical argument.)

Barriers to patient participation

Even though clinicians may agree to participate in clinical trials, the ultimate success of an RCT depends on patient participation. While their participation should be facilitated, individual patients must also be free to decline to participate or to withdraw from a study at any time.[45] Some commentators have blamed patients for poor recruitment and it is therefore important to understand the reasons why patients may withhold or withdraw their consent.

Additional demands of the RCT on the patient
The review confirmed that the additional demands of a study may cause concern for some patients (*Table 12*). The main causes of concern were additional procedures and appointments which may cause discomfort, inconvenience or additional expense.

Extra procedures and time pressures were identified as barriers to recruitment and causes of 'drop-out' of patients in 13 studies. When asked for reasons for not agreeing to participate in trials, patients cited higher numbers of appointments,[47,48] venepuncture and inpatient hospital stays,[37] discomfort from medical procedures[49] and worry about experimentation.[39,50] Lack of available time to take part was a deterrent for some patients.[51,52] Patients who 'dropped-out' of trials blamed uncomfortable procedures[53] or excessive study demands.[54] High frequency of visits and time spent at clinic visits were rated as negative experiences by participants in trials,[27,55] as were uncomfortable procedures,[49] length of the study and additional procedures.[56]

Travel and travel costs (*Table 13*) were an important reason for refusing to take part in a trial,[57] for missing appointments and dropping-out of a trial,[58] and were found to be disliked by participants in two studies.[27,56]

Extra costs incurred by participants (*Table 14*) were found to be a reason for not participating in four trials[39,47,48,59] and those who withdrew from one study indicated that they believed patients should be paid to participate.[58]

In contrast to the barriers to participation, the most commonly mentioned motivation for participation was altruism.[27,30,55,60–67] More detailed discussions of the ethical issues associated with patient recruitment are presented elsewhere.[68,69]

Recommendation

- The demands of a study on patients should be kept to the minimum consistent with the scientific purpose of the study. In addition, the extent and purpose of the investigations should be clearly explained at the start of the study. Patients should be financially compensated for any travel costs incurred. An appeal to altruism may also be effective if patients are genuinely uncertain whether to take part. (Basis: logical argument, anecdotal evidence.)

Recommendation for research

- The motives for taking part in clinical trials should be further examined with a view to designing protocols which are more acceptable to patients.

Patient preferences

Patient preferences for a particular treatment option (offered in the trial or outside it) were given as a reason for non-participation in 20 studies (*Table 15*). A strong preference for or against a particular treatment was mentioned by non-participating patients in five studies. Preferences reflected wishes not to change medication,[37,70] not to take placebo,[59,71] not to take an experimental medication,[37,67] or not to take any medication.[37] Some patients requested a specific intervention.[40,41,59,72,73] In four studies, patients chose not to take part at all.[30,39,72,74] In a further three, the reason given was an aversion to treatment choice by random allocation.[40,52,76]

Langley and colleagues reported that clinicians identified patient refusal as the most common reason why patients were not entered into clinical trials.[32] Clinicians believe some patients refuse because they do not like the idea of randomisation[12] (which is supported by several studies of patients themselves[40,52,75]) and do not want to be 'experimental' subjects[12] (confirmed by Hunter, *et al.*;[39] Simel, *et al.*;[50] Robinson, *et al.*[77]). Clinicians may also find it difficult to recruit patients who have a strong preference for one treatment option.[19,41]

Recommendation

- The possible benefits and adverse effects of the treatment options should be described in a balanced way, together with the rationale for random allocation when the best approach is not known. No coercion should be used, however, to persuade patients to participate and the arrangements for care of those who choose not to participate should be part of the study protocol. The trial design should take into

TABLE 12 *Additional demands on the patient: extra procedures and time pressures*

Study (country)	Title	Subject area	Setting	Method	Results
Atwood, *et al.,* 1992[53] (USA)	Reasons related to adherence in community-based field studies	Cancer	Community	Reasons for good, marginal or poor adherence or withdrawal from three studies; qualitative data extracted from progress notes and end-of-study evaluations (142 patients)	Uncomfortable procedures were a cause of drop-out, e.g. colon examination procedures.
Autret, *et al.,* 1993[47] (France)	Parental opinions about biomedical research in children in Tours, France	Child health	Maternity unit	Survey of parents in a maternity unit about possible participation of their children in clinical trials: 582/986 response	Higher number of appointments were a barrier for 8% of those who would not participate.
Bowen & Hirsch, 1992[37] (UK)	Recruitment rates and factors affecting recruitment for a clinical trial of a putative anti-psychotic agent in the treatment of acute schizophrenia	Mental health	Hospital	Data collected prospectively on 166 patients to establish reasons for not entering a trial; only one in seven entered trial	Reason for non-consent given as: venepuncture 15%; 3-week inpatient stay 9%; ECG 1.5%.
Cunny & Miller, 1994[49] (USA)	Participation in clinical drug studies: motivations and barriers	N/A	Clinical research organisation and students	Survey of 195 participants, plus 68 considering participating, in clinical drug studies (response rate not known but 1135 questionnaires were distributed)	Non-participants cited discomfort from medical procedures (32.4%); participants disliked discomfort from medical procedures (21.4%).
Harth & Thong, 1990[48] (Australia)	Sociodemographic and motivational characteristics of parents who volunteer their children for clinical research: a controlled study	Child health	Hospital	Survey of parents of children who had taken part in an RCT of ketotifen for asthma (68) and parents who had refused (42)	Reason for not volunteering: inconvenience of frequent visits 35%.
Henzlova, *et al.,* 1994[27] (USA)	Patient perception of a long-term clinical trial: experience using a close-out questionnaire in the Studies of Left Ventricular Dysfunction (SOLVD) trial	Cardio-vascular	Hospital	Close-out questionnaire to 5188 patients in treatment RCT or prevention RCT (74% response); patients were middle-aged or older and chronically ill; average follow-up was 37 months for treatment trial and 41 months for prevention trial	Frequency of visits and time spent at clinic was rated as a negative experience.

continued

TABLE 12 contd *Additional demands on the patient: extra procedures and time pressures*

Study (country)	Title	Subject area	Setting	Method	Results
Hunter, *et al.*, 1987[39] (USA)	Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log	Cancer	Community	Report on 9508 newly diagnosed patients eligible for cancer trials, of whom 3242 were on protocol	Discomfort about 'experimentation' accounted for 10% refusal of patients.
Moody, *et al.*, 1995[51] (USA)	Search and research: factors influencing post-menopausal African-American women's participation in a clinical trial	Osteo-porosis	Community	Study examining recruitment of post-menopausal African-American women from various sources (e.g. churches, nurse associations, media campaigns) for trial evaluating effect of alendronate on bone density; 21 women were recruited (a yield of about 1 in 10)	Reasons for not participating included: family pressures, child care or other family duties which prohibited participation, belief that osteoporosis would not affect them, too busy.
Morrison, 1994[56] (Canada)	Trials and tribulations: patients' perspectives of the BETASERON study	Multiple sclerosis	Multiple sclerosis clinics	50 patients who had taken part in an intensive trial of multiple sclerosis treatments: 38 completed questionnaires about their participation at end of study (reasons, experience, etc.)	30% of patients complained about length of study. A further 30% found increased procedures unattractive.
Schwartz & Fox, 1995[52] (USA)	Who says yes? Identifying selection biases in a psychosocial intervention study of multiple sclerosis	Multiple sclerosis	Hospital	Survey of patients who refused to participate in RCT of two psychosocial interventions for multiple sclerosis; reason for refusal obtained from 107 patients	Reasons for refusing included 'no time' (4%).
Simel & Feussner, 1991[50] (USA)	A randomised controlled trial comparing quantitative informed consent formats	N/A	Hospital	Study using 'sham' trial. Patients randomised to receive one of two consent forms: one described a treatment that may work 'twice as fast as usual treatment' (n = 52), the other 'half as fast as usual treatment' (n = 48)	Fear of experimentation given as reason for not participating.
Sutherland, *et al.*, 1993[55] (Canada)	A study of diet and breast cancer prevention in Canada: why healthy women participate in controlled trials	Cancer	Community	Postal survey of 90 women randomly selected from participants (n = 418) in a study of diet and breast cancer prevention; responses from 33 in control and 33 in intervention group	Attending appointments was given as main disadvantage of taking part in the trial.
Wingerson, *et al.*, 1993[54] (USA)	Personality traits and early discontinuation from clinical trials in anxious patients	Mental health	Trial participants	Survey of 112 patients participating in clinical trials for panic disorder and generalised anxiety disorder; comparison of those who completed the study and those who discontinued early	Authors suggest that excessive study demands and complex dosing contributed to patient drop-out.

TABLE 13 *Additional demands on the patient: travel problems*

Study (country)	Title	Subject area	Setting	Method	Results
Henzlova, *et al.*, 1994[27] (USA)	Patient perception of a long-term clinical trial: experience using a close-out questionnaire in the Studies of Left Ventricular Dysfunction (SOLVD) trial	Cardio-vascular	Hospital	Close-out questionnaire to 5188 patients in treatment RCT or prevention RCT (74% response); patients middle-aged or older, and chronically ill; average follow-up 37 months for treatment and 41 months for prevention trial	Travel to and from the clinic (21%) and parking (16%) were cited as a negative experiences for participants.
Morrison, 1994[56] (Canada)	Trials and tribulations: patients' perspectives of the BETASERON study	Multiple sclerosis	Hospital	Of 50 patients who took part in an intensive trial of treatments for multiple sclerosis, 38 completed questionnaires at end of study about their participation (reasons, experience, etc.)	Patients highlighted travel to clinic as an inconvenience; expenses associated with travel were disliked.
Orr, *et al.*, 1992[58] (USA)	Patient and clinic factors predictive of missed visits and inactive status in a multicenter clinical trial (Macular Photo-coagulation Study Group)	Ophthal-mology	Hospital	Survey of 19 patients lost to follow-up; 59 with missed visits; 97 with no missed visits	Travel and travel costs were identified as reasons for failure to participate in RCT. Clinic factors associated with drop-out included cost of travel (OR 8.9).
Schwartz & Fox, 1995[52] (USA)	Who says yes? Identifying selection biases in a psychosocial intervention study of multiple sclerosis	Multiple sclerosis	Hospital	Survey of patients who refused to participate in RCT of two psychosocial interventions for multiple sclerosis; reason for refusal obtained from 107 patients	36% of those who refused to take part cited 'logistics problems' (usually lack of transport); unwillingness to travel to Boston (2%).

TABLE 14 *Additional demands on the patient: extra costs*

Study (country)	Title	Subject area	Setting	Method	Results
Autret, *et al.*, 1993[47] (France)	Parental opinions about biomedical research in children in Tours, France	Child health	Maternity unit	Survey of parents in maternity unit about possible participation of their children in clinical trials: 582/986 response	Financial constraints given as reason by 4% of those who would not participate.
Harth & Thong, 1990[48] (Australia)	Sociodemographic and motivational characteristics of parents who volunteer their children for clinical research: a controlled study	Child health	Hospital	Survey of parents of children who had taken part in RCT of ketotifen for asthma (68) and parents who had refused (42)	Extra costs given as reason for not volunteering by 5%.
Hunter, *et al.*, 1987[39] (USA)	Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log	Cancer	Community	Report on 9508 newly diagnosed patients eligible for cancer trials, of whom 3242 were on protocol	Extra costs accounted for 2% of refusals by patients.
Orr, *et al.*, 1992[58] (USA)	Patient and clinic factors predictive of missed visits and inactive status in a multicenter clinical trial (Macular Photocoagulation Study Group)	Ophthal-mology	Hospital	Survey of 19 patients lost to follow-up: 59 with missed visits and 97 with no missed visits	Survey indicated that patients considered that they should be paid to participate. Clinic factors associated with drop-out included: cost of travel (OR 8.9); belief that patients should be paid for participation (OR 3.6).
Yeomans-Kinney, *et al.*, 1995[59] (USA)	Factors related to enrolment in the breast prevention trial at a comprehensive cancer center during the first year of recruitment	Cancer	Community	Factors predicting enrolment in an RCT: only 45% of 232 white women agreed to participate	Concern about personal expenses was more of an issue for non-participants (OR 3).

TABLE 15 *Patient preferences*

Study (country)	Title	Subject area	Setting	Method	Results
Bergman, *et al.,* 1994[71] (France)	A randomized clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain	Child health	Maternity unit	RCT to investigate effect of informed consent on analgesic activity of placebo and naproxen in cancer pain: 25 received treatments without information; 24 received information, of whom 6 refused to participate	6/24 patients refused to participate because they did not want to take placebo.
Bevan, *et al.,* 1992[70] (UK)	Patients' attitudes to participation in clinical trials	N/A	Community	Structured interviews with 66 patients who had taken part in clinical trials, 12 who had declined to take part and 119 who had never been invited	Patients who had not taken part in trials and would be unwilling to do so gave the following reasons: too ill (22%), worried about side-effects (17%), concerned about changing treatments (22%). Of the 12 who refused to take part in trials, three did not want to alter current therapy, three had insufficient time, three had relatives who objected to their participation.
Bowen & Hirsch, 1992[37] (UK)	Recruitment rates and factors affecting recruitment for a clinical trial of a putative anti-psychotic agent in the treatment of acute schizophrenia	Child health	Hospital	Data collected prospectively on 166 patients to establish reasons for not entering a trial; only one in seven entered trial	Reasons for non-consent given as did not want: any medication (6%); change to medication (1.5%); new experimental medication 9%.
Dahan, *et al.,* 1986[72] (France)	Does informed consent influence therapeutic outcome? A clinical trial of the hypnotic activity of placebo in patients admitted to hospital	Insomnia	Hospital	Study to examine effect of placebo when given with or without informed consent to patients suffering with insomnia; single blinded observer, blinded trial of patients paired according to sex, age and hospital environment; randomisation assigned first patient in pair to control or informed consent group. 26 patients refused consent, 30 pairs were given placebo treatment	26/86 patients refused to give consent because of: reluctance to participate (n = 12); fear of side-effects (n = 10); wishing to take a well-known drug (n = 4).
Foley & Moertel, 1991[12] (USA)	Improving accrual into cancer clinical trials	Cancer	Hospital	Survey of 209 participating clinicians	Some considered that patient refusal of randomisation (26%), refusal to be experimental subjects (24%), and patients unlikely to comply with protocol requirements (18%) were impediments to recruitment.
Hunter, *et al.,* 1987[39] (USA)	Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log	Cancer	Community	Report on 9508 newly diagnosed patients eligible for cancer trials, of whom 3242 were on protocol	32% of clinically eligible patients declined to participate. 'Experimentation' accounted for 10% of patients who refused.

continued

TABLE 15 contd *Patient preferences*

Study (country)	Title	Subject area	Setting	Method	Results
Jack, *et al.,* 1990[40] (UK)	Recruitment to a prospective breast conservation trial: why are so few patients randomised?	Cancer	Hospital	Of 3054 patients referred to one Edinburgh breast unit in 1988, 63 were eventually found to be eligible for a trial; 23 refused to take part and requested specific adjuvant treatment	A proportion of patients may refuse to take part in RCTs because they do not like the idea of a random decision (no evidence). 7/23 patients not entered requested specific intervention (role of the lay press in influencing patient preferences).
Langley, *et al.,* 1987[32] (Canada)	Why are (or are not) patients given the option to enter clinical trials?	Cancer	Hospital and general practice	Survey of 52 oncologists, 26 nurses and nine family physicians; asked reasons for offering patients entry into clinical trials and for not offering entry	Patient refusal identified as most common reason why patients were not entered into clinical trials (52% oncologists, 42% nurses, 56% family physicians).
Llewellyn-Thomas, *et al.,* 1991[76] (Canada)	Patients' willingness to enter clinical trials: measuring the association with perceived benefit and preference for decision participation	Cancer	Hospital	60 non-eligible patients were asked to decide whether or not he/she would be willing to enter trial; use of a probability trade-off method	22/35 who refused reported aversion to randomisation as their main reason for trial refusal.
Plaisier, *et al.,* 1994[41] (The Netherlands)	Unexpected difficulties in randomizing patients in a surgical trial: a prospective study comparing extracorporeal shock wave lithotripsy with open cholecystectomy	Surgery	Hospital	Case study of trial of lithotripsy vs. open cholecystectomy: over 3-year period only 8.3% of patients could be entered into the study	Patient preference played an important role, with patients requesting experimental technique and subsequently requesting new gold standard (laparoscopic cholecystectomy) (role of media in guiding patient preference).
Robertson, 1994[75] (USA)	Clinical trial participation. Viewpoints from racial/ethnic groups	Cancer	Community	Phone interviews on views and opinions about clinical trial participation with eight African-Americans, ten Hispanic and ten Native Americans	Reasons given for 'people of your race' not taking part in experimental studies: fear, lack of information, mistrust of being treated like guinea pigs, ethnic background, do not like to get involved.
Robinson, *et al.,* 1996[77] (USA)	Attitudes of African-Americans regarding prostate cancer clinical trials	Cancer	Community	Qualitative study using focus groups to examine attitudes of African-American males regarding prostate cancer clinical trials	Barriers included: suspicion associated with being used as guinea pig, fear of side-effects, mistrust of medical establishment, inexperienced or incompetent physician.
Schwartz & Fox, 1995[52] (USA)	Who says yes? Identifying selection biases in a psychosocial intervention study of multiple sclerosis	Multiple sclerosis	Hospital	Survey of patients who refused to participate in RCT of two psycho-social interventions for multiple sclerosis; reason for refusal obtained from 107 patients	6% of refusers cited unwillingness to be randomised.
Simel & Feussner, 1991[50] (USA)	A randomized controlled trial comparing quantitative informed consent formats	N/A	Hospital	Study using 'sham' trial; patients randomised to receive one of two consent forms: one described a treatment that may work 'twice as fast as usual treatment' (n = 52), the other 'half as fast as usual treatment' (n = 48)	Fear of experimentation given as the main reason for not participating.

continued

TABLE 15 contd *Patient preferences*

Study (country)	Title	Subject area	Setting	Method	Results
Siminoff, *et al.*, 1989[30] (USA)	Doctor–patient communication about breast cancer adjuvant therapy	Cancer	Hospital	Observation and audio-recording of 100 patient–doctor interactions and surveys of participants	12/22 patients preferred not to join trials they were offered.
Slevin, *et al.*, 1995[67] (UK)	Volunteers or victims: patients' views of randomised cancer clinical trials [see comments]	Cancer	Hospital	Questionnaire survey of 75 clinic patients about attitudes towards taking part in trials	Of those who were uncertain or would refuse to take part in trials, 51% would prefer doctor to make decision, 33% would worry about receiving new treatment.
Stone, *et al.*, 1994[74] (Australia)	Selection of patients for randomised trials: a study based on the MACOP-B vs. CHOP in NHL study	Cancer	Hospital	Comparison between eligible (except for consent) non-trial participants and all randomised patients	Of 32 patients considered eligible but who were not recruited, seven refused to take part.
Taylor, 1992[19] (various)	Physician participation in a randomized clinical trial for ocular melanoma	Cancer/ ophthal-mology	Hospital and private practice	Survey of 101 physicians in Collaborative Ocular Melanoma study	Difficulty of randomising patients who strongly prefer one treatment option.
Tindall, *et al.*, 1994[73] (Australia)	Effects of two formats of informed consent on knowledge amongst persons with advanced HIV disease in a clinical trial of didanosine	HIV/AIDS	Hospital	113 HIV patients in controlled trial were surveyed about informed consent procedure, then randomised to receive information about trial in written formal (n = 52) or written and verbal format (n = 61)	79% stated that patients should be allowed the choice between participating in a trial of unproved medications and receiving it outside trial.
Yeomans-Kinney, *et al.*, 1995[59] (USA)	Factors related to enrolment in the breast cancer prevention trial at a comprehensive cancer center during the first year of recruitment	Cancer	Community	Factors predicting enrolment in an RCT; only 45% of 232 white women agreed to participate	Logistic regression predicted non-participation from: concern about oestrogen replacement therapy contraindication (OR 12), tamoxifen side-effects (OR 5), possibility of placebo treatment (OR 8), concern about personal expenses (OR 3), concern about significant others not reassured if put on tamoxifen (OR 3).

account reasons why some patients may refuse participation, for example, change in medication should be minimised; the use of placebo must be justifiable both scientifically and ethically; and the process of randomisation should be presented as an extension of standard medical practice. (Basis: logical argument.)

Recommendation for research

- Further research is required to find the most appropriate methods to describe trials, particularly randomisation, to patients.

Worry about uncertainty

Patients may find the issues of uncertainty difficult to cope with (*Table 16*), particularly if they feel the efficacy of the treatment on offer is unproven.[47] Patients may have a distrust of hospital or medicine[47,48,76,77] and fear the unknown.[52] Under such circumstances, they find it unpleasant having to decide about taking part in a trial[62,78] and may even prefer the doctor to make the decision.[67]

Recommendation

- Patients deserve consideration and sympathy as they decide whether or not to take part in an RCT. They should not be pressurised into taking part. In the longer term, there is a need for ongoing public education about the appropriateness of RCTs when there is clinical uncertainty. (Basis: logical argument.)

Recommendation for research

- Further research is needed to identify the most appropriate methods to overcome patient worry about the uncertainties involved in taking part in RCTs.

TABLE 16 *Worry about uncertainty*

Study (country)	Title	Subject area	Setting	Method	Results
Autret, *et al.,* 1993[47] (France)	Parental opinions about biomedical research in children in Tours, France	Child health	Maternity unit	Survey of parents in a maternity unit about possible participation of their children in clinical trials; 582/986 response	Efficacy unproven was considered a barrier by 49% of those who would not participate.
Corbett, *et al.,* 1996[78] (UK)	Offering patients entry in clinical trials: preliminary study of the views of prospective participants	N/A	Community	Survey of members of the public, students and secretaries (n = 100) to find views on RCTs and randomisation	55% of respondents thought that they would find it upsetting to be asked to participate in RCT. 33% thought that participating in trial would affect recovery (63.5% for the worse, 36.4% for the better).
Harth & Thong, 1990[48] (Australia)	Sociodemographic and motivational characteristics of parents who volunteer their children for clinical research: a controlled study	Child health	Hospital	Survey of parents of children who had taken part in an RCT of ketotifen for asthma (68) and parents who had refused (42)	Reasons for not volunteering: distrust of modern medicine 22%, distrust of the hospital 8%.
Jensen, *et al.,* 1993[62] (Denmark)	Information for cancer patients entering a clinical trial – an evaluation of an information strategy	Cancer	Hospital	Interviews with 34 women invited to join three adjuvant therapy trials, 3 months after they were given information about trials (18 participants)	50% felt it difficult or unpleasant to decide about entering a trial.
Moody, *et al.,* 1995[51] (USA)	Search and research: factors influencing post-menopausal African-American women's participation in a clinical trial	Osteoporosis	Community	Study examining the recruitment of post-menopausal African-American women to a trial to evaluate effect of alendronate on bone density	Reasons for not participating included anxiety and fear of the unknown.
Robertson, 1994[75] (USA)	Clinical trial participation. Viewpoints from racial/ethnic groups	Cancer	Community	Phone interviews about views and opinions on clinical trial participation with eight African-Americans, ten Hispanic and ten Native Americans	Reasons given for 'people of your race' not taking part in experimental studies: fear, lack of information, mistrust of being treated like guinea pigs, ethnic background, do not like to get involved.
Robinson, *et al.,* 1996[77] (USA)	Attitudes of African-Americans regarding prostate cancer clinical trials	Cancer	Community	Qualitative study using focus groups to examine attitudes of African-American males regarding prostate cancer clinical trials	Barriers included: suspicion associated with being used as a guinea pig; fear of side-effects; mistrust of medical establishment; inexperienced or incompetent physician.
Slevin, *et al.,* 1995[67] (UK)	Volunteers or victims: patients' views of randomised cancer clinical trials	Cancer	Hospital	Questionnaire survey of 75 clinic patients about attitudes towards taking part in trials	Of those who were uncertain or would refuse to take part in trials, 51% would prefer doctor to make decision, 33% would worry about receiving new treatment.

Concerns about information and consent

Information and consent emerged as major concerns. These aspects of the conduct of clinical trials are issues which lend themselves relatively easily to publishable research (for example, RCTs of different designs of patient information and surveys of patient attitudes towards consent to participation in trials). It is therefore not surprising that the largest number of published papers were related to this area.

Information

Patients require full and open information to be able to make an informed choice about participating in an RCT.[45] Such information

should include the rationale for the study, details of the treatment options with the possible benefits and adverse effects, an explanation of the randomisation process, and the practical implications of participation. Patient information was covered in 33 papers in the main database: 15 surveys of patients or the public, ten RCTs and one non-randomised study of different ways of providing information to patients (*Table 17*), seven surveys of clinicians (*Table 18*).

Bergmann and colleagues[71] carried out a small RCT to investigate the effect of giving cancer inpatients information about a crossover, double-blind study of naproxen and placebo in mild or moderate cancer pain. Patients were randomised either to receive both treatments without any information, or to be given information about the study. Giving information to patients reduced recruitment. Among the patients who agreed to take part, the analgesic response was higher than in patients who had not received information. The authors point to the effect of information in altering apparent efficacy of some experimental treatments. In a study by Myers and colleagues, patients were more likely to discontinue treatment if they were informed about side-effects.[79] The study was confounded by the fact that different wordings were used at different study centres making interpretation difficult. However, wording of information is clearly important for recruitment, compliance and efficacy.

Some studies have shown that patients – both participating and non-participating – and healthy volunteers may want more information and that they tend to have a preference for written information on trials.[56,70,78,80] Some authors found the level of information given was acceptable to patients,[62,73] particularly those with higher levels of education.[81] Other studies provided information at too complex a level.[48,66]

Several studies investigated different forms of information presentation. Fetting and colleagues found that verbal presentation of information was better than numerical description.[82] The reverse was found in a study by Simel and Feussner, in which quantitative data increased the likelihood of agreeing to participate in a 'sham' trial.[50,83] Llewellyn-Thomas and colleagues[84] also investigated the use of numerical descriptions of rates of side-effects, phrased in negative, neutral or positive versions; the framing of information was not found to influence preferences for trial entry but this may have been because of the complexity of the descriptions. Another method applied by Llewellyn-Thomas and colleagues[76] to study the reasons why patients accept or refuse trial entry, was a probability trade-off approach to describing a hypothetical trial, in which patients could trade-off short-term toxicity in order to achieve a possible gain in long-term benefit. The method was difficult to understand, as indicated by the number of 'aberrant cases'. Another comparison carried out by this group[85] involved presenting teaching about a clinical trial as an audio tape or as an interactive computer program; there was no detectable difference in the numbers recruited using the two approaches (again the information presented was complex).

A combination of several information strategies (for example, written, oral, video) may improve awareness and knowledge, and knowledge retention.[64,86] Receiving information from the doctor and nurse, and being given sufficient time to read the information leaflet led to best perceived understanding (no comment on actual understanding).[87] DeLuca and colleagues' finding that patients do not read information may be as a consequence of having insufficient time.[60] Aaronson and colleagues found that patients who had supplementary information from nurses were better informed.[10] Conversely, Tindall and colleagues found no difference in knowledge between written and written plus oral consent procedures,[73] and Wadland and colleagues found no difference in recruitment rate between groups who had information read to them by a study coordinator, and those who read the information themselves.[23]

Recall of information may be poor and patients may consider themselves as better informed than they actually are.[88] In a randomised comparison of 'total disclosure' (a systematic way of delivering information to patients) with individual disclosure, total disclosure led to better understanding of treatments but no better understanding of the concept of randomisation.[33] Total disclosure also resulted in less willingness to take part in RCTs.[66] Illness severity affects the amount and type of information retained, with more severely ill patients retaining less information about the illness but more information about the protocol.[42]

Fetting and colleagues found that some patients who declined to take part in a trial spontaneously indicated that it was because of limited reading skills or because English was not their primary language.[82]

Many clinicians find giving information difficult (*Table 18*), both describing clinical trials in

TABLE 17 *Information: patient perspective*

Study (country)	Title	Subject area	Setting	Method	Results
Aaronson, *et al.,* 1996[10] (The Netherlands)	Telephone-based nursing intervention improves the effectiveness of the informed consent process in cancer clinical trials	Cancer	Hospital	RCT of telephone-based nursing information plus standard informed consent procedure vs. standard procedure; 180 patients took part from 346 eligible. Comparison of randomised and non-randomised patients to discuss possible selection bias and generalisability; physicians also asked for reasons patients not entered	Patients who received supplementary information from oncology nurse were significantly better informed about many aspects of the trial.
Bergmann, *et al.,* 1994[71] (France)	A randomised clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain	Cancer	Hospital	RCT to investigate effect of informed consent on analgesic activity of placebo and naproxen in cancer pain: 25 patients received treatment without information; 24 received information (6 refused to participate, 18 participated)	Giving information about the study reduced recruitment (6/24 with information refused to participate).
Bevan, *et al.,* 1992[70] (UK)	Patients' attitudes to participation in clinical trials	General medical	Hospital	Structured interviews with patients who had taken part in clinical trials (n = 66), 12 who had declined to take part and 119 who had never been invited	Of those who had taken part, 60% would have liked written information but only 38% had received information in this form.
Corbett, *et al.,* 1996[78] (UK)	Offering patients entry in clinical trials: preliminary study of the views of prospective participants	N/A	Community	Survey of members of the public, students and secretaries (n = 100) to find views on RCTs and randomisation	Great majority of respondents preferred written information (90.8%). Overwhelming preference was for more detailed information.
DeLuca, *et al.,* 1995[60] (USA)	Are we promoting true informed consent in cardiovascular clinical trials?	Cardiovascular	Hospital	Questionnaire administered to all eligible patients approached for participation in clinical trial (n = 247); non-consenters (n = 75) compared with consenters (n = 172)	30% of patients did not read given information. Authors suggest that: patients ignore information and adopt a 'physician-knows-best' strategy; lower education groups may exclude themselves from RCTs by not volunteering because they do not understand nature of trial. Enhanced readability of consent forms is suggested.
Dunbar, *et al.,* 1989[86] (USA)	Implementation of a multicomponent process to obtain informed consent in the diabetes control and complications trial	Diabetes	Clinical centres	Report of an informed consent education programme which formed part of a diabetes control and complications trial: only those who agreed to participate were followed-up (n = 278)	Use of multi-component presentation of information increased awareness and knowledge of procedures, and subsequent retention.
Fetting, *et al.,* 1990[82] (USA)	Effect of patients' expectations of benefit with standard breast cancer adjuvant chemotherapy on participation in a randomized clinical trial: a clinical vignette study	Cancer	Hospital	282 female cancer patients were randomised to receive vignettes with either a numerical description of disease-free survival or a description in words ('verbal'); randomisation 3:1 to verbal group	Patients considering clinical trial rely heavily on their doctor's recommendation. Information on trial and standard treatment should be included. Verbal description produced better recruitment than numerical description. 24 patients who would decline to enter trial spontaneously indicated that it was because of limited reading skills or English not being their primary language.

continued

TABLE 17 contd *Information: patient perspective*

Study (country)	Title	Subject area	Setting	Method	Results
Harth & Thong, 1995[61] (Australia)	Parental perceptions and attitudes about informed consent in clinical research involving children	Child health	Hospital	Questionnaire survey of 62/64 parents whose children had taken part in RCT of ketotifen for asthma	Most informed consent forms could only be understood by parents with a college education.
Jensen, *et al.*, 1993[62] (Denmark)	Information for cancer patients entering a clinical trial – an evaluation of an information strategy	Cancer	Hospital	Survey of patients eligible for three clinical trials (n = 34); 18 (53%) agreed to take part	'Evaluation' of new information procedure: 31/34 rated the information good or very good.
Llewellyn-Thomas, *et al.*, 1991[76] (Canada)	Patients' willingness to enter clinical trials: measuring the association with perceived benefit and preference for decision participation	Cancer	Hospital	60 non-eligible patients were asked to decide if they would be willing to enter a trial; used a probability trade-off method	Suggests the manner by which information is presented to patients affects their potential accrual.
Llewellyn-Thomas, *et al.*, 1995[84] (Canada)	Cancer patients' decision-making and trial-entry preferences: the effects of 'framing' information about short-term toxicity and long-term survival	Cancer	Hospital	90 patients, randomly allocated to receive different forms of information; to test the effects of different 'informational frames' (neutral, positive or negatively framed information) on treatment decision making and trial-entry decision	Authors found that context framing, i.e. setting side-effects in either a neutral, negative or positive frame, had no effect on hypothetical entry to a trial (however, information given was complex).
Llewellyn-Thomas, *et al.*, 1995[85] (Canada)	Presenting clinical trial information: a comparison of methods	Cancer	Hospital out-patients	100 patients randomised to receive information about hypothetical trial by audio a tape or by interactive computer program	No differences in understanding after 1 day, no difference in satisfaction after 1 week. 52% would agree to take part in the trial. Members of computer program group tended to report more positive attitudes toward trial entry. Refusers tended to be women with scores indicating higher understanding.
Lynoe, *et al.*, 1991[64] (Sweden)	Informed consent: study of quality of information given to participants in a clinical trial	Gynae-cology	Hospital	43/48 women traced 18 months after trial returned questionnaires: 35 remembered giving consent, seven did not, and one claimed non-consent	Information about conduct of trial was not always absorbed by patients; oral and written information better than just oral; signed consent could imply inability of patients to withdraw from study – 15/43 patients did not recall that they could withdraw at any time. 37 felt they had been given information at an appropriate time.
Maslin, 1994[80] (UK)	A survey of the opinions on 'informed consent' of women currently involved in clinical trials within a breast unit	Cancer	Hospital out-patients	300 women surveyed at a breast clinic, all with family history of breast cancer: 100 healthy volunteers not in a trial; 100 healthy volunteers in an RCT; 100 symptomatic women in an RCT	90% of survey respondents wanted all the information and support available; 75% wanted assurance that patient information is confidential; 66% wanted information on protection of patient's privacy. Healthy volunteers received more information (86% vs. 62%), were better informed, and more were informed that they could withdraw (84% vs. 46%) than breast cancer patients (but comparing completely different groups of patients/volunteers).

continued

TABLE 17 contd *Information: patient perspective*

Study (country)	Title	Subject area	Setting	Method	Results
Miller, *et al.,* 1994[88] (USA)	Comprehension and recall of the informational content of the informed consent document: an evaluation of 168 patients in a controlled clinical trial	Medical	Community	Phone survey of 168 patients participating in trial comparing two analgesic drugs: asked about recall and satisfaction with informed consent process	60 days after presentation of consent form more than half (52.4%) of patients were unable to recall any of 12 side-effects of drugs; 98% stated that they understood the information given in consent form. Authors found that education and age were not associated with understanding.
Morrison, 1994[56] (Canada)	Trials and tribulations: patients' perspectives of the BETASERON study	Multiple sclerosis	Multiple sclerosis clinics	50 patients who participated in intensive trial of multiple sclerosis treatments: 38 completed questionnaires at end of study about their participation (reasons, experience, etc.)	Information provided in both oral and written forms. 95% found this useful but 26% wanted more information; 5% were uncertain they had enough information to make a decision.
Myers, *et al.,* 1987[79] (USA)	The consent form as a possible cause of side-effects	Cardio-vascular	Hospital	Opportunistic study of two different patient information wordings	Over six times as many patients discontinued treatment in centres where side-effects were mentioned in consent form as opposed to centres where side-effects were not mentioned.
Olver, *et al.,* 1995[81] (Australia)	The adequacy of consent forms for informing patients entering oncological clinical trials	Cancer	Hospital	100 consecutive patients from 18 clinical trials; study looked at impact of information and consent forms	Patients reading form completely were higher educated, had English as first language and were inpatients at hospital; generally, patients who had read form thought content was about right and that the information had made no difference to their state of anxiety; younger patients tended to recall more information than older patients.
Ross, *et al.,* 1993[66] (USA)	Reasons for entry into and understanding of HIV/AIDS clinical trials: a preliminary study	HIV/AIDS	AIDS clinic	Survey of 32 trial participants regarding reasons for participating in trials	Total disclosure of information leads to better understanding of treatment side-effects, etc. but also results in less willingness to participate in RCTs.
Schaeffer, *et al.,* 1996[42] (USA)	The impact of disease severity on the informed consent process in clinical research	Various (HIV and cancer)	Medical research centre	Subjects from various trials (n = 73) and healthy volunteers (n = 52); questionnaires to assess knowledge of purpose and conduct of research, reading and comprehension levels	Illness severity affected amount of overall information retained by individuals: the more severe the illness, the less information is retained. Specific information about procedures showed reversal of this trend with severity of illness dictating procedural information retained; suggests clinician should reinforce information given during subsequent visits in order to maintain level of information available to patient. Severely ill patients want to be less involved in their health care.
Simel & Feussner, 1991[50] (USA)	A randomized controlled trial comparing quantitative informed consent formats	N/A	Hospital	Study using 'sham' trial. Patients randomised to receive one of two consent forms: one described treatment that may work 'twice as fast as usual treatment' (n = 52); the other 'half as fast as usual treatment' (n = 48)	Patients more likely to consent to RCT when information phrased in a positive way than when it was negatively presented; however, this was only the case for patients who suggested that quantitative data had influenced decision process.

continued

TABLE 17 contd *Information: patient perspective*

Study (country)	Title	Subject area	Setting	Method	Results
Simel & Feussner, 1992[83] (USA)	Suspended judgement. Clinical trials of informed consent	N/A	Hospital	Study using 'sham' trial. Patients randomised to receive one of two consent forms: one described treatment that may work 'twice as fast as usual treatment' (*n* = 52); the other 'half as fast as usual treatment' (*n* = 48)	Authors identified that information has to be clear, succinct and balanced. Results indicate that patients use quantitative data when making decisions.
Simes, *et al.,* 1986[33] (Australia)	Randomised comparison of procedures for obtaining informed consent in clinical trials of treatment for cancer	Cancer	Hospital	57 cancer patients recruited over 3-year period: randomised to receive total disclosure of information (*n* = 28) or individual approach (*n* = 29)	Total disclosure group more knowledgeable about side-effects, purpose of the research and mechanism for selection than individual approach group. More than 50% in each group failed to understand concept of randomisation.
Tankanow, *et al.,* 1992[87] (USA)	Patients' perceived understanding of informed consent in investigational drug studies	Various medical	Hospital	Interviews with 98 patients in investigational drug studies within 72 hours of enrolling	Best perceived understanding achieved by both nurse and doctor giving consent form to patient and by giving patient sufficient time to read it.
Tindall, *et al.,* 1994[73] (Australia)	Effects of two formats of informed consent on knowledge amongst persons with advanced HIV disease in a clinical trial of didanosine	HIV/AIDS	Hospital	113 HIV patients in controlled trial were randomised to receive information about trial in: written format (*n* = 52) or written and verbal format (*n* = 61)	Primary source of information used to influence decisions about their healthcare came from patients' specialist and GP. Majority considered that medical practitioner highlighted all possible treatments; 91% reported that they understood information sheet. No significant difference in knowledge found between written and written/verbal consent.
Wadland, *et al.,* 1990[23] (USA)	Recruitment in a primary care trial on smoking cessation	Smoking cessation	Primary care	Study to describe rates of recruitment during RCT of smoking cessation in two primary care practices	Interested smokers were randomly assigned to one of two methods of providing information: having form explaining study read by a study coordinator (*n* = 51); patients read form themselves (*n* = 53). No significant difference between groups in recruitment to trial.

general and specific trials.[15,30] This problem may cause clinicians to select patients with whom they are better able to communicate.[16] Even among clinicians who have agreed to collaborate in a trial, the level of information given to patients and its timing (pre-randomisation or after) may vary.[43] Clinicians may also find it difficult to assess the level of information required by patients[29] and worry that information may be frightening[26] or even lead to increased morbidity and mortality.[29] Such attitudes to patient information may lead to sub-optimum patient recruitment; however, 'on-scene educational materials' may assist clinicians in providing information.[12] Trials of different recruitment strategies are discussed in more detail by Edwards and colleagues.[68]

Recommendation

- Potential participants 'must be adequately informed' about the study in accordance with the Declaration of Helsinki.[46] Information should be carefully prepared in such a way that patients of all levels of education have sufficient understanding of the study and what is being asked of them. Information should be given in oral as well as written form, with special provision for those whose first language is not English or who do not find reading easy. (Basis: logical argument.)

TABLE 18 *Information: clinician perspective*

Study (country)	Title	Subject area	Setting	Method	Results
Fisher, *et al.,* 1991[15] (USA)	Clinical trials in cancer therapy: efforts to improve patient enrolment by community oncologists	Cancer	Community	Survey of 75 participating clinicians	Problems of explaining clinical trials to patients: in general terms (17%) and specific terms (29%).
Foley & Moertal, 1991[12] (USA)	Improving accrual into cancer clinical trials	Cancer	Hospital	Questionnaire about 'possible impediments to patient entry' sent to participating clinicians: 209/334 responded	55% felt 'on-scene educational materials' on clinical trials should be available.
Penn & Steer, 1990[26] (UK)	Reasons for declining participation in a prospective randomised trial to determine the optimum mode of delivery of the preterm breech	Obstetrics	Hospital	Survey of 11 centres who refused to collaborate in RCT of Caesarean section for preterm breech deliveries; reasons for non-participation	One centre stated that the information provided would frighten many mothers.
Siminoff, *et al.,* 1989[30] (USA)	Doctor–patient communication about breast cancer adjuvant therapy	Cancer	Hospital	Observation and audio-recording of 100 patient–doctor interactions and surveys of participants	Physicians generally recommended one treatment. (i.e. clinical trials or standard therapy or non-standard therapy). Clinical trial not communicated as effectively as recommendations for standard therapies. Patients over-estimated chances of cure by 20%. Authors suggest that if data are given in numeric format, patients may make more accurate prediction of success.
Taylor, 1985[16] (USA)	The doctor's dilemma: physician participation in randomized clinical trials	Cancer	Hospital	Survey of 91 participating clinicians	38% of doctors not enrolling all patients cited trouble with informed consent, e.g. disclosing information to patients. Appeared that surgeons may be differentially selecting patients whom they consider would have no difficulty with informed consent document.
Taylor, *et al.,* 1987[29] (Canada)	Physician response to informed consent regulations for randomized clinical trials	Cancer	Hospital	Survey of 170 participating clinicians	41% said it is difficult for physicians to assess patients' desire for information; 77% did not believe main purpose of informed consent was to inform patients; 73% believed information led to increased patient morbidity and mortality.
Williams & Zwitter, 1994[43] (various)	Informed consent in European multicentre randomised clinical trials – are patients really informed?	Cancer	Hospital	60/88 clinicians replied to a questionnaire about practice of obtaining informed consent	88% of clinicians informed patients prior to randomisation, 12% did not even though protocol would probably have specified this. 62% stated that they always told patients they had been randomised; 28% said they sometimes did; 10% said they did not. 58% gave full information; 42% gave information on trial arm only.

Recommendation for research

- Further research is needed to clarify our understanding of the way in which patients decide whether to enter clinical trials and to identify the best methods of providing information.

Consent procedure

The issue of consent is intimately linked to the provision of information, therefore many of the studies reported in this section are the same as those referred to in the previous section on patient information. Fourteen papers reported specifically on the consent procedure as a barrier to patient recruitment: three RCTs and one non-randomised comparison of different consent procedures, eight surveys of patients and two surveys of the public (*Table 19*).

TABLE 19 *Consent procedure: patient perspective*

Study (country)	Title	Subject area	Setting	Method	Results
Aaronson, *et al.,* 1996[10] (The Netherlands)	Telephone-based nursing intervention improves the effectiveness of the informed consent process in cancer clinical trials	Cancer	Hospital	RCT of telephone-based nursing information plus standard informed consent procedure vs. standard procedure; 346 patients eligible, 180 took part. Comparison of randomised and non-randomised patients to discuss possible selection bias and general-isability; physicians also asked for reasons patient not entered	No evidence that informed consent procedure increased anxiety.
Bergmann, *et al.,* 1994[71] (France)	A randomised clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain	Cancer	Hospital	RCT to investigate effect of informed consent on analgesic activity of placebo and naproxen in cancer pain: 25 received treatment without information; 18 of 24 who received information participated	Informed consent significantly modified results of controlled clinical trials (higher degree of analgesic effect in consent group). Informed consent reduced patient recruitment.
Corbett, *et al.,* 1996[78] (UK)	Offering patients entry in clinical trials: preliminary study of the views of pro-spective participants	N/A	Community	Survey of members of the public, students and secretaries (n = 100) to find views on RCTs and randomisation	Majority of respondents in this study would prefer to sign a consent form (86%).
Dahan, *et al.,* 1986[72] (France)	Does informed consent influence therapeutic outcome? A clinical trial of the hypnotic activity of placebo in patients admitted to hospital	Insomnia	Hospital	Study to examine effect of placebo when given with or without informed consent to patients suffering with insomnia. Single blinded observer, blinded trial of patients paired according to sex, age and hospital environment. Randomisation assigned first patient in pair to control or informed consent group; 26 refused consent, 30 pairs were given placebo treatment	26/86 patients refused consent because: reluctant to participate (n = 12), afraid of side-effects (n = 10), wished to take well-known drug (n = 4). Better hypnotic activity in the control group. Side-effects: four reported in consent group, none in control group (p not significant). Women over 60 years old less likely to give consent ($p < 0.02$). Duration of study was increased by consent procedure (having to replace non-consenters).
Gallo, *et al.,* 1995[45] (Italy)	Informed versus randomised consent to clinical trials	N/A	Community	2053 healthy visitors to scientific exhibition enrolled in hypothetical trial and randomly assigned to groups: one-sided informed consent (n = 622), two-sided consent (n = 376), randomised consent to experimental treatment (n = 730), randomised consent to standard treatment (n = 307)	Agreement to participate depended on type of consent procedure: one-sided informed consent (16% refused to participate); two-sided (20% refused); randomised consent experimental (12% refused); randomised consent standard treatment (49% refused). More of those pre-randomised to standard treatment refused consent than those pre-randomised to experi-mental (52% vs. 13%) (authors note that labelling treatments as A and B may have given different results).

continued

TABLE 19 contd *Consent procedure: patient perspective*

Study (country)	Title	Subject area	Setting	Method	Results
Harth & Thong, 1995[61] (Australia)	Parental perceptions and attitudes about informed consent in clinical research involving children	Child health	Hospital	Questionnaire survey of 64 parents whose children had taken part in RCT of ketotifen for asthma; 62 took part	14.5% of parents of children entered to RCT considered consent not necessary as the physician knew best. Twice as many parents thought informed consent document was to protect doctor rather than their child. Half the parents thought RCT was of no or low risk to their child. Attitudinal and psychological factors should be taken into consideration in design of informed consent to ensure it is informed.
Lynoe, *et al.,* 1991[64] (Sweden)	Informed consent: study of quality of information given to participants in a clinical trial	Gynae-cology	Hospital	43/48 women traced 18 months after trial returned questionnaires	35/43 remembered giving consent, seven did not recall and one patient said she had not given consent.
Miller, *et al.,* 1994[88] (USA)	Comprehension and recall of the inform-ational content of the informed consent document: an evaluation of 168 patients in a controlled clinical trial	Medical	Community	Phone survey of 168 patients participating in trial comparing two analgesic drugs asking about recall and satis-faction with informed consent process	Authors suggest breaking consent form up into less complex sections and frequent presentation may improve recall.
Myers, *et al.,* 1987[79] (USA)	The consent form as a possible cause of side-effects	Cardio-vascular	Hospital	Opportunistic study of two different patient information wordings	Consent form may alter results of a study. Consent form and content of inform-ation provided to a patient should be considered as a possible source of bias.
Olver, *et al.,* 1995[81] (Australia)	The adequacy of consent forms for informing patients entering oncological clinical trials	Cancer	Hospital	100 consecutive patients from 18 clinical trials; study considered the impact of information and consent forms	Authors suggest that less than 50% of patients understand the purpose of the informed consent form.
Ross, *et al.,* 1993[66] (USA)	Reasons for entry into and understanding of HIV/AIDS clinical trials: a preliminary study	Various (HIV, cancer)	Medical research centre	Survey of 32 trial participants regarding reasons for participating in trials	Informed consent form required read-ability level of a person at university (Flesch and Fry tests). Greatest source of influence on decision to enter trial was doctor at AIDS clinic.
Schaeffer, *et al.,* 1996[42] (USA)	The impact of disease severity on the informed consent process in clinical research	Various (HIV, cancer)	Medical research centre	Patients from various trials (n = 73) and healthy volunteers (n = 52); questionnaires to assess knowledge of purpose and conduct of research, reading and compre-hension levels	The greater time spent explaining consent form, greater level of immediate and long-term retention. Suggest written consent forms are limiting and research required on differing ways of providing patient with information and not just supplying the consent form, i.e. oral information.
Tankanow, *et al.,* 1992[87] (USA)	Patients' perceived understanding of informed consent in investigational drug studies	Various medical	Hospital	Interviews with 98 patients in investigational drug studies within 72 hours of enrolling	Best perceived understanding was achieved by both nurse and doctor giving patient consent form and by giving patient sufficient time to read it. 83% of patients perceived that they understood most of consent form.
Tindall, *et al.,* 1994[73] (Australia)	Effects of two formats of informed consent on knowledge amongst persons with advanced HIV disease in a clinical trial of didanosine	HIV/AIDS	Hospital	113 HIV patients in controlled trial were randomised to receive information about trial in written format (n = 52) or written and verbal format (n = 61)	96% believed informed consent document was necessary although majority felt it was to protect doctor. No significant difference in knowledge found between written and written plus verbal consent procedures.

In a small study to examine the effect of placebo when given with or without informed consent to patients suffering from insomnia, Dahan and colleagues[72] found that informed consent increased the duration of the trial recruitment period because a proportion of patients refused to give consent. Consent modified the characteristics of the study group because women aged over 60 years were less likely to give informed consent. Therapeutic response was also affected in the informed consent group, with better hypnotic activity reported in the control group. Dahan and colleagues highlighted the problem of a 'consent effect' in applying the results of trials to the general population. Bergman and colleagues[71] and Myers and colleagues[79] suggested that the consent process may be a source of bias, and Gallo and colleagues[45] found that the type of consent procedure affects patient recruitment. However, the majority of the general public[78] and of HIV trial participants[73] are reported to prefer to sign a consent form if asked to take part in a clinical trial.

Recall of, and satisfaction with, the consent procedure may be variable[64,88] and Miller and colleagues[88] suggested breaking the consent form up into less complex sections to improve recall. Ross and colleagues commented on the readability level of the consent form.[66] Schaeffer and colleagues found that the more time spent explaining the consent form, the greater the recall.[42] Harth and colleagues highlighted the importance of attitudinal and psychological factors in the design of informed consent forms,[61] although Aaronson and colleagues found no evidence that the consent procedure increased anxiety.[10]

Only half of patients may understand the purpose of the informed consent form,[81] with some patients believing that the primary reason for the consent document was to protect the doctor.[61,73] However, the majority of patients may believe that they understood the consent form, with those who had the trial procedures explained by both nurse and doctor, and who had been given time to read the information provided, perceiving their understanding was greatest.[87]

Recommendation

- The consent procedure be carefully designed so that patients give consent freely, having had the study described to them as fully as they wish. The consent process should be as simple and straightforward as possible and patients should be given a written copy of the information and consent forms. Staff should receive training in seeking informed consent. (Basis: empirical evidence and logical argument.)

Recommendation for research

- Further research, particularly using an RCT design, is needed to clarify the best ways of achieving these aims.

Other factors

It has been suggested that specific groups of patients are less likely than others to participate in trials. However, the evidence is equivocal. For example, Simel and Feussner found that more severely ill patients were less likely to give consent than patients who were less severely ill;[50,83] however, Bowen and Barnes did not find such a relationship.[89] Schwartz and Fox found that a moderate level of disability was associated with a willingness to be randomised.[52] Gallo and colleagues found that patients with a worse outlook towards disease were less likely to consent[45] but Lerman and colleagues found that perception of risk from disease increased the likelihood of taking part.[90]

The evidence that better-educated patients are more likely to participate is also inconclusive. For example, participants in cardiovascular trials tended to be more highly educated than non-participants[60] and parents who would agree to their child taking part in a hypothetical trial had higher levels of education.[47] However, the parents of children who volunteered for a placebo-controlled RCT of ketotifen tended to be less well-educated.[61]

Age has also been suggested as a problem in trial recruitment. Younger adult patients may select to take part in trials,[82] while older patients may be under-represented.[10,12,14,39,91] Taylor reported that clinicians found it more difficult to approach patients at the extremes of age, wealth and intelligence to ask for consent to take part in trials.[19]

The doctor has been reported to be the greatest influence on the decision to enter a trial,[27,61,66,92] with patients who refused consent expressing less confidence in the doctor.[33] Some parents felt consent unnecessary as the physician knew best.[61] Another source of influence on the decision was an 'important person' (e.g. spouse, family member, close friend), with patients unlikely to take part in a study if those close to them were against the idea.[59,60]

Recommendation

- The trial should be designed to recruit a representative group of patients. This should be monitored carefully to ensure later generalisability of the study findings. (Basis: logical argument.)

Systematic reviews

Five systematic reviews were identified from the main database. These naturally form two groups. Three relate to the recruitment of specific types of people to RCTs: the inclusion of women in clinical trials of antihypertensive medications;[5] the recruitment of older people to clinical trials of arthritis;[6] and the recruitment of ethnic or racial minorities.[7] Another study considered the problems associated with RCTs in surgery[8] and the last was a general review of recruitment experience.[9] The main findings of these papers are presented individually.

Inclusion of women in clinical trials of antihypertensive medications

This review[5] consisted of 24 papers describing trials of antihypertensive medications, retrieved from one pharmacology journal over a 2-year period. Although limited in its scope, the paper highlights an important issue – in spite of the lack of women (as well as elderly and minority groups) in these trials, generalisation of the findings is often implied or recommended.

Recruitment of older people to clinical trials of arthritis

In this study,[6] 83 RCTs (73 papers; none of the references to individual trials given) employing non-steroidal anti-inflammatory drugs (NSAIDs) in the management of arthritis were identified between September 1987 and May 1990, and studied in detail for age-related information. The goals of the study were to identify the age and number of older people (aged 65 years or more), to document the way in which age information was presented and to suggest ways in which the results of clinical trials could be more informative and applicable to older people. The authors found that 52 of 83 trials probably included people aged 65 years and over but that few provided useful age information. It appears, however, that the more frail, 75–84 year-olds were under-represented. Thus the age group who are most commonly treated with NSAIDs are generally omitted from trials. Rochon and colleagues[6] suggested that standardised reporting of age data would make the interpretation of trials easier.

Recruitment of ethnic or racial minorities

Swanson and Ward[7] carried out a systematic review to provide a summary of key issues in recruiting diverse populations into clinical trials, particularly ethnic and racial minorities. At the outset the authors were particularly interested in recruitment to cancer trials but their search expanded to include studies from other areas. Several search strategies were used, including searching large databases (e.g. MEDLINE, CANCERLIT), expanding the search by examining the bibliographies of the articles identified from the first search, including recently published books and conference proceedings, and reviewing the contents of relevant journals. It is not clear from the paper how many articles were identified but 107 are listed (11 in common with the papers from the present review). Few studies were found that formally tested specific strategies for recruiting defined populations and so the review includes a more general set of commentaries and descriptive studies.

The review found that the majority of clinical trials in the USA included a limited segment of the population: middle-class, married, white males. The trials faced many problems in their efforts to recruit participants. Barriers to physician participation included the time and effort involved, concern for the patient or the patient–clinician relationship, concern about conduct of studies, concern about the conflict between caregiver/scientist roles. Barriers to patient participation included the time and inconvenience involved in participation, negative personal and family attitudes, and inadequate evidence of benefits from trial participation. Swanson and Ward[7] noted that the time and costs of recruitment are often underestimated and that training is essential for staff. Investigators should be actively involved in managing clinical trial recruitment, with recruitment coordinators and experienced investigators producing the highest participation rates.

In recruiting diverse populations, similar issues are significant; however, Swanson and Ward[7] suggested that an important first step for investigators is to develop a relevant definition of the subject populations from literature and research (using surveys and focus groups). A recruitment plan should be developed to take account of access and economic issues. The importance of piloting the recruitment plan was stressed.

Swanson and Ward[7] specifically addressed the problem of approaching diverse populations in the USA, where minority groups tend to be less

accessible than in the UK health system. Nevertheless, the issues raised are important and many aspects of the review may usefully be applied to trials mounted outside the USA.

Problems associated with RCTs in surgery

This systematic review[8] set out to determine the number of published RCTs over a 1-year period (1990) performed by surgeons and containing a surgical arm, and to assess the quality of the trials. The search identified 202 RCTs in general surgery. Only 47 (23%) contained a surgical arm and the principal author was not a surgeon in 20 (43%) of these. The value of the review is limited by the fact that the references to the individual trials are not listed in the report.

Reasons suggested for the lack of surgical RCTs are surgeons' perception of the difficulties caused by the admission of uncertainty, both for themselves and their patients, created by randomisation of therapies. There are also specific problems in surgery which make randomisation difficult despite an equipoise in respect of two or more forms of management. Randomisation of patients to markedly different types of treatment, such as surgical versus medical (e.g. fundoplication versus histamine receptor blockade in reflux oesophagitis) or operations of differing extent or complexity (e.g. mastectomy versus lumpectomy in breast cancer), may be difficult for patients to accept because they are not only unequal in magnitude but the surgery is irreversible. Thus, if the alternative management is finally found to be more effective, the option of crossover to the more effective therapy once the trial is completed is unlikely to be available.

Acceptance of case series by surgeons and/or lack of training in clinical research may be reflected in the small number of surgeons who are the principal authors of the RCTs identified and the study authors suggest that clinical research may lack the esteem of basic research. Few surgeons appeared to have managed to obtain funding for an RCT, perhaps because of lack of government support or lack of training of surgeons in health services research methods. The quality of the surgical trials comparing two operations was reported to be significantly lower than those comparing two medical therapies, suggesting room for improvement in surgical trials.

A general review of recruitment experience

Hunninghake and colleagues[9] carried out a literature review covering the period 1966–86, using search terms such as 'clinical trials', 'randomised trials', 'screening programmes' or 'follow-up studies', combined with terms such as 'recruitment', 'accrual' or 'screening'. Having identified approximately 900 references, the published bibliography consists of 85 references, only one of which is cited in this chapter.[17]

The available information indicated that recruitment has almost always been a much greater problem than was expected by the investigator. Hunninghake and colleagues[9] noted that accrual rates are often not clearly reported and that most studies report that the period of accrual was extended. Delays in recruitment have an impact on costs and workload throughout the trial. Other problems with recruitment outcomes include miscalculations of the number of eligible participants in the total population, variations in yield from different types of recruitment sources and the level of community awareness. The authors pointed out that they were unable to estimate how many studies were terminated or had inconclusive results caused by inadequate participant recruitment.

The authors considered that complete, comprehensive and standardised reporting of all recruitment information for studies of all sizes is needed. Designing and planning the recruitment process should be an integral part of study design and may take up to 3 years for a major multicentre trial. Ideally, recruitment planning should be conducted in harmony between national and local levels in multicentre studies. Data-based management of recruitment is necessary and should include a comprehensive data monitoring system with strong administrative support at both central and local levels throughout the period of a study. Recruitment goals and provisions for corrective action if recruitment lags should be included. Lack of experience in trials was suggested as a major cause of recruitment failure.[9]

The authors consider that socio-economic characteristics and attitudes of patients and physicians should be considered in recruitment efforts because they can affect patient participation during all phases of the study. There is minimal information on the appropriate messages to use in recruitment strategies for different diseases or age groups.

Although desirable information is lacking in many areas, Hunninghake and colleagues[9] stated that there is sufficient information to permit

the design of recruitment strategies in a more effective way than has previously been employed.

Discussion

Delays and problems with recruitment of clinicians and patients to clinical trials have a major impact on costs and workload throughout trials.[7,9] In this systematic review an attempt has been made to identify barriers to recruitment, with the intention of making recommendations for good practice in the future conduct of trials.

Limitations of the findings

The findings presented in this chapter should be read with some caution. In particular, several sources of potential bias are inherent in the material included in the review. First are the biases of publishing and reporting; only the results which were available to us are reviewed and, hence, only those which make a particular point. It is also important that a systematic approach was adopted to identifying study reports. It was not intended to be comprehensive. Our searches were limited to recent reports and the authors are aware of a substantial amount of earlier relevant literature.

Second, the majority of papers are from cancer research. Although many of the issues are common to other disease states, it is likely that cancer patients are more distressed than patients with less life-threatening conditions and that this will influence their decisions whether or not to take part in clinical trials. The pressures on clinicians who treat cancer patients are also likely to be different.

Similarly, most of the research reported is from the USA, where particular healthcare issues apply. For example, patient recruitment is likely to be more complex where patient populations are not identifiable, contrasting with the UK where age, sex, and disease registers are available. An additional issue is that of finance – patients in the UK can expect to receive their treatment as part of a trial free-of-charge, while in the USA the treatment may require to be covered by health insurance or some other means. The expectation of payment for trial participants appears more prevalent in the USA than elsewhere. The finding that most of the research is hospital-based is unsurprising, since most clinical research is carried out in hospitals.

Many of the studies reviewed here appear to be unsatisfactory. An example is the body of work by Taylor and colleagues,[16–20,28,29] which investigated clinician participation in clinical trials in a number of settings. The studies are based on a series of very similar surveys of clinicians, consisting of what appear to be leading statements and questions about the problems of participating in clinical trials. Because these studies were set up in this way, many problems were indeed reported. Taylor and colleagues are not alone in using this approach.

Some more rigorous evaluations are included in the review but many of these are not suitable for practical application. For example, the work by Llewellyn-Thomas and colleagues in evaluating various forms of patient information and consent[76,84,85] applied sophisticated research methodologies to study highly complex alternatives for giving information to potential participants. Although increasing understanding of patient decision making, the methods used in these studies are unlikely to be applicable to UK trials because the information and consent procedures would be incomprehensible to patients under normal clinical trial conditions.

The issue of patient information and consent feature highly in the review. This may not only be because these are particularly contentious issues but also because they lend themselves relatively easily to the conduct of publishable research (for example, RCTs of different designs of patient information and surveys of patient attitudes towards consent to participation in trials). Edwards and colleagues commented in more detail about different methods of inviting people to participate in trials.[68]

This review set out to investigate barriers to recruitment. It has, therefore, found evidence of barriers. Evidence that strategies to avoid these factors encourages participation is very limited, however. The factors which act as incentives may be different and it may be that stressing incentives such as altruism would be more beneficial. Ideally, strategies to improve recruitment should be formally tested through nested studies within ongoing trials.

It appears from the review that few successful trialists write an account of how they overcome recruitment barriers. It would now be appropriate to find out how they do this. While some of the measures are likely to be commonsense good management, including obsessive attention to detail, and choosing 'appropriate' strategies for each individual set of circumstances, other effective strategies are likely to be generalisable to a wide range of trials mounted in the NHS. These should be identified.

References

1. Miles MB, Huberman AM. Qualitative data analysis: a sourcebook of new methods. 2nd ed. Newbury Park: Sage Publications; 1994.
2. Agar MH. Qualitative research methods: volume 2. Speaking of ethnography. Newbury Park: Sage Publications; 1986.
3. Drummond NA. Quality of life for asthma patients: an existential–aesthetic theory [PhD thesis]. Aberdeen: University of Aberdeen; 1997.
4. Chalmers I, Altman DG. Systematic reviews. London: BMJ Publications; 1995.
5. Nunnelee JD. The inclusion of women in clinical trials of antihypertensive medications: review of twenty-four trials in one pharmacology journal. *J Vasc Nurs* 1995;**13**(2):41–9.
6. Rochon PA, Fortin PR, Dear KBG, Minaker KL, Chalmers TC. Reporting of age data in clinical trials of arthritis: deficiencies and solutions. *Arch Intern Med* 1993;**153**:243–8.
7. Swanson GM, Ward AJ. Recruiting minorities into clinical trials: toward a participant-friendly system. *J Natl Cancer Inst* 1995;**87**:1747–59.
8. Solomon MJ, Laxamana A, Devore L, McLeod RS. Randomized controlled trials in surgery [see comments]. *Surgery* 1994;**115**:707–12.
9. Hunninghake DB, Darby CA, Probstfield JL. Recruitment experience in clinical trials: literature summary and annotated bibliography [review]. *Control Clin Trials* 1987;**8**(4 suppl):6–30S.
10. Aaronson NK, VisserPol E, Leenhouts GHMW, Muller MJ, van der Schot ACM, van Dam FSAM, *et al.* Telephone-based nursing intervention improves the effectiveness of the informed consent process in cancer clinical trials. *J Clin Oncol* 1996;**14**:984–6.
11. Dickinson CJ. Clinical research in the NHS today. *J R Coll Physicians Lond* 1994;**28**:460–3.
12. Foley JF, Moertel CG. Improving accrual into cancer clinical trials. *J Cancer Educ* 1991;**6**:165–73.
13. Smyth JF, Mossman J, Hall R, Hepburn S, Pinkerton R, Richards M, *et al.* Conducting clinical research in the new NHS: the model of cancer. United Kingdom Coordinating Committee on Cancer Research. *BMJ* 1994;**309**:457–61.
14. Benson AB, Pregler JP, Bean JA, Rademaker AW, Eshler B, Anderson, K. Oncologists' reluctance to accrue patients onto clinical trials: an Illinois Cancer Center study. *J Clin Oncol* 1991;**9**:2067–75.
15. Fisher WB, Cohen SJ, Hammond MK, Turner S, Loehrer PJ. Clinical trials in cancer therapy: efforts to improve patient enrollment by community oncologists. *Med Pediatr Oncol* 1991;**19**:165–8.
16. Taylor KM. The doctor's dilemma: physician participation in randomized clinical trials. *Cancer Treat Rep* 1985;**69**:1095–100.
17. Taylor KM, Margolese RG, Soskolne CL. Physicians' reasons for not entering eligible patients in a randomized clinical trial of surgery for breast cancer. *N Engl J Med* 1984;**310**:1363–7.
18. Taylor KM, Kelner M. Interpreting physician participation in randomized clinical trials: the physician orientation profile. *J Health Soc Behav* 1987;**28**:389–400.
19. Taylor KM. Physician participation in a randomized clinical trial for ocular melanoma. *Ann Ophthalmol* 1992;**24**:337–44.
20. Taylor KM, Feldstein ML, Skeel RT, Pandya KJ, Ng P, Carbone PP. Fundamental dilemmas of the randomized clinical trial process: results of a survey of the 1737 Eastern Cooperative Oncology Group investigators. *J Clin Oncol* 1994;**12**:1796–805.
21. Tognoni G, Alli C, Avanzini F, Bettelli G, Colombo F, Corso R, *et al.* Randomised clinical trials in general practice: lessons from a failure. *BMJ* 1991;**303**: 969–71.
22. Winn RJ, Miransky J, Kerner JF, Kennelly L, Michaelson RA, Sturgeon SR. An evaluation of physician determinants in the referral of patients for cancer clinical trials in the community setting. *Prog Clin Biol Res* 1984;**156**:63–73.
23. Wadland WC, Hughes JR, Secker Walker RH, Bronson DL Fenwick J. Recruitment in a primary care trial on smoking cessation. *Fam Med* 1990;**22**:201–4.
24. Shea S, Bigger JT Jr, Campion J, Fleiss JL, Rolnitzky LM, Schron E, *et al.* Enrollment in clinical trials: institutional factors affecting enrollment in the cardiac arrhythmia suppression trial (CAST). *Control Clin Trials* 1992;**13**:466–86.
25. Morse EV, Simon PM, Besch CL, Walker J. Issues of recruitment, retention, and compliance in community-based clinical trials with traditionally underserved populations. *Appl Nurs Res* 1995;**8**:8–14.
26. Penn ZJ, Steer PJ. Reasons for declining participation in a prospective randomized trial to determine the optimum mode of delivery of the preterm breech. *Control Clin Trials* 1990;**11**:226–31.
27. Henzlova MJ, Blackburn BH, Bradley EJ, Rogers WJ. Patient perception of a long-term clinical trial: experience using a close-out questionnaire in the Studies of Left Ventricular Dysfunction (SOLVD) trial. *Control Clin Trials* 1994;**15**:284–93.
28. Taylor KM. Integrating conflicting professional roles: physician participation in randomized clinical trials. *Soc Sci Med* 1992;**35**:217–24.

29. Taylor KM, Shapiro M, Soskolne CL, Margolese RG. Physician response to informed consent regulations for randomized clinical trials. *Cancer* 1987;**60**: 1415–22.

30. Siminoff LA, Fetting JH, Abeloff MD. Doctor–patient communication about breast cancer adjuvant therapy. *J Clin Oncol* 1989;**7**:1192–200.

31. Chang RW, Falconer J, Stulberg SD, Arnold WJ, Dyer AR. Prerandomization: an alternative to classic randomization. The effects on recruitment in a controlled trial of arthroscopy for osteoarthrosis of the knee. *J Bone Joint Surg Am* 1990;**72**:1451–5.

32. Langley GR, Sutherland HJ, Wong S, Minkin S, Llewellyn-Thomas HA, Till JE. Why are (or are not) patients given the option to enter clinical trials? *Control Clin Trials* 1987;**8**:49–59.

33. Simes RJ, Tattersall MH, Coates AS, Raghavan D, Solomon HJ, Smart H. Randomised comparison of procedures for obtaining informed consent in clinical trials of treatment for cancer. *BMJ* 1986;**293**:1065–8.

34. Antman K, Amato D, Wood W, Carson J, Suit H, Proppe K, *et al.* Selection bias in clinical trials. *J Clin Oncol* 1985;**3**:1142–7.

35. Deber RB, Thompson GG. Variations in breast cancer treatment decisions and their impact in mounting trials. *Control Clin Trials* 1990;**11**:353–73.

36. Klein MC, Kaczorowski J, Robbins JM, Gauthier RJ, Jorgensen SH, Joshi AK. Physicians' beliefs and behaviour during a randomized controlled trial of episiotomy: consequences for women in their care. *Can Med Assoc J* 1995;**153**:769–79.

37. Bowen J, Hirsch S. Recruitment rates and factors affecting recruitment for a clinical trial of a putative anti-psychotic agent in the treatment of acute schizophrenia. *Hum Psychopharmacol* 1992;**7**:337–41.

38. Coombs DW, Dimick A, Bronstein JM, Potts LH, Bowens B. Conceptual and methodologic problems in the evaluation of a new burns treatment modality. *J Burn Care Rehabil* 1993;**14**:568–71.

39. Hunter CP, Frelick RW, Feldman AR, Bavier AR, Dunlap WH, Ford L, *et al.* Selection factors in clinical trials: results from the Community Clinical Oncology Program physician's patient log. *Cancer Treat Rep* 1987;**71**:559–65.

40. Jack WJ, Chetty U, Rodger A. Recruitment to a prospective breast conservation trial: why are so few patients randomised? *BMJ* 1990;**301**:83–5.

41. Plaisier PW, Berger MY, van der Hul RL, Nijs HG, den Toom R, Terpstra, OT, *et al.* Unexpected difficulties in randomizing patients in a surgical trial: a prospective study comparing extracorporeal shock wave lithotripsy with open cholecystectomy. *World J Surg* 1994;**18**:769–73.

42. Schaeffer MH, Krantz DS, Wichman A, Masur H, Reed E, Vinicky JK. The impact of disease severity on the informed consent process in clinical research. *Am J Med* 1996;**100**:261–8.

43. Williams CJ, Zwitter M. Informed consent in European multicentre randomised clinical trials – are patients really informed? *Eur J Cancer* 1994;**30A**:907–10.

44. Zelen M. Alternatives to classic randomized trials. *Surg Clin N Am* 1981;**61**:1425–32.

45. Gallo C, Perrone F, de Placido S, Giusti C. Informed versus randomised consent to clinical trials. *Lancet* 1995;**346**:1060–4.

46. British Medical Association, Ethics, Science and Information Division. Medical ethics today: its practice and philosophy. London: BMJ Publications; 1993.

47. Autret E, Dutertre JP, Barbier P, Jonville AP, Pierre F, Berger C. Parental opinions about biomedical research in children in Tours, France. *Dev Pharmacol Ther* 1993;**20**:64–71.

48. Harth SC, Thong YH. Sociodemographic and motivational characteristics of parents who volunteer their children for clinical research: a controlled study. *BMJ* 1990;**300**:1372–5.

49. Cunny KA, Miller HW. Participation in clinical drug studies: motivations and barriers. *Clin Ther* 1994;**16**:273–82.

50. Simel DL, Feussner JR. A randomized controlled trial comparing quantitative informed consent formats. *J Clin Epidemiol* 1991;**44**:771–7.

51. Moody LE, Gregory SJ, Bocanegra T, Vasey F. Search and research: factors influencing post-menopausal African-American women's participation in a clinical trial. *J Am Acad Nurse Pract* 1995;**7**:483–8.

52. Schwartz CE, Fox BH. Who says yes? Identifying selection biases in a psychosocial intervention study of multiple sclerosis. *Soc Sci Med* 1995;**40**:359–70.

53. Atwood JR, Haase J, Rees-McGee S, Blackwell G, Giordano L, Earnest D, Alberts D, *et al.* Reasons related to adherence in community-based field studies. *Patient Educ Counsel* 1992;**19**:251–9.

54. Wingerson D, Sullivan M, Dager S, Flick S, Dunner D, Roy-Byrne P. Personality traits and early discontinuation from clinical trials in anxious patients. *J Clin Psychopharmacol* 1993;**13**:194–7.

55. Sutherland HJ, Carlin K, Harper W, Martin LJ, Greenberg CV, Till JE, *et al.* A study of diet and breast cancer prevention in Canada: why healthy women participate in controlled trials. *Cancer Causes Control* 1993;**4**:521–8.

56. Morrison W. Trials and tribulations: patients' perspectives of the BETASERON study. *Axon* 1994;**16**(2):51–4.

57. Schwartz JS. Prostate cancer intervention versus observation trial: economic analysis in study design and conditions of uncertainty. Oxford: Oxford University Press, National Cancer Institute Monograph no 19; 1995. p.73–5.

58. Orr PR, Blackhurst DW, Hawkins BS. Patient and clinic factors predictive of missed visits and inactive status in a multicenter clinical trial. The Macular Photocoagulation Study Group. *Control Clin Trials* 1992;**13**:40–9.

59. Yeomans-Kinney A, Vernon SW, Frankowski RF, Weber DM, Bitsura JM, Vogel VG. Factors related to enrollment in the breast cancer prevention trial at a comprehensive cancer center during the first year of recruitment. *Cancer* 1995;**76**:46–56.

60. DeLuca SA, Korcuska LA, Oberstar BH, Rosenthal ML, Welsh PA, Topol, EJ. Are we promoting true informed consent in cardiovascular clinical trials? *J Cardiovasc Nurs* 1995;**9**(3):54–61.

61. Harth SC, Thong YH. Parental perceptions and attitudes about informed consent in clinical research involving children. *Soc Sci Med* 1995; **40**:1573–7.

62. Jensen AB, Madsen B, Andersen P, Rose C. Information for cancer patients entering a clinical trial – an evaluation of an information strategy. *Eur J Cancer* 1993;**29A**:2235–8.

63. Larson E, McGuire DB. Patient experiences with research in a tertiary care setting. *Nurs Res* 1990; **39**:168–71.

64. Lynoe N, Sandlund M, Dahlqvist G, Jacobsson L. Informed consent: study of quality of information given to participants in a clinical trial. *BMJ* 1991;**303**:610–13.

65. Newburg SM, Holland AE, Pearce LA. Motivation of subjects to participate in a research trial. *Appl Nurs Res* 1992;**5**:89–93.

66. Ross MW, Jeffords K, Gold J. Reasons for entry into and understanding of HIV/AIDS clinical trials: a preliminary study. *AIDS Care* 1993;**6**:77–82.

67. Slevin M, Mossman J, Bowling A, Leonard R, Steward W, Harper P, *et al.* Volunteers or victims: patients' views of randomised cancer clinical trials. *Br J Cancer* 1995;**71**:1270–4.

68. Edwards SJR, Lilford RJ, Brannholtz DA, Jackson JC, Hewison J, Thornton J. Ethical issues in the design and conduct of randomised controlled trials. *Health Technol Assess* 1998;**2**(15).

69. Ashcroft RE, Chadwick DW, Clark SRL, Edwards RHT, Frith L, Hutton JL. Implications of socio-cultural contexts for the ethics of clinical trials. *Health Technol Assess* 1997;**1**(9):19–45.

70. Bevan EG, Chee LC, McInnes GT. Patients' attitudes to participation in clinical trials. *Br J Clin Pharmacol* 1992;**34**:156–7P.

71. Bergmann JF, Chassany O, Gandiol J, Deblois P, Kanis JA, Segrestaa JM, *et al.* A randomized clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. *Clin Trials Meta-Anal* 1994;**29**:41–7.

72. Dahan R, Caulin C, Figea L, Kanis JA, Caulin F, Segrestaa JM. Does informed consent influence therapeutic outcome? A clinical trial of the hypnotic activity of placebo in patients admitted to hospital. *BMJ* 1986;**293**:363–4.

73. Tindall B, Forde S, Ross MW, Goldstein D, Barker S, Cooper DA. Effects of two formats of informed consent on knowledge amongst persons with advanced HIV disease in a clinical trial of didanosine. *Patient Educ Counsel* 1994;**24**:261–6.

74. Stone JM, Page FJ, Laidlaw CR, Cooper I. Selection of patients for randomised trials: a study based on the MACOP-B vs. CHOP in NHL study. *Aust N Z J Med* 1994;**24**:536–40.

75. Robertson NL. Clinical trial participation. Viewpoints from racial/ethnic groups. *Cancer* 1994;**74**(suppl):2687–91.

76. Llewellyn-Thomas HA, McGreal MJ, Thiel EC, Fine S, Erlichman C. Patients' willingness to enter clinical trials: measuring the association with perceived benefit and preference for decision participation. *Soc Sci Med* 1991;**32**:35–42.

77. Robinson SB, Ashley M, Haynes MA. Attitudes of African-Americans regarding prostate cancer clinical trials. *J Community Health* 1996;**21**:77–87.

78. Corbett F, Oldham J, Lilford R. Offering patients entry in clinical trials: preliminary study of the views of prospective participants. *J Med Ethics* 1996;**22**:227–31.

79. Myers MG, Cairns JA, Singer J. The consent form as a possible cause of side effects. *Clin Pharmacol Ther* 1987;**42**:250–3.

80. Maslin A. A survey of the opinions on 'informed consent' of women currently involved in clinical trials within a breast unit. *Eur J Cancer Care* (English ed) 1994;**3**:153–62.

81. Olver IN, Buchanan L, Laidlaw C, Poulton G. The adequacy of consent forms for informing patients entering oncological clinical trials. *Ann Oncol* 1995;**6**:867–70.

82. Fetting JH, Siminoff LA, Piantadosi S, Abeloff MD, Damron DJ, Sarsfield AM. Effect of patients' expectations of benefit with standard breast cancer adjuvant chemotherapy on participation in a randomized clinical trial: a clinical vignette study. *J Clin Oncol* 1990;**8**:1476–82.

83. Simel DL, Feussner JR. Suspended judgment. Clinical trials of informed consent. *Control Clin Trials* 1992;**13**:321–4.

84. Llewellyn-Thomas HA, McGreal MJ, Thiel EC. Cancer patients' decision making and trial-entry preferences: the effects of 'framing' information about short-term toxicity and long-term survival. *Med Decis Making* 1995;**15**:4–12.

85. Llewellyn-Thomas HA, Thiel EC, Sem FWC, Woermke DEH. Presenting clinical trial information: a comparison of methods. *Patient Educ Counsel* 1995;**25**:97–107.

86. Dunbar J, Cleary PA, Siebert C, Baker L, Brink S, Nathan DM. Implementation of a multicomponent process to obtain informed consent in the diabetes control and complications trial. *Control Clin Trials* 1989;**10**:83–96.

87. Tankanow RM, Sweet BV, Weiskopf JA. Patients' perceived understanding of informed consent in investigational drug studies. *Am J Hosp Pharm* 1992;**49**:633–5.

88. Miller C, Searight HR, Grable D, Schwartz R, Sowell C, Barbarash RA. Comprehension and recall of the informational content of the informed consent document: an evaluation of 168 patients in a controlled clinical trial. *J Clin Res Drug Dev* 1994;**8**:237–48.

89. Bowen JT, Barnes TRE. The clinical characteristics of schizophrenic patients consenting or not consenting to a placebo controlled trial of antipsychotic medication. *Hum Psychopharmacol* 1994;**9**:423–33.

90. Lerman C, Rimer BK, Daly M, Lustbader E, Sands C, Balshem A, *et al.* Recruiting high risk women into a breast cancer health promotion trial. *Cancer Epidemiol Biomarkers Prev* 1994;**3**:271–6.

91. Nystrom KM, Forman WB, Holdsworth MT. Clinical research trials: evaluation of a recruitment strategy for healthy elders. *J Clin Res Pharmacoepidemiol* 1992;**6**:293–301.

92. Wilcox M, Schroer S. The perspective of patients with vascular disease on participation in clinical trials. *J Vasc Nurs* 1994;**12**:112–16.

Chapter 5

Limiting factors related to the conduct and structure of RCTs

Methods

All articles from the main database rated as being highly relevant to the conduct of randomised trials (coded 3) were read in full, as were those that were coded with the following keywords: administration, biased losses, blinding, compliance, context failure, data management, description failure, drop-outs, experience, failure to complete, failure to stop, follow-up, location, monitoring, organisation, poor outcome assessment, preparation, quality assurance, randomisation, recruitment, run-in, selection bias. The abstracts of all articles rated as moderately relevant (coded 2) were also read, together with the titles of uncoded articles (e.g. letters, news items), and any that appeared relevant were read in full. Other articles were identified from the reference lists of articles that were read and from personal knowledge of the subject.

In total, 270 complete articles were read by a single reviewer (CC) and each article was summarised. The summaries were then read again to produce a list of relevant problems and potential solutions, and these were then categorised to produce the final report. In this way, it was hoped to concentrate on problems that had been identified in the literature rather than discuss the authors' preconceived ideas. Many of the original articles simply reported problems and possible solutions experienced in individual trials, and most review articles were based largely on anecdotal discussion. These have been referenced in the text. Particularly relevant studies that used reasonable methodology (RCTs, other well-designed evaluation studies, systematic reviews of such studies, or large surveys) are summarised in the tables.

Once the protocol of a trial has been designed, a number of problems can arise during the conduct of the trial that prevent the trial starting or limit its progress and quality.

Failure to start a planned trial

There are no good data on the number of planned RCTs that fail to start. One survey of 487 clinical studies receiving ethical committee approval (40% of which were clinical trials) showed that 21% never started,[1] mainly because of lack of funding, the principal investigator leaving the institution, logistic problems or expected recruitment problems. It may have been appropriate that some of these trials were never started (e.g. if they were poorly designed) but this was not assessed. However, it is also possible that some important and well-designed trials may fail to get funding or ethical committee approval because of ill-informed or biased refereeing. Chalmers has suggested that some refereeing of manuscripts submitted to journals may not be objective[2] and similar problems could exist with refereeing trial protocols, although this is difficult to document. Trials of new technologies may fail to start because clinical enthusiasm promotes widespread use of a new procedure rather than adequate evaluation,[3] particularly if the protocol was designed without clinical input.

Recommendations for research

1. Prospective surveys are needed of the number of planned RCTs that fail to start, the reason(s) for failure and the quality and importance of those trials that fail.
2. Guidelines should be developed for the evaluation of new technologies before they come into widespread clinical practice.

Poor recruitment

Frequency and causes

Recruitment is often a major problem in RCTs.[4] It has been estimated that half of all clinical studies fail to reach their planned size[1] but this may be an overestimate caused by publication bias,[5] that is, only articles documenting problems in recruitment get published. Only one survey of recruitment, in a cohort of 41 RCTs in the USA, was identified; this showed that only 34% of trials achieved or surpassed their planned sample size while a further 34% recruited less than 75% of their planned size.[6] Many of the reported reasons for difficulties in recruitment related to problems in the design (e.g. narrow inclusion criteria) or

to obstacles to participation, as discussed earlier. Problems in recruitment stemming from the conduct of the trial included: narrowing the inclusion criteria after completing a successful pilot study;[7] overestimating the pool of eligible patients or the recruitment rate,[8,9] often because of poor planning or inexperience; using too few recruitment strategies; and lack of active involvement of the principal investigator and other staff in recruitment.[9]

Effects of poor recruitment

Poor recruitment has multiple effects on the progress of a trial, ranging from the inconvenience and increased cost of prolonging a trial to stopping a trial early.[4,5,9] One survey showed that of 195 clinical trials started, 25 (13%) were abandoned early, mainly because of poor recruitment.[1] Failure to reach the planned sample size results in reduced statistical power[4,5,9] but may also have other effects on trial quality. In one trial there was evidence of trial data being falsified to make ineligible patients appear eligible.[10] Poor recruitment can also preoccupy and demoralise staff,[4,9,11] which, in turn, may further reduce recruitment.[12] A sudden surge in recruitment towards the end of a trial in order to meet planned targets can also cause major problems, in terms of increased workload for trial staff at a time when resources may be short.[9]

Solutions to poor recruitment

Many different solutions have been proposed to ensure adequate recruitment but none have been properly evaluated in RCTs to our knowledge (although most are based on common sense). Such solutions include: careful planning and piloting of recruitment; the use of multiple recruitment strategies; close monitoring of recruitment so that poor strategies (or centres) can be dropped and more effort concentrated on effective methods (or centres); the use of recruitment coordinators; setting recruitment goals; having contingency plans in case of poor recruitment; and providing training in recruitment.[4,9,13–16] Computerised databases can help to identify patients eligible for trials that clinicians have missed.[17] Involving those who will enter patients into a trial in the development of the protocol may often but not always help to ensure adequate recruitment.[18] Any one centre in a clinical trial may suffer from 'trial fatigue', that is, recruitment decreases with time. In multicentre trials, one solution is to stagger the introduction of new centres in waves in order to achieve steady recruitment over time,[19] while for single centre trials, recruitment drives can also be staggered.[20]

Some consider that multicentre trials should be restricted to academic centres with special expertise in the condition of interest because district general hospitals may produce poor quality data.[21] This policy obviously limits the number of centres eligible to participate and, hence, recruitment, particularly for trials dealing with diseases for which most patients are treated in district hospitals. It also limits the overall generalisability of the results.[21] However, three trial groups have reported that, with appropriate training and support, the quality of data from non-expert centres was at least as good as those from the expert centres.[21–23] Multicentre trials should therefore include as many centres that want to participate as feasible, regardless of expertise, in order to maximise recruitment.

Methods of recruitment

Recruitment methods need to vary according to the type of trial and depending on factors such as the prevalence of the condition of interest, the stage of disease, the setting of the trial and the size of the trial. Studies looking at asymptomatic disease in the community will need to use different methods to those assessing the treatment of diseases with such serious *sequelae* that they usually lead to hospital admission. Different populations may also respond to similar strategies differently.[14] It is therefore difficult to compare the effectiveness of different recruitment strategies between different trials.

No randomised comparisons of different strategies were identified and such trials would be difficult to perform. However, several community-based trials and one hospital-based trial were identified that attempted to evaluate the relative effectiveness and efficiency of different recruitment methods. Their results are summarised in *Table 20*. However, there are several problems which make it difficult to compare different studies. Effectiveness was assessed by calculating the contribution of each method to the final number of patients randomised but this will be heavily dependent on the priority each method was given in each trial. The efficiency or yield of each method was calculated in terms of the proportion of patients initially contacted by that method who went on to be randomised, or the proportion of patients screened who were randomised. These figures could be more comparable than those for effectiveness but, for some methods, the denominator for the number of people contacted was unknown, that is, it was impossible to know how many people heard of a trial through the media. In addition, although the number screened was

TABLE 20 *Comparison of recruitment strategies: from case studies of RCTs*

Study (country)	Subject area: clinical activity	Intervention	Contribution to total number randomised (%)		Yield, %*		Cost per patient randomised (currency of country of study)	
Community-based trials								
Anderson, *et al.*, 1995[20] (USA)	Diabetes: treatment of elderly diabetic patients *n* = 103	Intensive monitoring/ education	Mass mail Mass media Community-based Medical referral Other	3† 80 3 7 7	Mass mail Mass media Community-based Medical referral Other	9 (S) 27 (S) 12 (S) 20 (S) 40 (S)	Mass mail Mass media Community-based Medical referral Other	$300 $37 $900 $40 –
Bjornson-Benson, *et al.*, 1993[24] (USA)	Respiratory: secondary prevention of lung disease *n* = 189	Bronchodilator plus smoking intervention	Mass mail Mass media Work-site Community-based Other	16 40 12 19 13	Mass mail Mass media Work-site Community-based Other	20 (S) 12 (S) 8 (S) 10 (S) –	Mass mail Mass media Work-site Community-based Other	$135 $54 $82 $670 –
Bradford, 1987[25] (USA)	Cardiology: primary prevention of cardiac disease *n* = 3810	Lipid lowering agent	Mass mail Mass media Work-site Community-based Medical referral Other	6 11 24 14 4 41	Mass mail Mass media Work-site Community-based Medical referral Other	0.8 (S) 2.3 (S) 0.6 (S) 0.6 (S) 1.4 (S) 9.0 (S)	– – – – – –	
Hollis, *et al.*, 1995[35] (USA)	General medicine: blood pressure reduction *n* = 2382	Weight loss/ diet	Mass mail Mass media Community-based Other	73 11 12 4	– – – –		– – – –	
Johnson, *et al.*, 1995[15] (USA)	Cardiology: prevention of cardiac disease in menopause *n* = 875	Hormone treatment	Mass mail Mass media Community-based Medical referral Other	12 49 3 1 35	– – – – –		– – – – –	
King, *et al.*, 1994[26] (USA)	Cardiology: reduction of cardiac risk in middle age *n* = 357	Exercise programme	Mass media Mass telephone	40 60	Mass media Mass telephone	32 (S) 1 (C) 11 (S)	Mass media Mass telephone	$69 $168
Maurer, *et al.*, 1995[36] (USA)	Orthopaedics: treatment of arthritis in elderly *n* = 108	Exercise programme	Mass mail Mass media Community-based Medical referral Other	6† 36 18‡ 30 10	– – – – –		– – – – –	
Rudick, *et al.*, 1993[30] (Canada)	Respiratory: secondary prevention of lung disease *n* = 577	Bronchodilator plus smoking intervention	Mass mail Mass media Mass telephone Work-site Community-based Other	33 7 20 18 13 9	Mass mail Mass media Mass telephone Work-site Community-based Other	6 (S) 7 (S) 6 (S) 8 (S) 4 (S) –	– – – – –	
Silagy, *et al.*, 1991[27] (Australia)	Geriatric: primary prevention of vascular disease *n* = 400	Aspirin	–		Mass mail (ER) Mass mail (GPL) Mass media Community-based	6 (C) 65 (S) 18 (C) 47 (S) 25 (S) 3 (C) 83 (S)	Mass mail (ER) Mass mail (GPL) Mass media Community-based	$60 $48 $48 $48

* *Percentage yield = number randomised ÷ number contacted (C) or screened (S) x 100*
† *Only patients attending clinics were mailed*
‡ *Posters placed in clinics rather than in community*

ER, mailing list generated from electoral role; GPL, mailing list generated from GPs' lists

continued

TABLE 20 contd *Comparison of recruitment strategies: from case studies of RCTs*

Study (country)	Subject area: clinical activity	Intervention	Contribution to total number randomised (%)	Yield, %*	Cost per patient randomised (currency of country of study)
Hospital-based study					
Schoenberger, 1987[37] (USA)	Cardiology: secondary prevention of myocardial infarction *n* = 4524	Aspirin	Physician referral Mass media Hospital record review Other	15 34 40 11	

* *Percentage yield = number randomised ÷ number contacted (C) or screened (S) x 100*

always available, the definition of the point at which screening occurred varied between studies. In some trials, most patients were initially quickly screened (e.g. by telephone) to discuss the trial and assess major eligibility criteria before being invited for full screening.[24,25] In such trials, the efficiency of a strategy could be reported either as the number randomised from those who were pre-screened[25] or as the number randomised from those attending the first full screening session.[24] The latter would obviously give a higher efficiency as some patients had already been excluded. Another problem is that although strategies in *Table 20* have been classified into seven groups, there were important differences within each group between different trials. For example, workplace methods could vary between asking employers to put up posters in their offices through to meetings held by trialists with the workforce. Finally, the cost of each strategy was assessed in some trials in terms of both time and materials but costings were probably calculated differently in each study. Owing to these problems, conclusions need to be guarded and based on comparisons within the same study rather than across studies. In addition, since none of the studies were performed in the UK, their results may be of limited generalisability to trials performed within the NHS.

Despite these difficulties, some broad patterns do emerge from the data. Both the effectiveness and efficiency of recruitment strategies are important. Efficient strategies which randomise a high proportion of those screened will not recruit many participants if they can only be applied to a small population. In general, most patients in community-based trials were recruited using mass media, mail or telephone campaigns because they could reach a large population (see *Table 20*). In the USA, media campaigns were usually cheaper than telephone or mail campaigns. Community-based strategies were much more expensive. One study found that patients recruited by random-digit telephoning produced higher-risk profile patients than those who volunteered after mass mailing.[26] Mailing those on general practitioner (GP) lists with a letter signed by the GP seemed to be more efficient than mailing people selected from an electoral role.[27] Conducting an initial interview on the telephone is a useful way of establishing preliminary eligibility.[15,28,29] Very little data was found on different recruitment methods for hospital-based studies (*Table 20*). However, recruitment strategies in which eligible patients are actively sought (e.g. hospital note review) are likely to be better than those that rely on referrals from other physicians.

Recruitment strategies must be adapted to suit the type of trial and the type of population. Different centres in the same study may need to use different methods.[24,30] Trial staff need to be experienced, flexible and show initiative. Large numbers of potential participants must be screened to achieve adequate recruitment. One study of 41 trials showed that 75% of trials in which more than twice the planned number were screened recruited more than 75% of the required number of patients, compared with only 25% of trials in which less than twice the planned number were screened.[6]

Recruitment logs

Some authors have recommended that all trials should record the details of both randomised patients and non-randomised patients who were screened.[6] They argued that this allows the generalisability of the final results to be assessed – if only a small percentage of eligible patients were randomised the results may not be generalisable.[6,31,32] However, others have argued that this increases the cost and workload of trials unnecessarily, thus possibly hindering recruitment, and does not help in assessing generalisability.[33,34] Data on the reasons for patient exclusion may help to improve recruitment by identifying overly

restrictive inclusion criteria which can then be modified.[11,32] It may therefore be useful to collect these data during a pilot trial. For rare diseases, data on non-randomised patients can also help in studying the epidemiology of the disease and management policies outside the trial.[11]

Recommendations

1. Pilot studies are necessary before starting most RCTs to check that recruitment strategies are adequate. (Basis: anecdotal evidence.)
2. Multiple recruitment strategies should be used with the aim of screening at least twice the planned sample size. (Basis: single survey of RCTs.)
3. Recruitment should be closely monitored and there should be contingency plans if fewer patients are randomised than expected. (Basis: anecdotal evidence.)
4. Staggering recruitment may help to prevent falling recruitment over time. (Basis: anecdotal evidence.)
5. Multicentre trials should not be restricted to expert academic centres. (Basis: case studies of two cancer trials' networks.)
6. Large trials should not be required to register details about those who are not randomised but recruitment logs can be useful and should be encouraged whenever feasible, especially in pilot trials or trials in rare conditions. (Basis: anecdotal evidence.)

Recommendation for research

- Further assessment is required of different recruitment strategies for community- and hospital-based RCTs and of the role of specific recruitment coordinators. For example, in multicentre trials, different centres could be randomised to have or not have a recruitment drive, or to have or not have a recruitment coordinator.

Subversion of randomisation

Random allocation is crucial to minimise selection bias and it is therefore vital to ensure that this is adequately performed. In particular, the next treatment allocation must be concealed from the person entering patients into a trial.[38] Unfortunately, some people succeed in breaking the allocation code before deciding whether to randomise a patient; anecdotal evidence suggests that this is not rare.[39]

Steps must therefore be taken to safeguard the concealment of allocation. Random number lists which are read by the doctor entering the patients are clearly useless; sealed envelopes may also not be secure (especially if they are not opaque) and so should also be avoided if possible.[39] The assignment code must be kept locked in a safe place and the person drawing it up should have nothing to do with the recruitment or assessment of patients.[39] If block randomisation is used, the size of the blocks should be varied randomly. A system of audit should be set up to check that allocation was performed adequately.[39] In multicentre trials, randomisation by contacting a central office is probably the best way to conceal allocation. This can be done by telephone but computerised fax systems and the Internet can also work well and avoid the need for 24-hour telephonists.[40,41]

Recommendation

- All trials within the UK should have access to a central randomisation service. Centres offering this service should be established around the country, with some offering 24-hour access. (Basis: anecdotal evidence.)

Although this recommendation is based on limited evidence, concealment of randomisation is crucial to prevent bias and thus efforts must be made to ensure randomisation is performed properly.

Recommendation for research

- Further surveys of trialists are needed to document how frequently randomisation is subverted.

Early stopping of trials

At least 10% of clinical trials may stop early for a variety of reasons.[1] A trial may have to stop early because of logistic problems, such as the principal investigator leaving the institution, withdrawal of funding, faults with the design, or irretrievably poor recruitment.[1,42,43] It may also be necessary to stop a trial because of the results of other, similar trials, or because the intervention being tested becomes superseded or withdrawn.[42] Interim analysis of the data can also provide good reasons for terminating a trial early by demonstrating unequivocal evidence of benefit or harm, unexpected severe side-effects, or that the trial has little chance of showing a clinically important benefit.[42–44] Such interim analyses should generally be planned, be based on good quality data and should use appropriate statistical techniques and stopping guidelines to avoid inappropriate conclusions (see also chapter 3).[42–44]

However, the decision to stop a trial early on the basis of an interim analysis is a complex one and should be based on clinical and ethical considerations as well as statistical ones.[42–45] For example, it would be unwise to stop a trial early simply because it had little chance of showing a statistically significant result, since lack of statistical significance does not imply lack of clinical significance.[44] Poorly conducted analyses can cause problems. Trials have been stopped inappropriately because of the use of spurious outcome measures, continuous data monitoring without appropriate statistical tests and inconclusive data.[42,44] In addition, interim analyses are often carried out on only a proportion of patients in the trial because some patients may not have reached the time of assessment. The final results can therefore be significantly different from those of an interim analysis.[46] Decisions about early termination should be delayed if current follow-up is not available for almost all randomised patients.[43]

Interim analyses can also cause problems in the conduct of a trial. They generate a significant workload for trial staff, particularly if unplanned, and their results can reduce recruitment dramatically if they become known to the investigators.[46] There are several problems associated with stopping a trial early.[44] The trial may be too small to give credible or precise results or the result may be too extreme to be realistic, especially if the trial was stopped on a 'random high' which increases the risk of a false-positive result for either benefit or harm. Early termination also attracts publicity and puts pressure on trialists to make incomplete or unnecessarily hasty recommendations.[42,44]

Given all the potential pitfalls of interim analyses and the problems of early termination of trials, it is surprising that one study found that few collaborative trial groups in Europe had drawn up a formal policy for interim analyses, although most recognised the need to do so.[45]

Recommendations

1. Interim analyses should be carefully planned. (Basis: anecdotal and empirical evidence.)
2. Trials should not be stopped on statistical grounds alone. (Basis: anecdotal evidence.)

Recommendation for research

- Studies are needed of the number of RCTs that terminate early, the reasons they do so, and whether the decision to stop was appropriate.

Compliance

Problems with compliance

Poor compliance with a trial protocol either by the investigator or the participant can have important effects on the overall result, although the effects vary depending on the disease and intervention under study.[47] Investigators may recruit ineligible patients, prescribe prohibited medication, fail to prescribe or carry out the allocated treatment adequately, or transfer the patient from one arm of the trial to the other (so-called crossovers).[48,49] Participants may: take too much or too little medication, take it at inappropriate intervals, fail to attend follow-up appointments, or fail to undergo certain tests.[48,50–52]

The reasons for non-compliance with a trial protocol are complex. Investigators may have insufficient time to comply with protocols, or lack motivation.[53] Their desire to give their patients the best possible treatment may mean that they do not comply with a treatment they consider is harmful or they want to give everyone a treatment that they consider is beneficial.[49,53] These are particular problems in unblinded trials.[48] This highlights the need for investigators to be truly uncertain about which treatment is best for their patients before they randomise them.[49] Even then, they may develop preferences during the trial based on the patients' experiences, for example, whether they develop side-effects.

Qualitative studies have shown that multiple factors influence participant compliance including disease severity, disease chronicity and the patient's view of their susceptibility to a poor outcome.[48,54] Barriers to accepting medical advice also include the financial cost of treatment, the risks and side-effects of treatment, the inconvenience in taking treatment, especially if it requires lifestyle changes, and a poor doctor–patient relationship.[55,56] Patients with strong support from their family or friends tend to be more compliant while those with impulsive or novelty-seeking personality traits may be less compliant.[48,57,58]

Effect of poor compliance on RCTs

Poor compliance reduces the power of a trial to identify a real treatment effect resulting in the need for increased sample sizes or an increased risk of a false-negative result (type II error).[56,59,60] A 50% compliance rate can increase five-fold the sample size required to show an expected treatment effect.[50] Differential compliance between two active treatments may also result in false-positive results, that is, the more effective treatment may

appear less effective if its compliance is poorer.[60] Thus it is important to measure compliance in RCTs, to assess both the risk of false-positive or -negative results and the acceptability of the treatment to patients. It may also allow side-effects associated with unusual dosing patterns to be identified.[60]

However, compliance appears rarely to be assessed in clinical trials.[47] This may be because of difficulties in defining and measuring it.[50,52] Classifying participants as compliers, partial compliers, overusers, erratic users, partial drop-outs and drop-outs has been suggested but without clear definitions being given.[60] It can be particularly difficult to define compliance with some non-pharmacological treatments. For example, how do you assess the adequacy of surgical procedures, or determine the difference between short and long interviews?[22,53] Various measures of compliance with pharmacological treatments exist but none are perfect. Indirect measures such as patient interviews, prescription monitoring, diaries, pill counts and electronic tagging of containers do not prove that medication has been ingested.[50,52,60,61] Direct measures, such as measuring drug or metabolite levels or an inert marker in the blood or urine, assess ingestion but not timing and not all drugs can be assayed in blood. Interviews, pill counts and diaries have been shown to overestimate compliance compared with electronic monitoring or assessment of blood levels[50,51,60,62] but electronic monitors and blood markers are more expensive and may not be practical.[50] Pill counting and blood assays of drugs with short half-lives are also subject to so-called 'white coat' compliance, that is, patients can appear compliant so long as they take the medication shortly before seeing the doctor.[50,60,61]

Methods to improve compliance in RCTs

The effort that should be expended to improve compliance of doctors and patients depends on the type of clinical trial. Explanatory or efficacy trials designed to assess whether the treatment has effect if taken or given as prescribed need to maximise compliance, whereas pragmatic or clinical management trials are interested in assessing whether the treatment is effective for the average patient under circumstances that reflect real life.[50,60] A number of anecdotal solutions have been proposed to improve compliance both with the intervention and with follow-up attendance including: education about the trial; assessing compliance during a run-in phase before randomisation so that obvious non-compliers are not randomised; simplifying treatment and follow-up regimens; and maintaining communication with participants by newsletters or telephone calls.[8,55,63–65]

There have been many RCTs in which methods of improving patient compliance with medication are assessed, although none specifically considered improving compliance in trials. A recent Cochrane systematic review summarised the results of 13 such trials and concluded that none of the methods were very effective and most required significant effort and resources (*Table 21*). However, because the review was aimed at improving clinical practice rather than RCTs, 19 trials were excluded because they only recorded details of compliance (adherence) and not clinical outcomes. These trials may provide useful information about improving compliance in RCTs but it was beyond the scope of this project to review these articles. Interested readers are referred to the Cochrane review.[66] Our search strategy identified two further RCTs of compliance but both were seriously flawed[67,68] (*Table 21*). No RCTs were identified that assessed ways of improving compliance with non-pharmacological interventions or ways of improving doctors' compliance with protocols. Several trials examining ways of improving clinic attendance in routine clinical practice (rather than in RCTs) have been reviewed[69] (*Table 21*).

Assessing the effect of compliance on the results of a trial

Patients with poor compliance are often excluded from the analysis of RCTs, especially those with an explanatory viewpoint.[59] Unfortunately, this approach is seriously flawed. First, adequate compliance can be difficult to define.[70] Second, excluding patients after randomisation may bias the results if there is a relationship between non-compliance and the outcome of interest (as is likely), particularly if the relationship differs between treatment groups.[50] For example, patients in the treatment group may be non-compliant because treatment is failing (i.e. they have a poor prognosis), while those in the control group may be non-compliant because they are doing well (i.e. they have a good prognosis). It is also not possible to compare the results of compliers and non-compliers from the same treatment group because they have been shown to differ in many ways.[57,59] One sensible way of assessing the effects of compliance might be to measure it before randomisation during a run-in period. Patients with poor compliance could then be excluded from explanatory trials, while in pragmatic trials of sufficient size pre-randomisation compliance could be used to define subgroups for analysis.[33]

TABLE 21 *Assessment of interventions designed to improve compliance*

Study (country)	Method	Subject area (number of trials)	Clinical activity	Setting (size)	Intervention	Results	Notes
Haynes, *et al.*, 1996[66] (various)	Systematic review of RCTs to improve compliance with medication in clinical practice: trials assessed both adherence and clinical outcome; follow-up of > 80% patients required	Hypertension (6) Schizophrenia (3) Infections (2) Asthma (1) Epilepsy (1)	Short- and long-term treatment	All trials based in community	Several interventions designed to improve compliance	Small improvements in compliance with: instructions/counselling; programmed learning on disease; increased convenience of care (e.g. simpler regimens); increased involvement of patients in their care; reminders; rewards.	Trials were small, with many methodological problems (poorly described randomisation, lack of blinding, poor assessment of compliance). No unpublished trials found. Interventions tested were complex.
Putnam, *et al.*, 1994[67] (USA)	RCT	Variety of infections	10-day course of antibiotics	University students (110)	Interview to highlight advantages of compliance, ways to achieve it and form contract	Improved mean adherence (pill count) in treatment group (92% vs. 81%).	Randomisation not described. 50 patients lost to follow-up.
Raynor, *et al.*, 1993[68] (UK)	RCT	Medical patients taking 2–6 regular medications	Various	Patients discharged from wards (210)	Standard nursing advice vs. nursing advice plus reminder chart vs. structured pharmacist's advice vs. pharmacist's advice plus reminder chart	Reminder chart increased good compliance (83% vs. 63%) on pill counts. Pharmacist advice no different from nursing advice.	Randomisation not described; 13 patients lost to follow-up; unblinded assessment; short follow-up (10 days).
Macharia, *et al.*, 1992[63] (USA, UK, Canada)	Systematic review of RCTs to improve keeping of appointments in clinical practice	Medical/ psychiatric (15) Cancer (8)	Initial diagnostic appointment (15); cancer screening (8); follow-up appointment (1)	Hospital	Various interventions designed to improve attendance	Improved attendance with: written (OR 2.0) or phoned (OR 2.8) prompts; provision of information about appointment (OR 2.9); contracts with patient (OR 1.9); follow-up clerk sending new appointment after missed appointment.	Randomisation poorly described; unblinded assessment; only three trials with long follow-up.

Recommendations

1. The possibility of sub-optimal compliance should be taken into account in planning sample sizes of pragmatic trials. (Basis: empirical evidence.)
2. Compliance with medication can be improved by patient education, simplifying regimens, and using reminder charts. (Basis: systematic review of small poor quality RCTs.)
3. Attendance at follow-up can be improved: by mailed or telephoned reminders (Basis: systematic review of seven small RCTs.); by providing information about the appointment (Basis: systematic review of three small RCTs.); or contracting with patients. (Basis: systematic review of two small RCTs.)
4. In explanatory trials, poor compliers should be identified before randomisation so that they can be excluded. (Basis: anecdotal evidence.)
5. Measurement of compliance during a trial can help explain a trial's result when a treatment appears not to work but complicated assessments may increase the workload unnecessarily. (Basis: anecdotal evidence.)

Recommendations for research

1. All existing RCTs of methods to improve compliance should be reviewed by extending the existing Cochrane systematic review.
2. Randomised trials to assess methods of improving health professionals' and patients' compliance with RCT protocols are required.
3. The most feasible methods of assessing compliance in RCTs, together with the impact of measuring compliance on the conduct (especially recruitment) and results of RCTs, need to be evaluated.

Drop-outs and losses to follow-up

The results of a trial can be invalidated if a large proportion of those randomised are excluded from the analysis because of protocol violation or because they drop out or are lost to follow-up. However, surveys of the reporting of trial results have shown that many trials do exclude a significant number of randomised patients from the analysis (*Table 22*).[71–77] Qualitative studies have identified a number of reasons why people drop out of trials, particularly psychosocial problems (e.g. marital problems), health-related problems, logistics problems (travel costs), and a poor relationship with the clinic.[78,79] Drop-outs can be identified by close monitoring. One trial retrieved 35 out of 36 drop-outs by employing an intensive counselling programme but this is probably not feasible for many trials.[79] It is obviously important to collect as many details as possible about the patient to help with follow-up (e.g. address, telephone number, GP, next of kin), and tagging records with national death registers can at least help to determine whether a patient has died.[80] Patients who withdraw from treatment or who violate the protocol should still be followed-up to allow an ITT analysis.

Recommendation

- Follow-up should be attempted on all patients who were randomised. (Basis: empirical evidence.)

Blinding

During a trial any of the following can be blinded to treatment allocation to reduce the risk of bias: the patient/participant, the health-care professional(s) performing the intervention, the person making the outcome assessment. It is not always possible to blind the patient or health professional (e.g. in surgical trials) but it is usually possible to blind the outcome assessor. Many trials report that they were double-blind, implying that both patients and health professionals were unaware of treatment allocation. However, it has been shown in several so-called double-blind trials (mainly involving psychotropic agents) that many doctors and patients (up to 100% in some studies) became unblinded during the trial because of treatment effects, side-effects or the lack of these.[81–85] This may be more likely to occur if patients had been exposed to a drug previously[83] or if an active agent was compared with an inactive placebo,[75] although some doctors could also distinguish between active agents.[82] Unblinding of both the patient or doctor could lead to bias; the doctor may monitor or treat patients in one group slightly differently, or the patient's knowledge of treatment may affect the outcome, particularly if the outcome is subjective. Indirect evidence of such bias was found in RCTs comparing new psychotropic agents with both older agents and placebo controls. The treatment effect of the older agents in these trials was less than seen in earlier trials in which the older agent was only compared with placebo. It was argued that this may be because if blinding had been broken in trials comparing both new and old agents with placebo, there was more bias in favour of the new treatment.[81] Kirsch and Rosadino[86] provided more direct evidence by showing that knowledge of treatment could affect subjective outcome measures in volunteers (*Table 23*). Unblinding the outcome assessor can also lead to bias if the outcome has a subjective element to it (*Table 23*).

It is therefore important to assess the effectiveness of blinding of the doctor, patient and outcome assessor during a trial, for example, by asking them to guess the treatment that the patient was taking.

TABLE 22 *Exclusions or losses following randomisation*

Study	Subject	Number and type of trials	Exclusions/losses after randomisation
Kleijnen, *et al.*, 1991[71]	Asthma	13 controlled clinical trials	8% reported > 20% patients lost
Liberati, *et al.*, 1986[72]	Breast cancer	63 RCTs	5% reported > 15% patients excluded
Lionetto, *et al.*, 1995[73]	Pancreatic cancer	27 controlled clinical trials	37% reported > 15% patients excluded
Rosetti, *et al.*, 1993[74]	Glaucoma	102 RCTs	34% reported > 15% patients lost
Schulz, *et al.*, 1996[75]	Obstetrics	110 RCTs	26% reported > 10% patients excluded
Solomon, *et al.*, 1994[76]	Surgery	202 RCTs	67% reported > 15% patients lost
Sonis & Joines, 1994[77]	Primary care	53 RCTs	73% reported > 15% patients excluded

TABLE 23 *Evaluations of the effect of blinding on the outcome of randomised trials*

Study (country)	Method	Subject area	Clinical activity	Setting	Intervention	Results
Carroll, *et al.*, 1994[87] (USA)	Case study of RCT; blinded outcome assessor guessed treatment assignment	Psychiatry	Treatment of cocaine dependence	Outpatients *n* = 73	Factorial design of psychotherapy and medication vs. control.	Significant interaction between accuracy of guess and effect of medication for subjective but not objective outcomes: those correctly guessed as placebo were rated as more severe than those correctly identified as active.
Kirsch & Rosidino, 1991[86] (USA)	RCT to assess effect of patient blinding on outcome measurement	Healthy volunteers	Effect of caffeine on physiological responses and mood	University psychology department *n* = 100	One group told they were drinking caffeinated coffee, another decaffeinated coffee, another that it could be either. In fact, all students were randomised between caffeine and no caffeine.	Significant interaction between what students were told and what they actually received for mood: only those told they were receiving caffeine and were receiving it reported increases in tension. No interactions for pulse or blood pressure.
Noseworthy, *et al.*, 1994[88] (Canada)	Case study of RCT; ratings of blinded assessor compared with those of unblinded assessor	Neurology	Treatment of multiple sclerosis	Hospital *n* = 165	Cytotoxic therapy ± plasma exchange vs. placebo.	Unblinded assessments suggested benefit from combination of cytotoxic therapy and plasma exchange; this was not confirmed in blinded assessments.

Demonstrating that blinding has been effective is reassuring. It is less clear what should be done if blinding was shown to be ineffective. Although there is the potential for bias in such cases, it does not follow that bias will definitely be present. It has been suggested that the level of unblinding be measured and the result corrected for this but it is unclear how this could be done.[81] It would be unwise to compare the results from participants in whom blinding was adequately maintained following randomisation with those in whom it was not because these two groups may differ in important ways.[32] For example, those on active treatment may guess correctly because they are doing well (i.e. they have a good prognosis) while those on inactive treatment may guess correctly because they are not responding (i.e. they have a poorer prognosis). One possibility would be to define the adequacy of blinding during a run-in period and then randomise patients. Patients with adequate and inadequate blinding would then be randomly allocated to the different treatment groups and sensible subgroup analysis may then be possible.

Several ways of improving blinding of patients and doctors during pharmaceutical trials have been suggested, including the use of placebos with similar side-effects to active drugs and blinding the investigators to both the trial design and the drug, but these are rarely feasible.[81] It may sometimes be feasible to blind outcome assessors completely to the nature of an intervention.

Recommendation

- Outcome assessments should be blinded. (Basis: two case studies.)

Recommendation for research

- Further evaluation of the impact of measuring the success of blinding on the conduct and results of RCTs is required.

Outcome measurement

Outcomes should be reliable, valid, clinically relevant, sensitive to important changes and assessed at relevant times (see chapter 3). They also need to be feasible, that is, it must be practical to collect good quality data for each outcome for every patient randomised (or nearly every patient). Trialists should only collect a small number of outcomes and should avoid outcomes that are too complicated or irrelevant.[7] It may be appropriate to collect data on surrogate outcomes (e.g. biochemical or radiological tests) in trials that are performed early in the development of a new intervention (Phase II trials) to assess whether

there is any evidence of efficacy but these trials should also measure clinically important outcomes. It is also important that those who collect the outcome data are adequately trained, otherwise they may collect useless information.[7] For serial measures of subjective states (e.g. quality of life), it may be appropriate to show patients their answers to previous assessments. With some measures, this may provide more valid results because it allows patients to calibrate their responses with respect to previous responses.[89]

Recommendations

1. Trialists should collect a small number of relevant and feasible outcome measures. (Basis: anecdotal evidence.)
2. All trials should measure clinically important variables. (Basis: anecdotal evidence.)
3. Those assessing outcomes should be properly trained. (Basis: anecdotal evidence.)

Recommendation for research

- Further study is required of the best way in which to assess repeated subjective outcome measures.

Data entry

Two studies evaluated single- versus double-entry of data during RCTs (*Table 24*). Double-entry does reduce the error rate but many of the errors probably do not affect data analysis.[90,91] Single-entry, with consistency checks, may be an appropriate alternative for small trials.

Recommendation

- Double-entry of data is preferable but single-entry with consistency checks may be appropriate. (Basis: one RCT and one non-random comparison).

Data monitoring and quality assurance

It is extremely important to ensure that the data collected and processed during a trial are accurate.[92,93] Good quality starts by having motivated, trustworthy investigators.[92] The Good Clinical Practice guidelines also emphasise the need for quality assurance through appropriate

TABLE 24 *Evaluation of double versus single data entry in randomised trials*

Study (country)	Method	Subject area	Clinical activity	Setting	Intervention	Results
Gibson, *et al.*, 1994[90] (UK)	Non-random comparison of single vs. double entry	Cancer	Radiotherapy	Hospital	Data from 44 patients (16,277 fields) entered by one person who was unaware data would be checked and then re-entered on same system by second person. Data files compared and differences examined.	122 differences between first and second entry (75 per 10,000); 56 were non-trivial, 31 due to errors on first entry (19 per 10,000) and 25 to errors on second. Of the 31 first entry errors, 25 were not found by subsequent consistency checks (15 per 10,000); of these 25, only four would have changed the analysis. Note: small non-random study with only two people entering data. Highlights that errors can also occur with second data entry.
Reynolds-Haertle & McBride, 1992[91] (USA)	Crossover RCT	Cardiology	Pharmacology	Hospital	26 centres allocated to 2 weeks of single or double data entry (at 23 centres same person entered); staff knew about trial. Data submitted for 4/25 forms (42,278 fields) were compared with original forms by two people blinded to nature of entry. Only errors that should have been prevented by double entry were counted.	Single entry resulted in error rate of 22/10,000 compared with 15/10,000 for double entry ($p < 0.001$); 33% of errors were on a single question; 50% of staff made no errors and two staff had very high error rates. Double entry increased average time for entering an item by only small amount (4.5 vs. 3.3 seconds).

quality controls and the audit of case records,[92,93] and these measures can identify fraudulent or careless practice.[94] Ideally, quality requirements should be defined and then the processes and the data should be monitored closely (e.g. by monitoring teams and site visits) to check that they meet these requirements.[92,93] There are a number of problems associated with quality control. Site visits are made more difficult by changes in hospital personnel.[92,93] More importantly, high levels of quality assurance can increase administrative costs and the investigators' workload dramatically and, hence, may not always be feasible, particularly in very large mega-trials.[95] One possible compromise would be to audit random samples of rather than all case records.[94]

Data monitoring can help to identify whether some unnecessary items are being collected and so help to simplify a trial.[11] The accuracy of the diagnosis and patient eligibility should be checked along with baseline and outcome data and the adequacy of follow-up.[11] Site monitoring should also assess the quality of the storage and dispensing of trial medication and the adequacy of blinding of the packaging.[96]

Recommendation

- There should be explicit mechanisms to monitor the quality of trial procedures and data. (Basis: anecdotal evidence.)

Recommendation for research

- Different intensities of quality control should be compared to see how they affect the conduct and quality of the trial. For example, centres in multicentre trials could be randomised to different levels of quality control.

Other problems with the conduct of trials

Problems can arise in trials that require collaboration between a number of different health professionals and agencies, although these can usually be solved by good communication.[4,97] Lack of experience on behalf of the trialists can also cause problems.[7] It would probably be helpful for experienced trialists to act as advisers to those with less experience. Staff turnover, particularly among junior hospital staff and nurses, can also hinder the progress of trials.[98] Trials in the community have specific problems compared with hospital-based trials.[99] There is no 24-hour care, communication between different health professionals and between hospital and the community can be difficult, more travelling is involved, and the provision of materials needs greater organisation. Multicentre trials also have specific problems such as the need for regional or national collaborators' meetings, site visits and translation of materials into appropriate languages.[100,101] Every trial needs an active and committed principal investigator to provide the necessary leadership and to see the project through to its completion. This is particularly necessary in multicentre trials in which many different people are involved.[101]

Investigators can be involved in several trials simultaneously, which could lead to conflicts of interest,[102] although not if the inclusion criteria for each trial were clearly different. Commercially sponsored trials, in which investigators receive a large honorarium for each patient entered, are likely to be given preference over unsponsored trials and this may not be in the public interest.[94] Some authors have suggested that investigators should only receive reimbursement for the extra work involved in entering patients into the trial and that no additional honoraria should be given.[94,102] However, calculating the cost of this extra work may be difficult. This is an area in which guidelines are required.

Recommendations

1. Inexperienced trialists should be supported by experienced trialists. (Basis: anecdotal evidence.)
2. Good communication is essential especially in multicentre trials. (Basis: anecdotal evidence.)

Recommendations for research

1. The impact and feasibility of having experienced trialists helping inexperienced trialists should be assessed.
2. Research is required into the effect of offering financial rewards to investigators for participating in trials and guidelines on the level of reimbursement in commercially sponsored trials are needed.

Limiting factors related to the structure of RCTs

RCTs probably need to have a good organisational and administrative base to succeed but little has been published about this. Large multicentre trials need a steering committee with overall responsibility for the trial, a coordinating centre to handle materials, data collection and communication with trialists, a trial coordinator to run the coordinating

centre and a DMC.[101] Smaller trials probably need a scaled-down but similar structure. Some trials may also need subgroups of the steering committee or DMC to monitor recruitment, data quality, compliance and follow-up. There have been concerns that some individuals may be on committees for several different trials and that this could lead to conflicts of interest.[102] No research relating to this issue or to how members of steering committees and DMCs are selected was identified.

The main role of the DMC is to review planned interim analyses to see if there is accumulating evidence to suggest that the trial should be stopped early or, in some instances, prolonged.[45] DMCs need:[42,43,45]

- to be independent from the trial and the sponsor (although the sponsor may be needed to provide important information about their product to the DMC)
- to include experts from all necessary disciplines
- to have explicit terms of reference
- to have access to the necessary data
- to collaborate with other similar RCTs
- to have appropriate resources.

However, it has been argued that it is not feasible for all trials to have a DMC because of their cost (although this can be minimised by, for example, using telephone conferencing) and the lack of experienced statisticians and independent experts.[43,103] Formal DMCs are recommended for large 'pivotal' trials, those measuring life-threatening end-points and those in which there is the potential for serious side-effects.[43,103] In many other trials, it has been suggested that interim analyses could be carried out by the trial statistician and principal investigator with the option of forming an *ad hoc* DMC if required.[45,103,104] However, interim analyses should not be released to other investigators without good reasons. One non-random comparison of 20 cancer trials showed that open reporting of interim results increased: the number of trials with declining recruitment over time; the number of trials that stopped without meeting defined objectives; and the number of trials in which early publications of results were inconsistent with the final results.[105]

The concern that involvement of a commercial sponsor in the design, data collection, analysis and reporting of a trial may lead to bias in favour of their product is difficult to substantiate on a wide scale. However, there are certainly examples in which bias has occurred, such as: the reporting of data from only favourable centres; the use of dosage regimens that favour the drug manufactured by the sponsor; the reporting of conclusions that do not match the data presented.[95,106] This is particularly worrying since a survey from one region of the UK showed that 65% of trials are sponsored by the pharmaceutical industry.[107] Some authors have recommended that all sponsored trials have independent steering committees responsible for the design, data collection, and analysis of the trial[94] but others have argued that this is not feasible or appropriate.[42,103] Large important commercially sponsored trials should have independent steering groups and DMCs whose members are not selected by the sponsor,[94] as well as independent data management and analysis. Alternative mechanisms for smaller trials include 'in-house' management of blinded data by the sponsor with independent analysis or 'in-house' analysis with restricted access to the unblinded results.[43]

Recommendations

1. Multicentre and important single centre trials need a steering group, an independent DMC and a trial coordinator. (Basis: anecdotal evidence.)
2. Small trials should establish an *ad hoc* DMC if interim analyses suggest that the trial should be stopped early or the design altered. (Basis: anecdotal evidence.)
3. DMCs should be multidisciplinary and have explicit terms of reference. (Basis: anecdotal evidence.)
4. Interim analyses should not be released to the study investigators unless absolutely necessary. (Basis: one small non-random comparative study.)
5. Important commercially sponsored trials should minimise the concerns about potential bias by having independent steering groups and DMCs, and independent data management and analysis. (Basis: anecdotal evidence.)

Recommendations for research

1. Specific aspects of trial structure, staffing and organisation that improve the quality and progress of large and small trials need further research.
2. Further research needed into whether small trials need formal structures such as steering groups and DMCs.
3. The risk of bias in commercially sponsored trials needs to be compared with non-sponsored trials so that the need for independent control of commercially sponsored trials can be assessed.
4. The selection of members for steering committees and DMCs needs further investigation.

Conclusion

Many trials probably fail to start, particularly due to of lack of funding. The progress of those that do start is often hindered by poor recruitment and other problems, while others are stopped prematurely because of inappropriate or inadequate interim analyses. A variety of problems during the conduct of trials can also affect their quality and increase the chance of a false-negative or false-positive result. These include subversion of randomisation, poor compliance with trial treatments, poor blinding, or too many withdrawals or losses to follow-up.

Although there have been many isolated reports of these problems occurring during RCTs, it is not known how common most of them are. Similarly, solutions to most of the problems are largely anecdotal. Large prospective surveys of trials are thus required to establish the prevalence of problems in the conduct of RCTs, and interventions to improve the conduct and quality of RCTs need to be properly evaluated. RCTs need to be well-organised to be successful but, again, there is little evidence on the specific organisational features that help to improve their progress and quality. It is likely that adequate piloting of methods would help to avoid many of the problems with the conduct of trials identified in this chapter but, again, the value of pilot studies has not been adequately evaluated.

Recommendation

- Pilot studies are required before starting most RCTs to refine important aspects of their design and conduct. (Basis: anecdotal evidence.)

Recommendations for research

1. Prospective surveys are needed to identify the main problems relating to the conduct of trials (especially problems other than recruitment) and the potential solutions.
2. Further well-designed evaluation studies of interventions designed to improve the conduct and quality of RCTs are required.

References

1. Easterbrook PJ, Matthews DR. Fate of research studies. *J R Soc Med* 1992;**85**:71–6.
2. Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA* 1990;**263**:1392–5.
3. Challah S, Mays NB. The randomised controlled trial in the evaluation of new technology: a case study. *BMJ* 1986;**292**:877–9.
4. Ashery RS, McAuliffe WE. Implementation issues and techniques in randomized trials of outpatient psychosocial treatments for drug abusers: recruitment of subjects. *Am J Drug Alcohol Abuse* 1992;**18**:305–29.
5. Holden G, Rosenberg G, Barker K, Tuhrim S, Brenner B. The recruitment of research participants: a review. *Soc Work Health Care* 1993;**19**(2):1–44.
6. Charlson ME, Horwitz RI. Applying results of randomised trials to clinical practice: impact of losses before randomisation. *BMJ* 1984;**289**:1281–4.
7. Coombs DW, Dimick A, Bronstein JM, Potts LH, Bowens B. Conceptual and methodologic problems in the evaluation of a new burn treatment modality. *J Burn Care Rehabil* 1993;**14**:568–71.
8. Diekmann JM, Smith JM. Strategies for assessment and recruitment of subjects for nursing research. *West J Nurs Res* 1989;**11**:418–30.
9. Hunninghake DB, Darby CA, Probstfield JL. Recruitment experience in clinical trials: literature summary and annotated bibliography [review]. *Control Clin Trials* 1987;**8**(4 suppl):6–30S.
10. Neaton JD, Grimm RH Jr, Cutler JA. Recruitment of participants for the multiple risk factor intervention trial (MRFIT). *Control Clin Trials* 1987;**8**(4 suppl):41–53S.
11. Collaborative Ocular Melanoma Study Group. Design and methods of a clinical trial for a rare condition. COMS report No 3. *Control Clin Trials* 1993;**14**:362–91.
12. Strandness DE Jr. Randomized clinical trials: enough bang for the buck [review]? *J Vasc Surg* 1991;**13**:544–8.
13. Carew BD, Ahn SA, Boichot HD, Dierenfeldt BJ, Dolan NA, Edens TR, *et al.* Recruitment strategies in the studies of left ventricular dysfunction (SOLVD): strategies for screening and enrolment in two concurrent but separate trials. *Control Clin Trials* 1992;**13**:325–38.
14. Durkin DA, Kjelsberg MO, Buist AS, Connett JE, Owens GR. Recruitment of participants in the Lung Health Study: I. Description of methods. *Control Clin Trials* 1993;**14**(2 suppl):20–37S.
15. Johnson S, Mebane-Sims I, Hogan PE, Stoy DB. Recruitment of postmenopausal women in the PEPI Trial (Postmenopausal Estrogen/Progestin Interventions). *Control Clin Trials* 1995;**16**(4 suppl):20–35S.

16. Loescher LJ, Graham VE, Aickin M, Meyskens FL Jr, Surwit EA. Development of a contingency recruitment plan for a Phase III chemoprevention trial of cervical dysplasia. *Prog Clin Biol Res* 1990;**339**:151–63.

17. Tu SW, Kemper CA, Lane NM, Carlson RW, Musen MA. A methodology for determining patients' eligibility for clinical trials. *Methods Inf Med* 1993;**32**:317–25.

18. Tognoni G, Alli C, Avanzini F, Bettelli G, Colombo F, Corso R, *et al.* Randomised clinical trials in general practice: lessons from a failure. *BMJ* 1991;**303**: 969–71.

19. Winkler G, Levin EA, Whalen EY, Larus J. Overcoming 'trial fatigue': a strategy for optimizing patient accrual speed and resource utilization. *Drug Inf J* 1996;**30**:35–40.

20. Anderson LA, Fogler J, Dedrick RF. Recruiting from the community: lessons learned from the diabetes care for older adults project. *Gerontologist* 1995;**35**:395–401.

21. Hjorth M, Holmberg E, Rodjer S, Taube A, Westin J. Patient accrual and quality of participation in a multicentre study on myeloma: a comparison between major and minor participating centres. *Br J Haematol* 1995;**91**:109–15.

22. Lawrence W Jr. Some problems with clinical trials. *Arch Surg* 1991;**126**:370–8.

23. Hawkins BS, Prior MJ, Fisher MR, Blackhurst DW. Relationship between rate of patient enrollment and quality of clinical center performance in two multicenter trials in ophthalmology. *Control Clin Trials* 1990;**11**:374–94.

24. Bjornson-Benson WM, Stibolt TB, Manske KA, Zavela KJ, Youtsey DJ, Buist AS. Monitoring recruitment effectiveness and cost in a clinical trial. *Control Clin Trials* 1993;**14**:52–67S.

25. Bradford RH. Participant recruitment to the Lipid Research Clinics Coronary Primary Prevention Trial. *Control Clin Trials* 1987;**8**(4 suppl):31–40S.

26. King AC, Harris RB, Haskell WL. Effect of recruitment strategy on types of subjects entered into a primary prevention clinical trial. *Ann Epidemiol* 1994;**4**:312–20.

27. Silagy CA, Campion K, McNeil JJ, Worsam B, Donnan GA, Tonkin AM. Comparison of recruitment strategies for a large-scale clinical trial in the elderly. *J Clin Epidemiol* 1991;**44**:1105–14.

28. Buchwald H, Matts JP, Hansen BJ, Long JM, Fitch LL. Program on Surgical Control of the Hyperlipidemias (POSCH): recruitment experience. *Control Clin Trials* 1987;**8**(4 suppl): 94–104S.

29. Daly M, Seay J, Balshem A, Lerman C, Engstrom P. Feasibility of a telephone survey to recruit health maintenance organization members into a tamoxifen chemoprevention trial. *Cancer Epidemiol Biomarkers Prev* 1992;**1**:413–16.

30. Rudick C, Anthonisen NR, Manfreda J. Recruiting healthy participants for a large clinical trial. *Control Clin Trials* 1993;**14**:68–79S.

31. Hudmon KS, Kinney AY. Issues of enrollment bias in cancer chemoprevention trials. *Cancer Bull* 1995;**47**:339–42.

32. Hunter CP, Frelick RW, Feldman AR, Bavier AR, Dunlap WH, Ford L, *et al.* Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log. *Cancer Treat Rep* 1987;**71**:559–65.

33. Collins R, Peto R, Gray R, Parish S. Large-scale randomized evidence: trials and overviews. In: Wetherall DJ, Ledingham JGG, Warrell DA, editors. Oxford textbook of medicine. 3rd ed. Oxford: Oxford University Press; 1996. vol. 1:21–32.

34. Peto R. Clinical trial reporting [letter]. *Lancet* 1996;**348**:894–5.

35. Hollis JF, Satterfield S, Smith F, Fouad M, Allender PS, Borhani N, *et al.* Recruitment for phase II of the trials of hypertension prevention. Effective strategies and predictors of randomization. Trials of Hypertension Prevention (TOHP) Collaborative Research Group. *Ann Epidemiol* 1995;**5**:140–8.

36. Maurer BT, Moreno SI, Pickard AR, Wurst BE, Norden DK, Schumacher HR Jr. A comparison of recruitment methods for an osteoarthritis exercise study. *Arthritis Care Res* 1995;**8**:161–6.

37. Schoenberger JA. Recruitment experience in the Aspirin Myocardial Infarction Study. *Control Clin Trials* 1987;**8**(4 suppl):74–8S.

38. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

39. Schulz KF. Unbiased research and the human spirit: the challenges of randomized controlled trials [editorial]. *Can Med Assoc J* 1995;**153**:783–6.

40. Akazawa K, Kamakura T, Sakamoto M, Odaka T, Nose Y. Patient registration and treatment allocation in multicenter clinical trials using a FAX-OCR system. *Methods Inf Med* 1994;**33**:530–4.

41. Hornbuckle J, Thornton JG, Kelly CR, Welch CR, Oldham J. Netpoints: patient randomisation on the web. *BMJ* 1997;**314**:1604.

42. Task Force of the Working Group on Arrhythmias of the European Society of Cardiology. The early termination of clinical trials: causes, consequences, and control. With special reference to trials in the field of arrhythmias and sudden death. *Eur Heart J* 1994;**15**:721–38.

43. Fleming TR, DeMets DL. Monitoring of clinical trials: issues and recommendations [review]. *Control Clin Trials* 1993;**14**:183–97.

44. Pocock SJ. When to stop a clinical trial. *BMJ* 1992;**305**:235–40.

45. Buyse M. Interim analyses, stopping rules and data monitoring in clinical trials in Europe. *Stat Med* 1993;**12**:509–20.

46. Sylvester R, Bartelink H, Rubens R. A reversal of fortune: practical problems in the monitoring and interpretation of an EORTC breast cancer trial. *Stat Med* 1994;**13**:1329–35.

47. Urquhart J. Role of patient compliance in clinical pharmacokinetics. A review of recent research. *Clin Pharmacokinet* 1994;**27**:202–15.

48. Besch CL. Compliance in clinical trials [editorial]. *AIDS* 1995;**9**:1–10.

49. Klein MC, Kaczorowski J, Robbins JM, Gauthier RJ, Jorgensen SH, Joshi AK. Physicians' beliefs and behaviour during a randomized controlled trial of episiotomy: consequences for women in their care. *Can Med Assoc J* 1995;**153**:769–79.

50. Melnikow J, Kiefe C. Patient compliance and medical research: issues in methodology. *J Gen Intern Med* 1994;**9**:96–105.

51. Rudd P, Ahmed S, Zachary V, Barton C, Bonduelle D. Improved compliance measures: applications in an ambulatory hypertensive drug trial. *Clin Pharmacol Ther* 1990;**48**:676–85.

52. Urquhart J. Ascertaining how much compliance is enough with outpatient antibiotic regimens. *Postgrad Med J* 1992;**68**(suppl 3):S49–59.

53. Richmond RL, Anderson P. Research in general practice for smokers and excessive drinkers in Australia and the UK: II. Representativeness of the results. *Addiction* 1994;**89**:41–7.

54. McWhinney IR, Bass MJ, Donner A. Evaluation of a palliative care service: problems and pitfalls. *BMJ* 1994;**309**:1340–2.

55. Morse EV, Simon PM, Besch CL, Walker J. Issues of recruitment, retention, and compliance in community-based clinical trials with traditionally under-served populations. *Appl Nurs Res* 1995;**8**:8–14.

56. Schechtman KB, Gordon MO. The effect of poor compliance and treatment side effects on sample size requirements in randomized clinical trials. *J Biopharm Stat* 1994;**4**:223–32.

57. Schron EB, Brooks MM, Gorkin L, Kellen JC, Morris M, Campion J, *et al.* Relation of socio-demographic, clinical, and quality-of-life variables to adherence in the Cardiac Arrhythmia Suppression Trial. *Cardiovasc Nurs* 1996;**32**(2):1–5.

58. Wingerson D, Sullivan M, Dager S, Flick S, Dunner D, Roy-Byrne P. Personality traits and early discontinuation from clinical trials in anxious patients. *J Clin Psychopharmacol* 1993;**13**:194–7.

59. Pledger GW. Compliance in clinical trials: impact on design, analysis and interpretation [review]. *Epilepsy Res Suppl* 1988;**1**:125–33.

60. Vander Stichele R. Measurement of patient compliance and the interpretation of randomized clinical trials. *Eur J Clin Pharmacol* 1991;**41**:27–35.

61. Stephenson BJ, Rowe BH, Haynes RB, Macharia WM, Leon G. Is this patient taking the treatment as prescribed? *JAMA* 1993;**269**:2779–81.

62. Maenpaa H, Manninen V, Heinonen OP. Comparison of the digoxin marker with capsule counting and compliance questionnaire methods for measuring compliance to medication in a clinical trial. *Eur Heart J* 1987;**8**(suppl 1):39–43.

63. Atwood JR, Haase J, Rees-McGee S, Blackwell G, Giordano L, Earnest D, *et al.* Reasons related to adherence in community-based field studies. *Patient Educ Counsel* 1992;**19**:251–9.

64. Dunbar J, Cleary PA, Siebert C, Baker L, Brink S, Nathan DM. Implementation of a multicomponent process to obtain informed consent in the diabetes control and complications trial. *Control Clin Trials* 1989;**10**:83–96.

65. Morrison W. Trials and tribulations: patients' perspectives of the BETASERON study. *Axon* 1994;**16**(2):51–4.

66. Haynes RB, McKibbon KA, Kanani R. Systematic review of randomised controlled trials of the effects on patient adherence and outcomes of interventions to assist patients to follow prescriptions for medications. (Cochrane review). In: Cochrane Library, Issue 4, 1996. Oxford: Update Software.

67. Putnam DE, Finney JW, Barkley PL, Bonner MJ. Enhancing commitment improves adherence to a medical regimen. *J Consult Clin Psychol* 1994; **62**:191–4.

68. Raynor DK, Booth TG, Blenkinsopp A. Effects of computer generated reminder charts on patients' compliance with drug regimens. *BMJ* 1993;**306**: 1158–61.

69. Macharia WM, Leon G, Rowe BH, Stephenson BJ, Haynes RB. An overview of interventions to improve compliance with appointment keeping for medical services. *JAMA* 1992;**267**:1813–17.

70. Lee YJ, Ellenberg JH, Hirtz G, Nelson KB. Analysis of clinical trials by treatment actually received: is it really an option? *Stat Med* 1991;**10**:1595–605.

71. Kleijnen J, ter Riet G, Knipschild P. Acupuncture and asthma: a review of controlled trials [review]. *Thorax* 1991;**46**:799–802.

72. Liberati A, Himel HN, Chalmers TC. A quality assessment of randomized control trials of primary treatment of breast cancer. *J Clin Oncol* 1986;**4**:942–51.

73. Lionetto R, Pugliese V, Bruzzi P, Rosso R. No standard treatment is available for advanced pancreatic cancer [review: see comments]. *Eur J Cancer* 1995;**31A**:882–7.

74. Rossetti L, Marchetti I, Orzalesi N, Scorpiglione N, Liberati A. Is proper methodology associated with the use of a clinically relevant outcome measure? The case of randomized clinical trials on medical treatment of open-angle glaucoma. *Online J Current Clin Trials* [serial online] 1993 Nov 11; Doc No 100: [6512 words; 199 paragraphs].

75. Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. *BMJ* 1996;**312**:742–4.

76. Solomon MJ, Laxamana A, Devore L, McLeod RS. Randomized controlled trials in surgery. *Surgery* 1994;**115**:707–12.

77. Sonis J, Joines J. The quality of clinical trials published in the *Journal of Family Practice,* 1974–1991. *J Fam Pract* 1994;**39**:225–35.

78. Orr PR, Blackhurst DW, Hawkins BS. Patient and clinic factors predictive of missed visits and inactive status in a multicenter clinical trial. The Macular Photocoagulation Study Group. *Control Clin Trials* 1992;**13**:40–9.

79. Probstfield JL, Russell ML, Henske JC, Reardon RJ, Insull W Jr. Successful program for recovery of dropouts to a clinical trial. *Am J Med* 1986;**80**: 777–84.

80. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Stat Med* 1995;**14**:2057–79.

81. Fisher S, Greenberg RP. How sound is the double-blind design for evaluating psychotropic drugs [review]? *J Nerv Ment Dis* 1993;**181**:345–50.

82. Margraf J, Ehlers A, Roth WT, Clark DB, Sheikh J, Agras WS, *et al.* How 'blind' are double-blind studies? *J Consult Clin Psychol* 1991;**59**:184–7.

83. Morin CM, Colecchi C, Brink D, Astruc M, Mercer J, Remsberg S. How 'blind' are double-blind placebo-controlled trials of benzodiazepine hypnotics? *Sleep* 1995;**18**:240–5.

84. Moscucci M, Byrne L, Weintraub M, Cox C. Blinding, unblinding, and the placebo effect: an analysis of patients' guesses of treatment assignment in a double-blind clinical trial. *Clin Pharmacol Ther* 1987;**41**:259–65.

85. Talley NJ. A critique of therapeutic trials in Helicobacter pylori-positive functional dyspepsia [review]. *Gastroenterology* 1994;**106**:1174–83.

86. Kirsch I, Rosadino MJ. Do double-blind studies with informed consent yield externally valid results? An empirical test. *Psychopharmacology (Berl)* 1993;**110**:437–42.

87. Carroll KM, Rounsaville BJ, Nich C. Blind man's bluff: effectiveness and significance of psychotherapy and pharmacotherapy blinding procedures in a clinical trial. *J Consult Clin Psychol* 1994;**62**:276–80.

88. Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994;**44**:16–20.

89. Guyatt GH, Townsend M, Keller JL, Singer J. Should study subjects see their previous responses? Data from a randomized controlled trial. *J Clin Epidemiol* 1989;**42**:913–20.

90. Gibson D, Harvey AJ, Everett V, Parmar MK. Is double data entry necessary? The CHART trials. [Continuous, Hyperfractionated, Accelerated RadioTherapy]. *Control Clin Trials* 1994;**15**:482–8.

91. Reynolds-Haertle RA, McBride R. Single vs. double data entry in CAST. *Control Clin Trials* 1992;**13**: 487–94.

92. Poy E. Objectives of QC systems and QA function in clinical research. *Qual Assur* 1993;**2**:326–31.

93. Wood DJ. Quality assurance in Phase II and III studies. *Qual Assur* 1993;**2**:353–8.

94. van Gijn J. Who is in charge? Investigator and sponsor alliances. *Cerebrovasc Dis* 1995;**5**:22–6.

95. Julian D. Trials and tribulations. *Cardiovasc Res* 1994;**28**:598–603.

96. Jeiven ML. Patient and clinical site compliance: preparing user-friendly clinical trial materials. *Clin Res Reg Affairs* 1994;**11**:67–79.

97. Cohen MC, McKenna C, Lewis SM, Muller JE. Requirements for controlled clinical trials of preoperative cardiovascular risk reduction. *Control Clin Trials* 1995;**16**:89–95.

98. McCandlish R, Renfrew M. Trial and tribulation ... treatments for inverted nipples ... difficulties involved in research. *Nurs Times* 1991;**87**(38):40–1.

99. Banks V, Bale SE. Practical problems of undertaking clinical trials in the community. *J Wound Care* 1994;**3**:301–4.

100. Hogg RJ. Trials and tribulations of multicenter studies. Lessons learned from the experiences of the Southwest Pediatric Nephrology Study Group (SPNSG) [review]. *Pediatr Nephrol* 1991;**5**:348–51.

101. Warlow C. How to do it. Organise a multicentre trial. *BMJ* 1990;**300**:180–3.

102. Bogousslavsky J. The jungle of acute stroke trials, trialists and sponsors. *Cerebrovasc Dis* 1995;**5**:1–2.

103. Rockhold FW, Enas GG. Data monitoring and interim analyses in the pharmaceutical industry: ethical and logistical considerations [review]. *Stat Med* 1993;**12**:471–9.

104. Parmar MK, Machin D. Monitoring clinical trials: experience of, and proposals under consideration by, the Cancer Therapy Committee of the British Medical Research Council. *Stat Med* 1993;**12**: 497–504.

105. Green SJ, Fleming TR, O'Fallon JR. Policies for study monitoring and interim reporting of results. *J Clin Oncol* 1987;**5**:1477–84.

106. Rochon PA, Gurwitz JH, Simms RW, Fortin PR, Felson DT, Minaker KL, *et al.* A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med* 1994;**154**:157–63.

107. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;**337**:867–72.

Chapter 6

The analysis of RCTs

Methods

The objectives, in this chapter of the review, are limited. Clearly, it would not be sensible to attempt to review appropriate methods of analysis for all the variety of trial designs and types of data which exist. Also, no attempt is made to examine the arguments for and against Bayesian and frequentist philosophies, as Bayesian methods are being examined in another review within the NHS HTA programme. There is a great deal of literature in which deficits in the quality of statistical reporting are examined and this is not repeated here. Instead, those areas of analysis which have attracted criticism in the literature are considered, in which there may be empirical evidence leading to a set of recommendations.

There is, of course, a close association between how a trial is analysed and how it is designed (see chapter 3), conducted (see chapter 5) and reported (see chapter 7). Correspondingly, some of the topics below could equally well have been covered in any of those chapters. As far as possible, duplication between the chapters has been avoided and so the reader is also referred to chapters 3, 5 and 7 for additional topics within the broad area of analysis.

The strategy employed here was initially to produce a list of those articles with positive scores for analysis, subdivided by keywords. The keywords 'ITT analysis' (33 articles), 'multiple end-points' (four articles), 'subgroup analyses' (14 articles), and 'significance level' (six articles) were used directly for four of the following sections and all of these articles were read in full. Other articles scoring 2 or 3 for analysis were also scanned for relevance, given the limited objectives of this chapter. Articles dealing with the 'comprehensive cohort follow-up' design were included as were those relating to changes in quality over time.

Intention-to-treat analysis

The principle of analysis by ITT arises from RCTs in which, for a variety of reasons, the treatment to which a patient is randomised might not be received, in full or part, by that patient. In these circumstances, those patients who do not receive the allocated treatment may well differ systematically in some way from those who do receive the allocated treatment. Furthermore, these selection processes may differ between treatment groups. An analysis based only on patients actually receiving the treatment as randomised would therefore give a biased estimate of treatment differences. This bias can be mitigated by analysing the RCT according to the groups into which patients were randomised. If outcome observations are available for all patients, treatment comparisons are potentially unbiased, although bias may arise from some other aspect of the trial's design or conduct. Another advantage of ITT is that it mimics 'real life' with respect to errors in treatment and poor adherence to treatment, and gives a realistic assessment of a treatment policy.

The vast majority of papers dealing with general issues in the analysis of RCTs, emphasise the importance of ITT analysis.[1–3] The importance of the ITT principle is also implicit within systematic reviews, in which the proportion of trials reporting an ITT analysis may be reported (see chapter 7).

Peduzzi and colleagues[4,5] presented a case study of the Coronary Artery Surgery Study (or CASS), in which they compared ITT analysis with a range of alternative strategies for analysing data which take into account adherence to the randomised treatment. They concluded that only ITT is recommended. Lee and colleagues[6] presented a case study based on a trial to evaluate phenobarbital in children with febrile seizures and came to the same conclusion.

May and colleagues[7] made the point that an ITT analysis may imply the need for a larger sample size than the 'true' treatment difference would suggest. Despite widespread support for the ITT principle, there have been papers which are critical of this approach. Kannel[8] commented on the results of the Coronary Artery Surgery Study. The original analysis had emphasised difficulties with approaches other than ITT but, Kannel argued, at a time when 40% of patients in the medical arm of this trial have received the surgical treatment, the difficulties with ITT are greater:

"The huge crossover bias in this study seems to outweigh any possible bias of on-treatment analysis, especially because those crossing over to surgery were likely to have deteriorated clinically."

Sheiner[9] argued that ITT analysis is not optimal because it ignores compliance, power may be reduced and the question is totally changed from whether the drug itself has efficacy to whether ITT with the drug has effectiveness, that is, it assesses treatment policy rather than just treatment. Although Sheiner considered this a weakness, it does reflect real life – and the authors of this review consider this to be a strength and also note that compliance is a useful outcome in itself. If there is poor compliance, it suggests that the treatment as used in the RCT is not practical.

Rabeneck and colleagues[10] were also critical of the standard ITT approach. They stated that when there are significant modifications to the randomly assigned treatments, clinical treatment of patients is not advanced by the policy of entirely excluding treatment changes from consideration in the analysis. To enhance the relevance of assessments of effectiveness, the analysis should attempt to use information on changes in treatment. Sheiner and Rubin[11] proposed using measures of compliance with treatment to obtain estimates of the true treatment effect or method effectiveness. The general approach, based on an instrumental variable analysis is attractive in principle; however, it should be stressed that this is a theoretical paper and the method does not appear to have been used in practice. In criticism of ITT, Sheiner and Rubin argued that ITT does not even give reliable estimates of the outcomes associated with the prescription of alternative treatments, as the pattern of non-compliance will differ in future non-trial patients.

An alternative approach to ITT is the 'per protocol' analysis, in which only subjects conforming to the protocol are subject to analysis. In addition to the concerns about selection bias outlined above, the danger inherent in this approach is its susceptibility to manipulation and data dredging, especially if not pre-specified.[6]

Recommendation

- In the view of the authors, ITT will be the method of choice when an unbiased estimate of treatment effects are required. This will apply particularly when the objectives of the trial are pragmatic in terms of treatment effectiveness. There are, however, situations in which the aims of the trial are more explanatory than pragmatic and, in these circumstances, a biased estimate of treatment effectiveness may be acceptable; however, the rules for defining a 'per protocol' analysis need to precede the analysis. (Basis: case studies and logical argument.)

The comprehensive cohort follow-up study design

One criticism made of RCTs is that they often lack generalisability. This may arise because only a small proportion of the available population may be successfully randomised into an RCT. Olschewski and colleagues[12] described and illustrated the comprehensive cohort follow-up approach for the Coronary Artery Surgery Study. The basis of this approach is that everybody eligible for entry into a trial should be followed-up in a similar manner, irrespective of whether or not they accepted randomisation. The treatment difference between randomised treatments A and B will then be estimated in those accepting randomisation, after incorporating into the analysis relevant covariates of prognostic value. A separate estimate of the treatment difference will also be obtained from those rejecting randomisation but selecting treatment A or B. Analysis of the study results focuses on inferences about a parameter representing the difference between the size of the treatment effects in the randomised and non-randomised subgroups, with formal significance testing being based on a treatment by randomisation interaction. Schmoor and colleagues[13] reported a similar approach applied to three trials in breast cancer. They pointed out that the sample size requirement for randomised patients is no less than with a conventional approach and that the number of non-randomised patients must be substantially greater if interactions between treatment and prognostic factors are to be investigated. Weighing the advantages and disadvantages of a comprehensive cohort study, which included an increased workload for trialists, they could not recommend its routine use, a view which is endorsed by Silverman and Altman.[14] The new information management strategy in the NHS may make the establishment of such studies easier in the future.

Multiple end-points

Multiple end-points are one of several ways in which the interpretation of significance tests

can be affected by a multiplicity of tests being performed. The problem of multiple end-points has been highlighted in a systematic review[15] and in two non-systematic reviews[16,17] of RCTs in rheumatoid arthritis. The authors of the latter papers identified a problem with the use of measures which are insensitive to change and suggested that fewer measures would improve quality. They recommended using a combined score. Blair and Silman[16] also identified the need to identify outcome variables which are sensitive and suggested the establishment of a central core of sensitive measures for use in RCTs in rheumatoid arthritis. In a systematic review of 45 RCTs and comparative trials in prestigious journals, Pocock and colleagues[18] found a median of six end-points being reported, with a median of four significance tests being performed. This review also identified the particular multiple-testing problem of repeated measurements over time. In this survey, 40% of trials had such data. Of these 18 trials, descriptive methods only were used in ten but in eight the results of significance tests at several time points were reported, with a serious increase in the risk of a type I error.

Recommendation

- Regulatory bodies already impose on pharmaceutical companies the requirement to identify, at the stage of preparing the trial protocol, a primary outcome variable, which may be supplemented by a limited number of secondary outcome variables plus safety variables. A statistical analysis plan also has to be prepared, indicating the plan of analysis. This will deal in detail with such aspects of the data as repeated observations and how these will be analysed to avoid problems of multiple testing.

  Such an approach should be followed in the wider population of RCTs. Some grant-awarding bodies already require this but many trials are not subject to the discipline of seeking grant funding. However, all RCTs in the UK will be presented to ethics committees and the authors consider that they offer the most promising route for improving standards. (Basis: logical argument.)

Subgroup analysis

An understandable objective of trialists may be the identification of subgroups of patients who respond particularly well to one or other treatment. Statistical considerations indicate, however, that there are dangers of multiple-testing arising from this approach. That is, there may be a considerable number of variables recorded about a patient before entry into a trial and any of these could conceivably be used as the basis of a subgroup analysis. If multiple tests are then performed, it can be expected, by chance alone, that some will show apparent enhanced treatment differences within subgroups.

This has been recognised in papers which are predominantly educational in content (see, for example, those by Bigby and Gadenne[2] and Tannock[19]). Objective information about subgroup analysis from systematic reviews was sparse in the papers identified by our search strategy. Pocock and colleagues,[18] in a review of 45 papers in prestigious journals, found that 23 (51%) reported subgroup analyses. Ten of these included more than one prognostic factor in their subgroup analyses and one of these reported only within-group *p*-values, with no clear idea of the real effects of the prognostic factors on the value of treatment. In a systematic review of 196 double-blind trials of NSAIDs in rheumatoid arthritis, Gøtzsche[15] found that results from subgroup analyses were stressed in four trials, in which the main analysis had not favoured the new drug.

Taking a different approach, Counsell and colleagues[20] used a simulation method to illustrate the problem. Participants in a practical class in statistics rolled possibly weighted (but, in reality, unweighted) dice of different colours to simulate mortality in a clinical trial, with a six denoting 'death'. Red, green and white dice represented 'active' treatments. Each participant conducted two 'trials' of an 'active' treatment against 'control', with variable sample sizes. The resultant trial results were then subjected to a simulated publication bias, similar in magnitude to that experienced in real trials. A meta-analysis was performed with the inclusion of some pre-specified subgroup analyses. The simulation demonstrated a significant treatment effect in 'published' trials with 'experienced' clinicians (i.e. using the results of their second trial), illustrating how plausible subgroup analyses can generate false-positive findings.

A totally different aspect of subgroup analysis relevant to clinical trials, sponsored by the National Institutes of Health (NIH) in the USA, was highlighted by Freedman and colleagues.[21] They described the implication of the NIH Revitalization Act of 1993, which requires that "the trial is designed and carried out in a manner sufficient for valid analysis of whether

the variables being studied in the trial affect women or members of minority groups, as the case may be, differently than other subjects in the trial". The interpretation of this phrase was crucial because it could impair the ability to carry out clinical trials at all. They calculated that if trials were powered to compare the degree by benefit in men and women, the sizes of trials would have to increase by a factor of four to 16 and that, with five racial/ethnic subgroups, the inflation factor would range from 10 to 40. Freedman and colleagues interpreted the 'valid analysis' specified by the Act instead to require:

(i) that allocation from subgroups to intervention and control groups be by an unbiased procedure such as randomisation
(ii) unbiased assessment of the outcome of study participants
(iii) use of unbiased statistical analyses and proper methods of inference to estimate and compare the intervention effects between the gender and racial/ethnic subgroups.

The authors described the philosophy of the analysis of Phase III clinical trials as having two components. The first is to test the primary question posed by the trial over the total set of subjects. The other is the conduct of secondary analyses to identify questions with sufficient scientific basis to be tested as primary questions in future trials. Such secondary analyses may involve the examination of intervention effects in subgroups which may be defined by gender, race or ethnicity, among other factors. The NIH Revitalization Act is therefore interpreted as requiring appropriate representation according to gender and race/ethnicity, allowing investigation of these factors in subgroup analyses. The general philosophy, as outlined above, would then apply to gender and race/ethnicity, as it would to any other subgrouping of the population. Subsequent responses to the paper by Freedman and colleagues[21] gave general endorsement to their approach.[22–29]

Recommendations

1. Any subgroup analyses which are to be considered as hypothesis testing should be specified in the trial protocol. These should be limited to subgroups in which there is an *a priori* reason to expect a subgroup by treatment interaction.
2. All other subgroup analyses should be regarded as hypothesis-generating and their interpretation in this more limited role should depend on the number of subgroups examined.
3. Subgroup analyses should be approached through a comparison of the estimates of treatment effects across the levels of a subgroup variable. The practice of carrying out significance tests within levels of a subgroup variable should be forcefully discouraged. Any significance tests should be based on subgroup by treatment interactions.
4. Medical journal editors should be encouraged to implement the above recommendations in their journal policy, as there is current evidence from case studies that even journals with a positive attitude to statistical refereeing occasionally allow possibly misleading subgroup analyses to feature strongly in the abstract of a paper.[30]

(Basis: strength of the logical arguments and their widespread endorsement.)

Recommendation for research

- Analysis is required of how frequently the results from subgroup analyses have been misleading.

Significance levels

Criticisms of the use of significance levels identified by this review have been focused on the interrelationship between statistical significance and clinical significance. As such, it is as much a design issue as an issue for analysis. In a paper which is essentially educational, aimed at improving research in the area of rehabilitation, Ottenbacher[31] identified confusion of clinical and statistical significance as one of three problems in the analysis and interpretation of investigations based on statistical testing of hypotheses. Allied to this, he identified low statistical power and the importance of replicating results as the other two main problems. Klein[32] identified statistical analyses focused on detecting statistical significance, rather than clinical significance in Phase II trials, as one of ten correctable flaws. Bigby and Gadenne[2] also emphasised that statistical significance does not equate to medical significance and give further references in support of this. Lindgren and colleagues[33] undertook a systematic review of clinical trials based on the text-word 'FVC' (forced vital capacity) focusing on the contrast between clinical and statistical significance. In the 121 articles they identified, 92.6% discussed statistical significance but only 13.2% discussed clinical significance. None of the papers discussed clinical significance without also discussing statistical significance. However, although 21.5% of the articles discussed sample size, only 5% specifically discussed power. This is disappointing because

a well-designed trial will have a high power to detect as statistically significant a difference which is clinically significant.

Two-tailed tests of significance are, by consensus and convention, usually recommended in preference to one-tailed tests. However, Overall[34] argued that when one wishes to support the directional claim of superiority of drug over placebo for regulatory purposes, one-sided significance tests or CIs should be used. He argued that the additional 25% of patients required for comparable power are being unnecessarily required to forego effective treatment in order to serve as experimental controls, and emphasised that this should be a design issue rather than just an issue for analysis. This argument discounts the possibility of a treatment being harmful. The difference of opinion is based on philosophical issues and the views of the authors of this report favour the widely-held position that two-tailed tests would underpin the usual analysis of choice.

Recommendation

- Two-tailed tests and corresponding two-tailed CIs should be used unless one is so convinced that differences can only occur in one pre-specified direction and that differences in the opposite direction, however large or unlikely, would irrevocably be interpreted as due to chance. (Basis: philosophical argument.)

Other areas of analysis

Many of the deficiencies in the analysis of RCTs are well-recognised and uncontroversial. Numerous surveys have revealed that the proportion of published papers with erroneous statistical analyses is disturbingly large, and little would be gained from reinvestigating this area. Similarly, it is widely recognised that the use of CIs is underemployed. These issues have been addressed extensively in the literature by a plethora of educational articles in most specialist areas of medicine, as well as in the leading general medical journals. The volume of evidence which has emerged about poor practice, and the strength of the philosophical arguments, have convinced many medical journals of the need for expert statistical referees and for statisticians on editorial boards. It is an area in which systematic reviews have already produced change but there is conflicting evidence of improvements in the quality of published RCTs over time. Assessments of quality cover design as well as analysis but Balas[35] and Sonis and Joines[36] reported significant improvements in quality over time, while Marsoni and colleagues,[37] reviewing chemotherapy trials in advanced ovarian cancer, and Rosetti and colleagues,[38] reviewing RCTs in glaucoma, detected no improvement. In his systematic review in rheumatoid arthritis, Gøtzsche[15] examined separately design aspects and analysis aspects. Between 1959 and 1984, he found a statistically significant worsening in the design score ($p = 0.02$) but a statistically significant improvement in the analysis score ($p = 0.002$).

Thus, although it may appear that the authors of this report put relatively less weight on the inadequacies of analyses, these are often found and the emphasis should not be taken as indicative that there is cause for complacency. The authors consider that ignorance of the problem has been addressed and that the messages about good practice are widely accepted. The difficulty is turning goodwill and good intentions into good practice. The observed quality implies a need for statistical training for trialists. Journals are restricted by the shortage of referees with sufficient expertise to identify possible misuses of techniques and sufficient time to dedicate to refereeing. The recent introduction of small fees for referees may help those journals whose finances can afford this policy but it cannot be adopted by all. No literature has been identified on the subject but it seems apparent to the authors that until statistical refereeing is valued in career terms, both in the academic world and in the pharmaceutical industry, the maintenance of quality will depend on individuals putting altruism above self-interest. This does not seem a solid basis on which to build.

Recommendations

1. The status of refereeing should be elevated so that it is of positive value in career development.
2. There is a need for improvement in the statistical training of trialists, with evaluation of the effect of training on the adequacy of the statistical reporting.

(Basis: logical reasoning.)

References

1. The early termination of clinical trials: causes, consequences, and control. With special reference to trials in the field of arrhythmias and sudden death (Task Force of the Working Group on Arrhythmias of the European Society of Cardiology). *Eur Heart J* 1994;**15**:721–38.

2. Bigby M, Gadenne AS. Understanding and evaluating clinical trials. *J Am Acad Dermatol* 1996;**34**:555–94.

3. Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research [review]. *Int J Epidemiol* 1992;**21**:837–41.

4. Peduzzi P, Detre K, Wittes J, Holford T. Intent-to-treat analysis and the problem of crossovers. An example from the Veterans Administration Coronary Bypass Surgery Study. *J Thorac Cardiovasc Surg* 1991;**101**:481–7.

5. Peduzzi P, Wittes J, Detre K, Holford T. Analysis as randomized and the problem of non-adherence: an example from the Veterans Affairs Randomized Trial of Coronary Artery Bypass Surgery. *Stat Med* 1993;**12**:1185–95.

6. Lee YJ, Ellenberg JH, Hirtz G, Nelson KB. Analysis of clinical trials by treatment actually received: is it really an option? *Stat Med* 1991;**10**:1595–605.

7. May GS, DeMets DL, Friedman LM, Furberg C, Passamani E. The randomized clinical trial: bias in analysis [review]. *Circulation* 1981;**64**:669–73.

8. Kannel WB. Coronary artery surgery study revisited. Limitation of the intent-to-treat principle [editorial; comment]. *Circulation* 1990;**82**:1859–62.

9. Sheiner LB. The intellectual health of clinical drug evaluation. *Clin Pharmacol Ther* 1991;**50**:4–9.

10. Rabeneck L, Viscoli CM, Horwitz RI. Problems in the conduct and analysis of randomized clinical trials. Are we getting the right answers to the wrong questions? *Arch Intern Med* 1992;**152**:507–12.

11. Sheiner LB, Rubin DB. Intention-to-treat analysis and the goals of clinical trials [review]. *Clin Pharmacol Ther* 1995;**57**:6–15.

12. Olschewski M, Schumacher M, Davis KB. Analysis of randomized and non-randomized patients in clinical trials using the comprehensive cohort follow-up study design. *Control Clin Trials* 1992;**13**:226–39.

13. Schmoor C, Olschewski M, Schumacher M. Randomized and non-randomized patients in clinical trials: experiences with comprehensive cohort studies. *Stat Med* 1996;**15**:263–71.

14. Silverman WA, Altman DG. Patients' preferences and randomised trials. *Lancet* 1996;**347**:171–4.

15. Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of non-steroidal anti-inflammatory drugs in rheumatoid arthritis [published erratum appears in *Control Clin Trials* 1989;**50**:356]. *Control Clin Trials* 1989; **10**:31–56.

16. Blair PS, Silman AJ. Can clinical trials in rheumatology be improved [review]? *Curr Opin Rheumatol* 1991;**3**:272–9.

17. Felson DT, Anderson JJ, Meenan RF. Time for changes in the design, analysis, and reporting of rheumatoid arthritis clinical trials [review]. *Arthritis Rheum* 1990;**33**:140–9.

18. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987;**317**:426–32.

19. Tannock IF. Assessment of study design in clinical trials for bladder cancer. *Urol Clin North Am* 1992;**19**:655–62.

20. Counsell CE, Clarke MJ, Slattery J, Sandercock PA. The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? *BMJ* 1994;**309**:1677–81.

21. Freedman LS, Simon R, Foulkes MA, Friedman L, Geller NL, Gordon DJ, *et al.* Inclusion of women and minorities in clinical trials and the NIH revitalization act of 1993 – the perspective of NIH clinical trialists. *Control Clin Trials* 1995;**16**:277–85.

22. Bourassa MG, Roubin GS, Detre KM, Sopko G, Krone RJ, Attabuto MJ, *et al.* Bypass Angioplasty Revascularization Investigation: patient screening, selection, and recruitment. *Am J Cardiol* 1995;**75**(9): 3–8C.

23. Brawley OW. Response to 'Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993 – the perspective of NIH clinical trialists'. *Control Clin Trials* 1995;**16**:293–5.

24. Buist SA, Greenlick MR. Response to 'Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993 – the perspective of NIH clinical trialists'. *Control Clin Trials* 1995;**16**:296–8.

25. Fraser WD. Response to 'Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993 – the perspective of NIH clinical trialists'. *Control Clin Trials* 1995;**16**:299.

26. Baldwin W. Impact of the NIH inclusion guidelines. *Control Clin Trials* 1995;**16**:300.

27. Woolson RF, Jones MF, Clarke WR, Torner JC. Discussion of 'Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993 – the perspective of NIH clinical trialists'. *Control Clin Trials* 1995;**16**:301–3.

28. Meinert CL. The inclusion of women in clinical trials. *Science* 1995;**269**:795–6.

29. Piantadosi S. Commentary regarding 'Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993 – the perspective of NIH clinical trialists'. *Control Clin Trials* 1995;**16**:307–9.

30. Adachi JD, Bensen WG, Brown J, Hanley D, Hodsman A, Josse R, *et al.* Intermittent etidronate therapy to prevent corticosteroid-induced osteoporosis. *N Engl J Med* 1997;**337**:382–7.

31. Ottenbacher KJ. Why rehabilitation research does not work (as well as we think it should). *Arch Phys Med Rehabil* 1995;**76**:123–9.

32. Klein DF. Improvement of Phase III psychotropic drug trials by intensive Phase II work [review]. *Neuropsychopharmacology* 1991;**4**:251–71.

33. Lindgren BR, Wielinski CL, Finkelstein SM, Warwick WJ. Contrasting clinical and statistical significance within the research setting. *Pediatr Pulmonol* 1993;**16**:336–40.

34. Overall JE. Statistical efficiencies that the FDA should encourage [review]. *Psychopharmacol Bull* 1991;**27**:211–16.

35. Balas EA, Austin SM, Ewigman BG, Brown GD, Mitchell JA. Methods of randomized controlled clinical trials in health services research. *Med Care* 1995;**33**:687–99.

36. Sonis J, Joines J. The quality of clinical trials published in the *Journal of Family Practice*, 1974–1991. *J Fam Pract* 1994;**39**:225–35.

37. Marsoni S, Torri W, Taiana A, Gambino A, Grilli R, Liati P, *et al.* Critical review of the quality and development of randomized clinical trials (RCTs) and their influence on the treatment of advanced epithelial ovarian cancer. *Ann Oncol* 1990;**1**:343–50.

38. Rossetti L, Marchetti I, Orzalesi N, Scorpiglione N, Liberati A. Is proper methodology associated with the use of a clinically relevant outcome measure? The case of randomized clinical trials on medical treatment of open-angle glaucoma. *Online J Curr Clin Trials* [serial online] 1993 Nov 11; Doc No 100: [6512 words; 199 paragraphs].

Chapter 7

Limiting factors relating to the reporting of RCTs

Methods

The methods applied were the same as those described in chapter 5. Articles that were rated as being highly relevant to the reporting of randomised trials (coded 3) were read in full, as were those that were coded with the following keywords: description failure, impact, inappropriate conclusions, publication bias. The abstracts of all articles rated as moderately relevant (coded 2) were also read, as were the titles of uncoded articles (e.g. letters, news items), and those that appeared relevant were read in full. Other articles were identified from the reference lists of articles that were read and from personal knowledge of the field. The search resulted in about 100 complete articles that were read and summarised by a single reviewer (CC).

Poor reporting of trial details

Of the many articles highlighting the poor reporting of RCTs, many fail to differentiate between poor reporting and poor conduct. There may well be a correlation between the quality of the design and conduct of a trial and its reporting – good quality trials might be expected to be better reported – but there is little evidence to support this. If some aspect of conduct is not reported, it does not necessarily mean that it did not occur, and the only way to find out is to contact the authors. The trial may then turn out to be of higher quality than reported, although one survey found that, in most cases, if an item of conduct was not reported it did not occur.[1] Some trials in which details are poorly reported may actually be of poorer quality than appears from the report. For example, some trials described simply as 'randomised' have, in fact, been shown not to use true random allocation.

Many surveys of the quality of reporting of randomised or controlled clinical trials were identified from the literature, hence, only those that assessed ten or more trials are included in this report. In most of these surveys, the trials were identified by various combinations of searching MEDLINE, searching journals by hand and searching reference lists. The various aspects of reporting were divided into those relating to internal validity, that is, the reduction of bias (*Table 25*),[1–24] and those relating to external validity or generalisability (*Table 26*).[1,3,5–9,12,14,16,17,22,23,25] The quality of the detailed reporting of statistical items is considered in the statistical section of this report.

No surveys reported all the items in the tables and many used slightly different definitions for the adequacy of reporting of certain features. For some of these items, many surveys simply assessed whether trials reported any details at all (e.g. baseline comparability of groups, sample size calculations, blinding), and it is likely that the details given were often inadequate or inappropriate.[2,13,19] For other items, the reliability of the results is questionable. For example, several surveys assessed whether ITT analyses were performed. However, when two statisticians independently scored trials on whether they were analysed on an ITT basis, their agreement was so poor that they decided not to collect these data.[26]

Despite these problems, *Tables 25* and *26* show that no features were consistently well reported and some items were very poorly reported (sample size, method of randomisation). Some studies have suggested that the quality of reporting has improved over time,[1,3,14,23] while others have not.[10,12,17,21] The type of author may influence the quality of reporting, as may the journal. One study showed that surgical trials were reported less well if a surgeon rather than a physician was first author and if they were published in specialist surgical journals.[22] Another study showed that the quality of reporting was better in one journal that had specifically tried to improve reporting, although it was still far from perfect.[18]

Recommendation for research

- The relationship between the quality of reporting and the quality of the design and conduct of RCTs needs further research.

TABLE 25 *Quality of reporting of the internal validity of randomised trials*

Study	Subject (number and type of trials)	Description of randomisation adequate (percentage of trials)	Adequate reporting of following features (percentage of trials)
Altman & Doré, 1990[2]	Various (80 RCTs)	Sequence generation: 49% Concealment: 26% Stratified or blocked: 40% unclear*	Number randomised: 77% Baseline comparability: 59% Blinding: – Compliance: – Sample size: 39% Exclusions after randomisation: poorly described ITT analysis: – Analyses repeatable: –
Balas, *et al.,* 1995[3]	Health services (101 RCTs)	Sequence generation: 24%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: 51% Blinding: – Compliance: – Sample size: poorly reported Exclusions after randomisation: 25% ITT analysis: – Analyses repeatable: –
Claessen *et al.,* 1992[4]	Otitis media (50 controlled clinical trials)	Sequence generation: 12%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: 74% Blinding: double-blind, 52% Compliance: 40% Sample size: 6% Exclusions after randomisation: – ITT analysis: – Analyses repeatable: –
Hall, *et al.,* 1996[5]	Surgery (346 RCTs)	Sequence generation: 27%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: 67% Blinding: outcome assessor, 48% Compliance: – Sample size: 19% Exclusions after randomisation: – ITT analysis: – Analyses repeatable: –
Kleijnen & Knipschild, 1991[6]	Mental function (53 controlled clinical trials)	Sequence generation: – Concealment: – Stratified or blocked: 30% stratified	Number randomised: – Baseline comparability: 38% Blinding: patient, 96%; outcome assessor, 94% Compliance: – Sample size: 8% Exclusions after randomisation: – ITT analysis: – Analyses repeatable: 30%
Kleijnen, *et al.,* 1991[7]	Asthma (13 controlled clinical trials)	Sequence generation: – Concealment: – Stratified or blocked: 8% stratified	Number randomised: – Baseline comparability: 54% Blinding: patient, 61%; outcome assessor, 54% Compliance: – Sample size: – Exclusions after randomisation: – ITT analysis: – Analyses repeatable: 46%

* *Treatment groups too similar in size in trials reporting no blocking; –, data not reported*
† *No differentiation between sequence generation and concealment*

continued

TABLE 25 contd *Quality of reporting of the internal validity of randomised trials*

Study	Subject (number and type of trials)	Description of randomisation adequate (percentage of trials)	Adequate reporting of following features (percentage of trials)
Kleijnen & Knipschild, 1992[8]	Cerebrovascular (40 controlled clinical trials)	Sequence generation: 10%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: 30% Blinding: patient, 48%; outcome assessor, 48% Compliance: – Sample size: – Exclusions after randomisation: – ITT analysis: – Analyses repeatable: 50%
Klein, 1988[9]	Irritable bowel (43 RCTs)	Sequence generation: – Concealment: – Stratified or blocked: –	Number randomised: – Baseline comparability: 27% Blinding: – Compliance: – Sample size: – Exclusions after randomisation: – ITT analysis: 39% Analyses repeatable: –
Koes, *et al.,* 1995[10]	Back pain (89 RCTs)	Sequence generation: <50%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: poor Blinding: patient, 22% Compliance: – Sample size: – Exclusions after randomisation: 29% gave reasons for drop-outs ITT analysis: 52% Analyses repeatable: 51%
Liberati, *et al.,* 1986[1]	Breast cancer (63 RCTs)	Sequence generation: – Concealment: 25% Stratified or blocked: –	Number randomised: – Baseline comparability: 60% Blinding: double-blind, 8% Compliance: 46% Sample size: 32% Exclusions after randomisation: 14% stated none; 21% unclear ITT analysis: 14% Analyses repeatable: –
Lionetto, *et al.,* 1995[11]	Pancreatic cancer (27 controlled clinical trials)	Sequence generation: 11%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: – Blinding: – Compliance: – Sample size: 18% Exclusions after randomisation: – ITT analysis: 41% Analyses repeatable: –
Marsoni, *et al.,* 1990[12]	Ovarian cancer (38 RCTs)	Sequence generation: 21%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: – Blinding: – Compliance: 24% Sample size: – Exclusions after randomisation: – ITT analysis: 13% Analyses repeatable: –

† *No differentiation between sequence generation and concealment;* –, *data not reported*

continued

TABLE 25 contd *Quality of reporting of the internal validity of randomised trials*

Study	Subject (number and type of trials)	Description of randomisation adequate (percentage of trials)	Adequate reporting of following features (percentage of trials)
Moher, *et al.*, 1994[13]	Various (102 RCTs with negative results)	Sequence generation: – Concealment: – Stratified or blocked: –	Number randomised: – Baseline comparability: – Blinding: – Compliance: – Sample size: 32% gave details but often poor Exclusions after randomisation: – ITT analysis: – Analyses repeatable: –
Nicolucci, *et al.*, 1989[14]	Lung cancer (150 RCTs)	Sequence generation: – Concealment: 12% Stratified or blocked: –	Number randomised: – Baseline comparability: 77% Blinding: double-blind 9%; blinding of outcome assessors not reported Compliance: 39% Sample size: 6% Exclusions after randomisation: 5% stated none; unclear in 10% ITT analysis: 5% Analyses repeatable: –
O'Donovan, *et al.*, 1993[15]	Male infertility (174 RCTs)	Sequence generation: 35%† Concealment:† Stratified or blocked: –	Number randomised: – Baseline comparability: – Blinding: double-blind, 36% Compliance: – Sample size: 4% Exclusions after randomisation: – ITT analysis: mostly not Analyses repeatable: –
Petersen & Kristensen, 1992[16]	Psoriasis (62 controlled clinical trials)	Sequence generation: – Concealment: – Stratified or blocked: –	Number randomised: – Baseline comparability: – Blinding: – Compliance: – Sample size: 0% Exclusions after randomisation: – ITT analysis: – Analyses repeatable: –
Rosetti, *et al.*, 1993[17]	Glaucoma (102 RCTs)	Sequence generation: – Concealment: – Stratified or blocked: –	Number randomised: – Baseline comparability: 41% Blinding: double-blind, 77% Compliance: – Sample size: 4% Exclusions after randomisation: – ITT analysis: 47% Analyses repeatable: –
Schulz, *et al.*, 1994[18]	Obstetrics (206 controlled clinical trials)	Sequence generation: 32% Concealment: 52% Stratified or blocked: Poor*	Number randomised: – Baseline comparability: 59% gave details but often poor Blinding: – Compliance: – Sample size: 24% Exclusions after randomisation: – ITT analysis: – Analyses repeatable: –

* *Treatment groups too similar in size in trials reporting no blocking; –, data not reported*
† *No differentiation between sequence generation and concealment*

continued

TABLE 25 contd *Quality of reporting of the internal validity of randomised trials*

Study	Subject (number and type of trials)	Description of randomisation adequate (percentage of trials)	Adequate reporting of following features (percentage of trials)
Schulz, *et al.*, 1996[19]	Obstetrics (110 RCTs)	Sequence generation: – Concealment: – Stratified or blocked: –	Number randomised: 92% Baseline comparability: – Blinding: double-blind, 48%‡; outcome assessor, 14%; only 2% tested success of blinding Compliance: – Sample size: – Exclusions after randomisation: poorly described; no exclusions, 10% ITT analysis: – Analyses repeatable: –
Silagy, *et al.*, 1994[20]	Primary care (55 RCTs)	Sequence generation: – Concealment: 25% Stratified or blocked: –	Number randomised: – Baseline comparability: – Blinding: outcome assessor, 75% Compliance: – Sample size: 9% Exclusions after randomisation: – ITT analysis: 69% Analyses repeatable: –
Silagy & Jewell, 1994[21]	Primary care (90 RCTs)	Sequence generation: – Concealment: 28% Stratified or blocked: –	Number randomised: – Baseline comparability: – Blinding: outcome assessor, 56% Compliance: – Sample size: – Exclusions after randomisation: – ITT analysis: 38% Analyses repeatable: –
Solomon, *et al.*, 1994[22]	Surgery (202 RCTs)	Sequence generation: – Concealment: 86% Stratified or blocked: –	Number randomised: – Baseline comparability: 73% Blinding: patient, 36%; outcome assessor, 34% Compliance: – Sample size: 11% Exclusions after randomisation: – ITT analysis: – Analyses repeatable: –
Sonis & Joines, 1994[23]	Primary care (53 RCTs)	Sequence generation: – Concealment: 26% Stratified or blocked: –	Number randomised: – Baseline comparability: 51% Blinding: patient, 66%‡; outcome assessor, 55%; success of blinding tested in 0% Compliance: 36% Sample size: 9% Exclusions after randomisation: – ITT analysis: 11% Analyses repeatable: –
Williams & Davis, 1994[24]	Various (200 controlled clinical trials)	Sequence generation: 9%† Concealment:† Stratified or blocked: 9%	Number randomised: – Baseline comparability: – Blinding: – Compliance: – Sample size: – Exclusions after randomisation: – ITT analysis: – Analyses repeatable: –

† *No differentiation between sequence generation and concealment;* –, *data not reported*
‡ *Percentage of trials in which patients could have been blinded*

TABLE 26 *Quality of reporting of external validity of randomised trials*

Study	Subject (number/ type of trials)	Percentage of trials with adequate reporting of each item					
		Case selection process	Rejection log	Eligibility/ diagnostic criteria	Description of included patients	Description of intervention	Reporting side-effects
Balas, *et al.*, 1995[3]	Health services (101 RCTs)	Poor	–	–	–	–	Poor
Hall, *et al.*, 1996[5]	Surgery (346 RCTs)	75	–	–	–	94	77
Kleijnen & Knipschild, 1991[6]	Mental function (53 controlled clinical trials)	–	–	40	–	–	–
Kleijnen, *et al.*, 1991[7]	Asthma (13 controlled clinical trials)	–	–	–	–	77	38
Kleijnen & Knipschild, 1992[8]	Cerebrovascular (40 controlled clinical trials)	–	–	–	30	95	–
Klein, 1988[9]	Irritable bowel (43 RCTs)	–	–	42	–	–	44
Liberati, *et al.*, 1986[1]	Breast cancer (63 RCTs)	46	17	–	–	87	52
Marsoni, *et al.*, 1990[12]	Ovarian cancer (38 RCTs)	–	–	34	–	84	–
Nicolucci, *et al.*, 1989[14]	Lung cancer (150 RCTs)	–	9	–	63	–	79
Petersen & Kristensen, 1992[16]	Psoriasis (62 controlled clinical trials)	5	–	11	–	–	–
Rochon, *et al.*, 1993[25]	Arthritis (83 RCTs)	–	–	–	Age and sex not reported in 12% and 22%, respectively	–	–
Rosetti, *et al.*, 1993[17]	Glaucoma (102 RCTs)	–	–	–	61	–	–
Solomon, *et al.*, 1994[22]	Surgery (202 RCTs)	69	37	–	–	–	60
Sonis & Joines, 1994[23]	Primary care (53 RCTs)	70	15	–	–	94	18

Completeness of reporting of trial results

Four studies from Europe and North America have analysed clinicians' views of the effectiveness of treatments and their likelihood to use them based on the way the results of the same trial were reported (*Table 27*).[27–30] The studies were not ideal. In particular, they were restricted to trials of the prevention of cardiovascular disease; they selected unrepresentative samples of mainly hospital doctors rather than surgeons or GPs; and they certainly simplified the clinical decision-making processes that occur in real life. However, they did give consistent results. Clinicians were more likely to view a treatment as effective or hazardous if presented with its relative benefits rather than its absolute benefits. However, absolute

TABLE 27 *Impact of reporting absolute versus relative treatment effects*

Study (country)	Study participants	Methods	Results	Notes
Bobbio, *et al.,* 1994[27] (Italy)	Physicians attending 'refresher courses' asked to rate likelihood of prescribing treatment (using visual analogue scale) based on trials' results.	Results of single beneficial outcome from same trial* presented in four ways (relative and absolute risk reduction, % patients event-free, NNT) and another statement gave relative risk reduction for both beneficial and harmful outcomes. Each result disguised as a different trial and order of results randomly varied.	148 respondents (50% response). For single outcome, mean likelihood to prescribe was higher with relative risk reduction (77%) than with other methods (which were all similar; 24–37%). Mean likelihood lower if adverse as well as beneficial relative risk reduction reported (23%).	Poor response. Non-representative sample. No doctors realised results were from same trial. Inappropriate use of parametric statistics – data not normally distributed.
Bucher, *et al.,* 1994[28] (Switzerland)	Random sample of internists and GPs asked to rate effectiveness of treatment (11-point scale) and likelihood of prescribing treatment for specified patient (7-point scale) based on trials' results.	Doctors randomly allocated to receive results of 3 outcomes (1 significant benefit, 1 non-significant benefit, 1 non-significant harm) from same trial* as either relative or absolute risk reduction. Both also received result for one outcome as NNT. Each result disguised as a different trial.	499 respondents (62% response). Effectiveness of treatment and likelihood of treatment rated significantly higher with relative than with absolute risk reduction. Effectiveness and likelihood to treatment rated lower with NNT than both relative and absolute risk reduction.	Poor response.
Forrow & Taylor, 1992[29] (USA)	Hospital doctors attending educational conferences and epidemiology training fellows asked to rate likelihood of prescribing treatment (7-point scale) based on trials' results.	Doctors received six results from trials of treatment of either hypertension or hyperlipidaemia, two of which gave results for same beneficial outcome as either relative or absolute risk reduction. Order of these two was reversed for half of the doctors.	235 responses (30–75% response). Hyperlipidaemia: 49% doctors more likely to treat given relative than absolute risk reduction; 48% gave same response for relative and absolute risk reduction. Hypertension: 33% doctors more likely to treat given relative than absolute risk reduction, 62% gave same result.	Poor response. Non-representative sample. No difference in results based on speciality, level of training, academic vs. non-academic practice.
Naylor, *et al.,* 1992[30] (Canada)	Hospital physicians known to authors or attending rounds asked to rate effectiveness of treatment (11-point scale) based on trials' results.	Doctors randomly allocated to receive results of 3 outcomes (1 significant benefit, 1 non-significant benefit, 1 non-significant harm) from same trial* as either relative or absolute risk reduction. Both also received result for one outcome as NNT. Each result disguised as a different trial.	100 responses (75% response to mailing). Effectiveness of treatment rated higher for significantly beneficial outcome and lower for non-significant adverse outcome with relative than with absolute risk reduction. Effectiveness rated lower with NNT than both relative and absolute risk reduction.	Non-representative sample.

* *Helsinki Heart Study*

treatment effects are usually the most important with regards to treatment of individual patients, although relative treatment effects may be more generalisable.[31] The 'number-needed-to-treat' (or NNT) statistic is directly related to the absolute risk[31] but two of the studies showed that clinicians were less likely to prescribe a treatment on the basis of the number-needed-to-treat than with the absolute risk reduction.[28,30] None of the studies tried to identify why relative benefits were more attractive than absolute ones although it was probably because absolute benefits were numerically much smaller (a treatment that reduces the risk of death from 2% to 1% has a relative risk reduction of 50% but an absolute risk reduction of only 1%). One study also showed that presenting both the benefits and hazards of a treatment reduced a clinician's enthusiasm for prescribing the treatment.[27]

Recommendation

- Results should be presented as both absolute and relative treatment effects and should include both benefits and side-effects or hazards. (Basis: two randomised and two non-random comparisons of the impact of reporting absolute versus relative treatment effects.)

Recommendations for research

1. The impact of presenting absolute versus relative treatment effects needs to be assessed on more representative samples of doctors and surgeons, and for a wider variety of trials than previously.
2. Randomised trials are needed to assess whether educating clinicians about the use of absolute risks can alter their interpretation of trial results and clinical practice.

Reporting of conclusions

The conclusions of a trial should be supported by the data presented. Only one study was identified in which this issue was assessed.[32] In this study, the results and conclusions of 61 trials of NSAIDs were assessed in a blinded fashion and then compared to see if the conclusions matched the data. Most (85%) of these trials were sponsored by pharmaceutical companies. Only 88% of positive conclusions were supported by the data (conclusions always favoured the sponsor's agent). Similarly, only 45% of reports of reduced toxicity were supported by the data. Again, these usually favoured the sponsor's drug.

The result of a trial should also be put in context by comparing it with all similar trials to give a balanced, authoritative discussion. This is best done by quoting or performing relevant systematic reviews (meta-analyses); however, this is rarely done. A recent survey of the reports of 26 trials published in five major journals in May 1997 showed that only two trials (8%; 95% CI, 1–25) presented their results in the context of an updated systematic review.[33] As a result, many trial reports have biased discussions. Two surveys have shown that positive trials tend to cite other positive trials rather than co-existing negative ones.[34,35]

Recommendation

- Conclusions should be supported by data and include a balanced discussion of all relevant evidence, preferably by quoting up-to-date systematic reviews. (Basis: three surveys of trial reports.)

Recommendation for research

- Further studies, similar to the one already undertaken by Rochon and colleagues,[32] are required on the frequency with which the conclusions of RCTs (both commercially sponsored and unsponsored) are supported by the data presented.

Reporting of commercially sponsored trials

The role any sponsor plays in the funding, design, conduct, analysis and reporting of a trial should be clearly stated in the reports of RCTs,[36] since some sponsored trials have been shown to be biased in favour of the manufacturer's product.[37] Rochon's survey of mainly commercially sponsored trials of NSAIDs showed that the trial design often favoured the company's agent (e.g. by comparing it with a suboptimal dose of another agent) and that conclusions were often biased in favour of the company's agent.[32] No studies were identified that specifically considered the adequacy of reporting of a company's involvement in trials or that considered the reporting of financial rewards given to investigators in sponsored trials. It is also important to note that even trials without commercial sponsorship may be associated with pressures to produce positive results in order to secure future funding and status.[37]

Recommendation

- Reports should make explicit the role of any sponsor in the design, conduct, data management, analysis and reporting of RCTs, and any financial inducements to participation. (Basis: one survey of the design and reporting of sponsored trials, plus further anecdotal evidence.)

Recommendation for research

- The adequacy of reporting of the role of sponsors in RCTs needs to be studied and further assessments made of the presence of bias in reports of sponsored trials.

Guidelines for the reporting of RCTs

Because of the deficiencies in the reporting of RCTs demonstrated above, several groups of experts have developed guidelines (evidence-based where possible) to improve reporting. The SORT (Standards of Reporting Trials) guidelines contain 32 items relating to the reporting of the internal validity of the trial.[38] There have been complaints that these guidelines reduced the readability of a report because of their rigid format and that no details of external validity were included.[39] The Asilomar group produced a list of 33 items relating to both the internal and external validity of a trial which, they hoped, would improve reporting, help meta-analysts, educate new trialists about good

methodology and improve peer review.[40] Sensibly, both these groups met to produce a single set of guidelines which incorporated the best points from both lists. The final Consolidated Standards of Reporting Trials (CONSORT) guidelines include 21 items to be included in a report of RCTs, together with a flow chart to document patient selection and follow-up, and recommendations relating to the title, abstract, introduction, and discussion.[41] However, it has been argued that the items relating to data collection on those not randomised are unnecessary,[42,43] while others have suggested that some important items relating to data quality have been omitted.[44] Details relating to the reporting of commercial involvement are also not included. The CONSORT guidelines do not cover the statistical aspects of trials and so should be used in conjunction with existing statistical guidelines.[45]

Recommendations

1. Authors and editors should use the CONSORT guidelines to improve the reporting of RCTs, although certain items may not be necessary and details of data quality checks should also be included. (Basis: evidence of poor reporting from many surveys of trial reports.)
2. The CONSORT guidelines should be regularly updated in the light of valid criticism and new research. (Basis: anecdotal evidence of problems with the CONSORT guidelines.)

Recommendations for research

1. The effect of the CONSORT guidelines on the quality of reporting of RCTs needs to be assessed.
2. The compliance of journals and authors with the CONSORT guidelines should be monitored.

Failure to publish completed RCTs

Failure to publish the results of completed trials is scientific misconduct because it deprives clinicians and patients of information they need to make rational decisions.[46] There are a number of reasons for failure to publish. Publication bias is the tendency to publish studies based on the 'direction or strength of study findings':[47] trials showing large, statistically significant treatment effects tend to be published more frequently than those showing small, non-significant effects. Dickersin and Min[47] studied retrospectively a cohort of 200 RCTs from two major centres in the USA (of which 92.5% were published) and showed the odds of publication were about nine times greater for trials with significant results than those with non-significant results – 98% of significant trials were published compared with 86% of non-significant ones (odds ratio 8.92; 95% CI 1.96–40.65). Trials with external rather than internal funding were also more likely to be published (94% versus 81%; odds ratio 3.59; 95% CI 1.03–12.52). There were non-significant trends for multicentre trials to be published more often than single centre ones (95% versus 90%) and for larger trials (at least 100 patients) to be published more frequently than smaller ones (94% versus 90%).

These analyses are limited by the small numbers of RCTs (particularly unpublished ones) and because they were based on trials from major centres. The publication of trials from lesser centres may differ. The use of odds ratios also exaggerates the apparent effect of publication bias, because so many RCTs are, in fact, published. The equivalent relative risk of publication of positive as opposed to negative RCTs is 1.15. Dickersin and Min[47] found that the main reason for trials remaining unpublished was because the authors did not write them up and submit them to journals, either because the results were 'not interesting' or because of problems with co-investigators or because of lack of time. No trial in this cohort remained unpublished because of rejection by a journal, although this does occur.[37,48,49] In another study of 148 clinical trials (not just RCTs) from the UK, similar results were obtained, although the odds ratio for the publication of significant versus non-significant trials was smaller (2.10).[50] This study also showed that trials sponsored by pharmaceutical companies were published less often than other trials (odds ratio 0.17; 95% CI 0.05–0.53); one of the reasons cited for lack of publication was that sponsors had control of the data. Publication bias is likely to be one of the main reasons why trials sponsored by pharmaceutical companies usually report results that favour the company's product. One survey of 107 trials published in major journals found that 89% (33/37) of trials supported by pharmaceutical companies reported results that favoured the company's product and none reported results favouring another company's product. By comparison, 61% (43/70) of trials which did not acknowledge commercial support reported results in favour of a new therapy over a traditional therapy.[51]

Many studies, including RCTs, are only published in abstract form and therefore may be difficult to identify. A systematic review showed that about

35% of RCTs initially presented as abstracts were never published in full;[52] this was more likely if they were small, with non-significant results. As a result of publication bias, many people have called for worldwide prospective registration of randomised trials,[46,53] since retrospective identification of unpublished trials is unsatisfactory.[54] Three recent initiatives may help to bring about a comprehensive register of completed and ongoing trials: the Cochrane Collaboration has established a Controlled Trials Register using the specialised registers of trials developed by the Cochrane Review Groups;[55] some journals have undertaken to publish protocols of ongoing trials;[56] and a trial 'amnesty' has recently been announced in which authors of unpublished trials have been encouraged to send details of their trials to various journals for registration.[57,58] Even if the results of trials do not appear in print, trialists should be encouraged to make their results publicly available, for example, by providing the authorising ethics committee with a copy of the results or by publishing them in electronic format on a trials' register.

Recommendations

1. Trialists should submit the results of all RCTs for publication.
2. Trialists should make their results publicly available.
3. Efforts to establish prospective registration of RCTs should continue.

(Basis: all recommendations are supported by two systematic reviews of surveys of final publication status of RCTs in general and those initially published in conference proceedings.)

Multiple publications of the same trial

Failure to publish trial results causes problems but, paradoxically, so too does multiple publication of the same trial. Again, this is more common for trials showing statistically significant results.[50] Multiple reports can cause confusion since it may be unclear that they all relate to the same trial and there may be discrepancies in the results and conclusions of different reports.[59,60]

Recommendation

- Multiple publications of the same results should be avoided and journal editors should insist that any previous publications of the same trial are clearly referenced. (Basis: anecdotal evidence.)

Timeliness of publication

For effective dissemination of the results of RCTs, it is important that publication should occur as soon as possible after the results have been properly analysed. In particular, there are ethical reasons for large well-planned trials, designed to answer clinically important questions, to be fast-tracked for publication. No articles were identified that specifically addressed this issue. However, some journals have now instituted a rapid review policy for important articles, including RCTs.[61]

Recommendation

- More journals should adopt a fast-track publication policy for important RCTs. (Basis: anecdotal evidence.)

Conclusions

Randomised trials are usually reported poorly which makes it difficult to assess their internal and external validity. The conclusions may not always match the data presented (this may be more of a problem with commercially sponsored trials) and the discussion of the trial results can be biased because of failure to quote all the existing evidence. In addition, trial reports can exaggerate the benefits of treatment by concentrating on the relative rather than the absolute effects of treatment. A proportion of completed RCTs are never published (perhaps 5–15%); this is more likely to occur if the trial shows no differences or has been sponsored by the pharmaceutical industry. Small negative trials are also more likely to be published only as abstracts and so may be difficult to find. In contrast, some trials (particularly those with positive results) can lead to multiple publications which can cause confusion.

References

1. Liberati A, Himel HN, Chalmers TC. A quality assessment of randomized control trials of primary treatment of breast cancer. *J Clin Oncol* 1986;**4**:942–51.
2. Altman DG, Doré CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990;**335**: 149–53.
3. Balas EA, Austin SM, Ewigman BG, Brown GD, Mitchell JA. Methods of randomized controlled clinical trials in health services research. *Med Care* 1995;**33**:687–99.

4. Claessen JQ, Appelman CL, Touw-Otten FW, de Melker RA, Hordijk GJ. A review of clinical trials regarding treatment of acute otitis media [review]. *Clin Otolaryngol* 1992;**17**:251–7.

5. Hall JC, Mills B, Nguyen H, Hall JL. Methodologic standards in surgical trials. *Surgery* 1996;**119**:466–72.

6. Kleijnen J, Knipschild P. Niacin and vitamin B6 in mental functioning: a review of controlled trials in humans [review]. *Biol Psychiatry* 1991;**29**:931–41.

7. Kleijnen J, ter Riet G, Knipschild P. Acupuncture and asthma: a review of controlled trials [review]. *Thorax* 1991;**46**:799–802.

8. Kleijnen J, Knipschild P. Ginkgo biloba for cerebral insufficiency [review]. *Br J Clin Pharmacol* 1992;**34**:352–8.

9. Klein KB. Controlled treatment trials in the irritable bowel syndrome: a critique [review]. *Gastroenterology* 1988;**95**:232–41.

10. Koes BW, Bouter LM, van der Heijden GJ. Methodological quality of randomized clinical trials on treatment efficacy in low back pain. *Spine* 1995;**20**:228–35.

11. Lionetto R, Pugliese V, Bruzzi P, Rosso R. No standard treatment is available for advanced pancreatic cancer [review]. *Eur J Cancer* 1995;**31A**:882–7.

12. Marsoni S, Torri W, Taiana A, Gambino A, Grilli R, Liati P, *et al.* Critical review of the quality and development of randomized clinical trials (RCTs) and their influence on the treatment of advanced epithelial ovarian cancer. *Ann Oncol* 1990;**1**:343–50.

13. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994;**272**:122–4.

14. Nicolucci A, Grilli R, Alexanian AA, Apolone G, Torri V, Liberati A. Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer. A lost opportunity for meta–analysis. *JAMA* 1989;**262**:2101–7.

15. O'Donovan PA, Vandekerckhove P, Lilford RJ, Hughes E. Treatment of male infertility: is it effective? Review and meta-analyses of published randomized controlled trials. *Hum Reprod* 1993;**8**:1209–22.

16. Petersen LJ, Kristensen JK. Selection of patients for psoriasis clinical trials: a survey of the recent dermatological literature. *J Dermatol Treat* 1992;**3**:171–6.

17. Rossetti L, Marchetti I, Orzalesi N, Scorpiglione N, Liberati A. Is proper methodology associated with the use of a clinically relevant outcome measure? The case of randomized clinical trials on medical treatment of open-angle glaucoma. *Online J Curr Clin Trials* [serial online] 1993 Nov 11; Doc No 100: [6512 words; 199 paragraphs].

18. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994;**272**:125–8.

19. Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. *BMJ* 1996;**312**:742–4.

20. Silagy CA, Jewell D, Mant D. An analysis of randomized controlled trials published in the US family medicine literature, 1987–1991. *J Fam Pract* 1994;**39**:236–42.

21. Silagy CA, Jewell D. Review of 39 years of randomized controlled trials in the *British Journal of General Practice* [review]. *Br J Gen Pract* 1994;**44**:359–63.

22. Solomon MJ, Laxamana A, Devore L, McLeod RS. Randomized controlled trials in surgery. *Surgery* 1994;**115**:707–12.

23. Sonis J, Joines J. The quality of clinical trials published in the *Journal of Family Practice,* 1974–1991. *J Fam Pract* 1994;**39**:225–35.

24. Williams DH, Davis CE. Reporting of assignment methods in clinical trials. *Control Clin Trials* 1994;**15**:294–8.

25. Rochon PA, Fortin PR, Dear KBG, Minaker KL, Chalmers TC. Reporting of age data in clinical trials of arthritis: deficiencies and solutions. *Arch Intern Med* 1993;**153**:243–8.

26. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

27. Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians' willingness to prescribe. *Lancet* 1994;**343**:1209–11.

28. Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. *BMJ* 1994;**309**:761–4.

29. Forrow L, Taylor WC, Arnold RM. Absolutely relative: how research results are summarized can affect treatment decisions. *Am J Med* 1992;**92**:121–4.

30. Naylor CD, Chen E, Strauss B. Measured enthusiasm. Does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992;**117**:916–21.

31. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;**318**:1728–33.

32. Rochon PA, Gurwitz JH, Simms RW, Fortin PR, Felson DT, Minaker KL, *et al.* A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med* 1994;**154**:157–63.

33. Clarke M, Chalmers I. Discussion sections in reports of controlled trials published in five general medical journals: islands in search of continents. *JAMA* 1998;**280**:280–2.

34. Gøtzsche PC. Reference bias in reports of drug trials. *BMJ* 1987;**295**:654.

35. Ravnskov U. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome [published erratum appears in *BMJ* 1992;**305**:505]. *BMJ* 1992;**305**:15–19.

36. Waldron HA, Cookson RF. Avoiding the pitfalls of sponsored multicentre research in general practice. *BMJ* 1993;**307**:1331–4.

37. Julian D. Trials and tribulations. *Cardiovasc Res* 1994;**28**:598–603.

38. Andrew E, Anis A, Chalmers T, Cho M, Clarke M, Felson D, *et al.* A proposal for structured reporting of randomized controlled trials. *JAMA* 1994;**272**: 1926–31.

39. Rennie D. Reporting randomized controlled trials. An experiment and a call for responses from readers [editorial]. *JAMA* 1995;**273**:1054–5.

40. Asilomar Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Checklist of information for inclusion in reports of clinical trials. *Ann Intern Med* 1996;**124**:741–3.

41. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, *et al.* Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;**276**:637–9.

42. Peto R. Clinical trial reporting [letter]. *Lancet* 1996;**348**:894–5.

43. Warlow C, Sandercock P. Carotid endarterectomy vs carotid angioplasty [letter;]. *Lancet* 1997;**349**:880.

44. Marciniak TA. Recommendations for reporting clinical trials [letter]. *Ann Intern Med* 1997;**126**:411.

45. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement [editorial]. *BMJ* 1996;**313**:570–1.

46. Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990;**263**:1405–8.

47. Dickersin K, Min YI. Publication bias: the problem that won't go away. *Ann N Y Acad Sci* 1993;**703**: 135–48.

48. Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA* 1990; **263**:1392–5.

49. Parmley WW. Publication bias [editorial]. *J Am Coll Cardiol* 1994;**24**:1424–5.

50. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;**337**:867–72.

51. Davidson RA. Source of funding and outcome of clinical trials. *J Gen Intern Med* 1986;**1**:155–8.

52. Scherer RW, Dickersin K, Langenberg P. Full publication of results initially presented in abstracts. A meta-analysis [published erratum appears in *JAMA* 1994;**272**:1410]. *JAMA* 1994;**272**:158–62.

53. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;**263**: 1385–9.

54. Hetherington J, Dickersin K, Chalmers I, Meinert CL. Retrospective and prospective identification of unpublished controlled trials: lessons from a survey of obstetricians and pediatricians. *Pediatrics* 1989;**84**:374–80.

55. The Cochrane Collaboration. The Cochrane Controlled Trials Register [database]. In: Cochrane Library Issue 4, 1997. Oxford: Update Software.

56. McNamee D. Protocol reviews at *The Lancet*. *Lancet* 1997;**350**:6.

57. Horton R. Medical editors trial amnesty. *Lancet* 1997;**350**:756.

58. Smith R. An amnesty for unpublished trials. *BMJ* 1997;**315**:622.

59. de Bellefeuille C, Morrison CA, Tannock IF. The fate of abstracts submitted to a cancer meeting: factors which influence presentation and subsequent publication. *Ann Oncol* 1992;**3**:187–91.

60. Keirse MJ, Van Hoven M. Reanalysis of a multireported trial on home uterine activity monitoring. *Birth* 1993;**20**(3):117–22.

61. McNamee D. Fast-track to publication in *The Lancet*. *Lancet* 1997;**349**:970.

Chapter 8
Costs

Introduction

Costs affect the number, progress and quality of randomised trials in two basic ways.

1. In trials that seek to investigate both the effectiveness of a healthcare technology and its cost, and to compare these by economic analysis, there may be flaws in the trial that have a bearing on the study of these costs, or on the collection or analysis of relevant data; such flaws are likely to limit the economic quality of the trial and thus its value to those responsible for decisions about the provision of the technology.
2. Absence of resources to undertake trials is certain to reduce their number, and shortage of resources is likely to limit their progress or general quality or both.

Thus this chapter is concerned with both the study within or alongside trials of the costs of using a technology (described in this chapter and elsewhere as 'treatment' rather than 'technology' costs) and the reimbursement of the costs of conducting trials to evaluate that technology. The latter costs include both the treatment costs of patients within the trial and the marginal costs of conducting the research (described here and elsewhere as 'trial costs'); these trial costs include both direct research costs and 'service support costs', that is, the marginal cost of clinical services which are provided in the trial but which would not be provided in normal clinical practice.

The relevant publications relating to the costs of trials and evaluating the costs of treatment within trials are summarised in *Table 28*.

Study of treatment costs

Although the aim of all trials is to improve the welfare of patients in the long run, they have two different types of short-term aims – to influence healthcare decision making or to add to knowledge, for example, about the pharmacology of a drug. In a seminal paper,[1] Schwartz and Lellouch described knowledge-seeking trials as 'explanatory' and decision-oriented trials as 'pragmatic' and argued that more trials should be pragmatic and have some or all of the following features.

1. They should be conducted under normal clinical conditions rather than strict laboratory-like conditions.
2. They should be guided by a flexible protocol that permits the treatment to be adapted to the patient, thus optimising its placebo effect, rather than by a rigid protocol that equalises the placebo effect between treatments.
3. They should be analysed on an ITT basis rather than by excluding patients who diverge from the protocol.
4. They should be evaluated not by narrow biomedical criteria but by a single comprehensive criterion that reflects as wide a range as possible of the benefits and costs of the treatments being compared.

Although Hansson and colleagues[2] did not cite Schwartz and Lellouch,[1] they developed the analogous argument that double-blind, placebo-controlled trials are unrealistic and expensive, and recommend that, in order to minimise bias, only the assessors of patient outcomes need to be blinded to the allocated treatment.

The case for pragmatic trials in general, and for the evaluation of benefits and costs in particular, was largely neglected in the UK until the need for health services research and development was recognised, notably by the Medical Research Council gradually throughout the 1980s and by the NHS suddenly from 1991. Since then the case for including treatment costs as an end-point in trials has been cogently argued by an increasing number of authors.[3–10] Nevertheless, in the majority of trials no attempt is made still to address economic issues; in particular, only 5% of recent surgical trials analysed costs or resources.[10]

Although the case for economic analysis within or alongside trials is increasingly accepted, there are many practical problems to be overcome. Thus, many publications reviewed in this study recommend that clinicians considering the inclusion of economic analysis in their trials should work closely with health economists.[6,7,11,12] The

TABLE 28 *Publications concerning costs of trials (limiting number and progress) and evaluating costs of treatment with trials (limiting quality)*

Reference	Title	Classification	Design	Summary of paper
Anderson, *et al.,* 1995[13]	Recruiting from the community: lessons from the Diabetes Care for Older Adults Project	Trial costs (recruitment)	Case study of RCT	Compares effectiveness and cost of concurrent community-based recruitment methods. Press releases and newspaper advertisements were most effective (providing 80% of subjects) and cost-effective ($37 per subject). In contrast, posters and brochures cost $900 per subject recruited.
Bennett, *et al.,* 1994[6]	Economic analyses of clinical trials in cancer: are they helpful to policy makers?	Economic analysis (treatment costs)	Case study of RCT	Reports Phase III trial of haematopoietic growth factors at two hospitals in Paris and New York; the two centres had similar clinical results but very different economic results. Under these circumstances one should report on each site separately rather than pool data. In future, more sites or larger samples at each site should be used.
Bennett, *et al.,* 1994[14]	Economic analysis in Phase III cancer trials	Economic analysis	Economic theory	Contrasts clinical and economic goals; expounds economic aspects of each stage of a clinical trial – study design, data collection (both cost data and extra data needed for economic analysis), outcome measures, analysis and interpretation. Concludes that to follow standard Phase III trials by cost-focused trial is inefficient.
Bigorra & Banos, 1990[15]	Financial reward in the decision by medical students and experienced volunteers to participate in clinical trials	Trial costs (recruitment)	Questionnaire surveys	Survey of medical students showed that about one-third would not be willing to take part in a trial of 'a new drug with clear therapeutic potential', about one-third would do so for scientific reasons, and the rest would need a financial incentive. Survey of experienced volunteers showed that the main incentive for 90% was financial and that, given this, they were willing to take part again.
Bjornson-Benson, *et al.,* 1993[16]	Monitoring recruitment effectiveness and cost in a clinical trial	Trial costs (recruitment)	Case study of RCT	Compares cost per participant recruited of five concurrent recruitment methods in a trial of lung disease treatment. The media were most cost-effective, followed by workplace recruitment, referral by other participants, direct mail and neighbourhood recruitment. These results are not generalisable as continuous feedback allowed redirection of resources between the five methods.
Bonsel, *et al.,* 1993[4]	Economic evaluation alongside cancer trials: methodological and practical aspects	Economic analysis	Economic theory	Describes cost-effectiveness analysis and analyses advantages (convenience, cost, time) and disadvantages of combining economic analysis with a clinical trial: patients, staff and methods may be unrepresentative; follow-up may be too brief; outcome measures may not include quality of life or utility.
Bothani, *et al.,* 1989[17]	Recruitment in the Hypertension Prevention Trial	Trial costs (recruitment)	Case study of RCT	Analyses recruitment in a four-centre trial. Most effort was put into direct mailing: 220,000 brochures generated 10,000 (85%) of 12,000 initial contacts. Newspapers were only other substantial source. Tabulates procedures and costs during each step of reducing these contacts to final 841 participants, averaging $950 per participant. Total costs closely depend on the order of these steps.
Buxton & Hannay, 1996[18]	How can payback from health services research be assessed?	Trial benefits	Economic theory and eight case studies	Proposes nine-stage input–output model of research utilisation: (0) Research needs assessment (a) Project definition (I) Inputs (II) Processes (III) Primary outputs (b) Dissemination (IV) Secondary outputs (V) Applications (VI) Payback – knowledge, research benefits, policy benefits, health benefits and economic benefits. Examines feasibility of applying model using eight case studies.
Cohen, *et al.,* 1995[19]	Requirements for controlled clinical trials of preoperative CV risk reduction	Trial benefits and costs	Case study of RCT	Because of rarity of adverse outcome, trials of cardiac risk assessment and reduction before non-cardiac surgery are expensive. If such a trial showed that present high-cost strategy could safely be modified, however, potential savings would also be large. Hence funding organisations should support such trials.

continued

TABLE 28 contd *Publications concerning costs of trials (limiting number and progress) and evaluating costs of treatment with trials (limiting quality)*

Reference	Title	Classification	Design	Summary of paper
Detsky, 1989[20]	Are clinical trials a cost-effective investment?	Trial benefits and costs	Non-systematic review	Reports cost-effectiveness of seven randomised trials ranging from $2–3 to $400–700 per life-year saved – much better than that of many established treatments or preventive measures; trials range from bypass surgery ($5000) to liver transplant ($250,000).
Diekmann & Smith, 1989[21]	Strategies for accessment and recruitment of subjects for nursing research	Trial costs (recruitment)	Non-systematic review	Reviews strategies for obtaining access through professionals, community and media, and for recruiting: reports on Penman *et al.* (1984)[22] on cost per subject recruited by personal contact ($90), media ($58) and mail, etc. ($42); and Baines (1984)[23] on telephone follow-up after letter, which reduced cost per response from $9 to $6.40.
Drummond & Davies, 1991[24]	Economic analysis alongside clinical trials: revisiting the methodological issues	Economic analysis	Economic theory	Reviews four methodological issues: (i) trial design, viz. choice of therapies, sample size, location; (ii) collecting data on resource use; (iii) choice of outcome measures, including disease-specific vs. general profiles vs. utility measures; (iv) interpretation and extrapolation of results. Makes seven clear recommendations and recognises need for further research and debate to resolve these issues.
Drummond, *et al.,* 1992[25]	Funding research and development in the NHS	Trial costs	Economic theory	Discusses effect of purchaser–provider split on research and development: financial pressure may encourage inappropriate use of research money for routine treatment of trial patients or quality assurance. Suggests strategies to maintain quality of research and development.
Drummond & O'Brien, 1993[5]	Clinical importance, statistical significance and the assessment of economic and quality-of-life outcomes	Economic analysis	Economic theory	Observes that most economic studies are deterministic. However, costs and cost-effectiveness ratios are subject to random variation and need statistical tests and CIs. Economic as well as clinical outcomes should be considered when fixing sample size but an arbitrary 'minimum important change' of 10% or 15% from baseline is no more justified for economic than for clinical outcomes.
Drummond, 1994[7]	Economic analysis alongside controlled trials: an introduction for clinical researchers	Economic analysis	Economic theory	This guide, commissioned by the NHS R&D Directorate for clinical researchers contemplating complementary economic analysis, has four main chapters. Which trials need economic analysis? Which form of economic analysis? How should economic data be collected? How should they be analysed?
Drummond, 1995[9]	Economic analysis alongside clinical trials: problems and potential	Economic analysis	Economic theory	Whether to undertake concurrent economic analysis of a trial depends on its quality, economic importance of the question, additional cost (especially in international trials), and whether trial conditions are realistic and generalisable. Extra data is needed on direct and indirect benefits and costs of treatment, including quality of life, either as part of trial or by abstracting other data, typically routine.
Easterbrook & Matthews, 1992[26]	Fate of research studies	Trial costs (funding)	Telephone interview survey	Reports history of 487 trial protocols approved by Central Oxford Research Ethics Committee. Of 100 trials not started, 40 had not obtained funding; of 58 abandoned, six had stopped because funding had been withdrawn.
Fetter, *et al.,* 1989[3]	Randomised clinical trials: issues for researchers	Trial and treatment costs	Non-systematic review	Introduces basic issues in RCTs. Distinguishes between cost as outcome measure in economic analysis and as issue in trial design: indirect costs of trial must be considered; multicentre trials are not necessarily more expensive.
Fleming, 1994[27]	Barriers to clinical trials: I – reimbursement problems	Trial and treatment costs	Opinion	Research costs in cancer trials in USA are always paid by sponsor, while 'usual care' but not extra care due to trial is traditionally borne by health insurers. Several insurers are now refusing trial patients even 'usual care'. When care contracts are awarded on lowest tender, research activity is ignored and therefore discouraged.

continued

TABLE 28 contd *Publications concerning costs of trials (limiting number and progress) and evaluating costs of treatment with trials (limiting quality)*

Reference	Title	Classification	Design	Summary of paper
Foley & Moertel, 1991[28]	Improving accrual into cancer trials	Trial costs (to patients)	Questionnaire survey	Survey of health professionals in USA identified many obstacles to patient recruitment, notably extra costs to patients, and recommended that drugs and extra research tests should be subsidised nationally rather than charged to patient.
Hall, *et al.,* 1996[10]	Methodologic standards in surgical trials	Treatment costs	Systematic review	Reviews 346 surgical trials published in major surgical or medical journals, 1988–94: only 19 (5%) analysed costs or resource use and only six (2%) reported formal measures of quality of life.
Hansson, *et al.,* 1992[2]	Prospective randomised open blinded endpoint (PROBE) study: novel design for trials	Trial costs	Methodology paper	Shows that large trials are costly and suggests 'new' study design to reduce costs by increasing realism. Blinding only endpoint assessor replaces double-blind drug administration by cheaper standard prescribing and reduces other trial costs. Argues that only drawback is susceptibility to investigator bias.
Ho, 1994[29]	The future direction of clinical trials	Trial costs (staff)	Opinion and two case studies of RCTs	Reports that American Cancer Society has used volunteers in several epidemiological studies; in 1990 it began two trials staffed mainly by volunteers, including physicians and nutritionists. Suggests that high cost of trials can be reduced by using volunteers without compromising reliability; however, does not estimate savings.
Johansen, *et al.,* 1991[30]	Obstacles to implementing cancer clinical trials	Trial cost (to patients)	Non-systematic review	Describes who funds cancer trials in USA – National Cancer Institute, drug industry, health insurers and patients themselves – and what is covered. Reports that patients are deterred by extra expense from taking part and suggests how nurses can help to overcome these and other obstacles.
King, *et al.,* 1994[31]	Effect of recruitment strategy on types of subjects entered into a primary prevention trial	Trial costs (recruitment)	Case study of RCT	Reports that random telephone dialling recruited 214 participants and media campaign 143. Although cost per person assessed was slightly lower with random dialling, cost per person randomised was $168 compared with $70 for media campaign. However, more high-risk subjects were recruited, which may compensate for extra cost.
Lawrence, *et al.,* 1993[32]	The impact of clinical trial protocols on patient care systems	Trial costs (to patients)	Opinion	Reports that health insurers in USA may refuse to reimburse patients for treatment costs if they participate in trial and discusses how this might be avoided by careful drafting of trial protocols. Also suggests trials of prevention which could benefit poor or other uninsured Americans.
Mansour, 1994[33]	Barriers to clinical trials: III – knowledge and attitudes of health care providers	Trial costs (to patients and doctors)	Opinion	Discusses range of barriers including: (i) reimbursing participating doctors does not cover time spent recruiting patients, which apparently averages 4 hours per recruit; (ii) many health insurers will not reimburse for trials; there is even inconsistency between insurers over what is covered.
NHS Research & Development Task Force, 1994[34]	Supporting research and development in the NHS [the 'Culyer Report']	Trial and treatment costs	Report of UK Government committee	Set up to review funding of research and development and associated service costs within NHS, the Task Force makes 22 recommendations under three main headings: (i) determining what research should be supported; (ii) funding mechanisms, including proposal that there should be a single explicit stream to fund three main activities – research and development projects, service support and research facilities; (iii) costing and accounting.
Oddone, *et al.,* 1995[35]	Measuring activities in clinical trials using random work sampling: implications for analysis	Trial and treatment costs	Case study of RCT	Describes use of portable random alarms by nurses with dual roles in trial: they recorded their current activity when alarm sounded. On average they spent 42% of their time on trial activities and 58% caring for intervention patients. Thus, in economic analysis, 58% of their salaries was attributed to the intervention.

continued

TABLE 28 contd *Publications concerning costs of trials (limiting number and progress) and evaluating costs of treatment with trials (limiting quality)*

Reference	Title	Classification	Design	Summary of paper
Schwartz & Lellouch, 1967[1]	Explanatory and pragmatic attitudes in therapeutic trials	Trial design	Methodology paper	Argues that most trialists fail to distinguish between: (i) explanatory trials to understand differences between rigidly defined treatments under controlled conditions using biological criteria; and (ii) pragmatic trials to decide between flexible treatments under clinical conditions using clinical, social and economic criteria.
Schwartz, 1995[11]	PIVOT: economic analysis of study design in conditions of uncertainty	Economic analysis	Case study of RCT	Advocates economic analysis to decide between alternative study designs and to consider treatment costs when determining sample size; for example, accurate staging of prostate cancer increases both trial costs and generalisability – economic modelling can decide which is more valuable. However, problems exist in eliminating protocol-induced costs from estimation of true costs for economic analysis.
Silagy, *et al.,* 1991[36]	Comparison of recruitment strategies for a large-scale clinical trial in the elderly	Trial costs (recruitment)	Comparative study	Reports cost-effectiveness of three methods of recruiting set number of patients: mail using family practice registers recruited 1/6 of those approached ($48 per patient); mail using electoral roll recruited 1/17 ($59 per patient); community approach was cheapest ($43) but only 1/29 recruited.
Smyth, *et al.,* 1994[37]	Conducting clinical research in new NHS: the model of cancer	Trial costs	Questionnaire survey	Describes costs of conducting research – direct excess service costs, infrastructure costs, central organisational costs and general costs including indemnity. Concludes with eight recommendations for meeting these costs and encouraging research.
Steward, *et al.,* 1993[38]	Chemotherapy administration and data collection in an EORTC collaborative group	Trial costs	Case study of RCT	Describes quality control system of large multicentre trial and justifies cost of occasional site visits by reference to resulting improvements in quality.
Swanson & Ward, 1995[39]	Recruiting minorities into clinical trials: towards a participant-friendly system	Trial costs (recruitment)	Systematic review	Notes that recruitment to trials is expensive and strategies to target minorities are likely to be even more expensive for both trial and patients. Recommends that definition of trial costs should be extended to include transport, meals, child care and social services for patients, and educational materials for use by trial team.
Tannock, 1994[8]	New perspectives in combined radiotherapy and chemotherapy treatment	Economic analysis	Non-systematic review	Argues that combined therapy should be evaluated against standard treatment in large randomised trials. Proposes guidelines for conduct of such trials: in particular, small improvements in quality or quantity of life should be converted into quality-adjusted life-years and compared with economic cost of achieving them.
Tannock, 1995[40]	The recruitment of patients into clinical trials	Trial and treatment costs	Opinion	Identifies an expensive breast cancer treatment widely used in USA with no evidence of increased survival; some health insurers will fund it only within trial. Asserts that good trials are cheaper than uncritical adoption of unproven treatments.
Tilley, *et al.,* 1990[41]	Designing clinical trials of treatment for osteoporosis: recruitment and follow-up	Trial costs (recruitment and to patients)	Case study of RCT	Identifies reasons for poor recruitment including: (i) as screening for eligibility was underfunded, patients were pre-screened at own expense by their own doctors; (ii) once women were in trial, costs not covered by insurance met from trial funds; however, they were concerned that treatment costs would exceed annual limit on insurance claims, leaving them without cover for rest of year.
Torgerson, *et al.,* 1995[12]	Economics in sample size determination for clinical trials	Economic analysis	Economic theory	Three case studies illustrate use of data, on cost as well as effectiveness, to determine effect size that trials should aim to identify: a cheap measure to prevent osteoporosis justifies a large trial, while a smaller study is needed to investigate *in vitro* fertilisation. Thus inclusion of simple economic analysis in trial design can lead to more efficient allocation of scarce research resources.

continued

TABLE 28 contd *Publications concerning costs of trials (limiting number and progress) and evaluating costs of treatment with trials (limiting quality)*

Reference	Title	Classification	Design	Summary of paper
Welsh, *et al.*, 1994[47]	Issues affecting minority participation in research studies of Alzheimer disease	Trial costs (to patients)	Non-systematic review	Notes that minorities have lower incomes, less insurance cover and less chance of being treated in tertiary care, where most trials take place. Recommends use of incentives to ensure balanced participation in trials, notably assessment in community or reimbursement of transport costs.
Wineman & Durand, 1992[43]	Incentives and rewards for subjects in nursing research	Trial costs (to patients)	Case study of interview survey	Argues that financial incentives can encourage participation in trials but effect may be coercive if amount is too large. Authors used letter emphasising personal choice and offering $15 per interview to achieve response rates of 65% in surveys of multiple sclerosis and spinal cord injury.
Winn, 1994[44]	Obstacles to the accrual of patients to clinical trials in the community	Trial costs	Non-systematic review	Notes that community trials are particularly likely to be affected by reluctance of insurers to pay patient costs. Minimising protocol costs by using only essential tests can reduce costs by 60% but trialists should continue to pursue third-party reimbursement, as care within trials is usually best available treatment.

methodological issues to be addressed by the resulting multi-disciplinary teams can be grouped under four main headings: trial design, collection of data on resource use, choice of patient outcome measures, and the interpretation and extrapolation of findings.[24] All of the seven basic recommendations made by Drummond and Davies,[24] and listed below, are reinforced by at least three other peer-reviewed papers.

1. Treatments to be compared should be those likely to be considered by future decision makers.[1,2,11]
2. Because resource use may be more or less variable than patient outcomes, sample size calculations should consider economically important effect sizes, notably in costs and patient utilities, as well as clinically important effect sizes.[5,11,12]
3. The estimation of treatment costs should exclude all trial costs, although this is rarely easy.[3,11,35]
4. The reporting of treatment costs should separate the quantities of resources consumed from the prices paid.[6,9,14]
5. Patient outcome measures should be rigorously chosen and should generally include clinical measures, disease-specific measures of quality of life, and generic measures of patient utility.[4,8,10]
6. If the economic value of different technologies is to be compared across trials, findings about resources and outcomes should be reported in standard form; for example, net direct costs (i.e. direct costs minus direct financial benefits) and net gain in quality-adjusted life years.[4,7,8]
7. Extrapolating the findings of the trial to other settings should generally be based on economic modelling that takes account of differences between trial centres and uses data from other centres, for example, on treatment costs.[4,9,37]

Reimbursement of treatment costs

In April 1998 a new system was introduced in the UK to ensure that both treatment and service support costs (as defined in chapter 1) of peer-reviewed trials are met by the NHS, thus implementing the recommendations of the Research and Development Task Force (the 'Culyer report').[34] Although the effectiveness of this system in overcoming previous financial barriers to trials[25,37] is untested, many papers identified in this review have been overtaken by this development and thus are not reported here. Given the level of concern expressed in these papers, however, monitoring the success of the Culyer reforms in overcoming financial barriers to trials is recommended. To improve the quality of trials, it is suggested that NHS trusts should acquire access to (or establish) trials offices.

Most of the remaining literature on treatment costs discusses American studies. Cost to patients is a major theme, because it can be a major deterrent to recruitment.[30,39,41,42] In particular, while insurers have traditionally funded 'usual care' of patients in trials, costs of extra tests necessary for the trial are met by the patient unless there is provision for this in the trial budget;[27,28] even

when the cost of usual care is met, patients with managed-care contracts may exceed their allocated budget.[44] Several insurers now refuse to reimburse even usual care.[27,32,33]

In response to these financial barriers, many authors argued for mechanisms to ensure that patients participating in trials are fully reimbursed for all their resulting costs, and some authors even proposed stipends or fees for participants.[15,39,43] While some proposed that these patient costs should be fully covered by research funds,[39,42] others advocated national initiatives as far-reaching as the Culyer reforms.[27,28] Winn proposed that health insurers should meet the full cost of routine care in trials, since this represents the best treatment for most patients.[44] Taking this argument farther, Tannock in Canada[40] and Smyth and colleagues in the UK[37] urged that unproven treatments should be funded only for patients in peer-reviewed trials. Despite the differences between the USA and the UK, the authors consider that concerns about patient recruitment are universal and recommend that serious consideration be given in the UK to reimbursing trial patients for the full cost of taking part.

Trial costs

The Culyer report also recommended that research which is not worth supporting should be stopped.[34] Although the effectiveness of the new system in eliminating poor or repetitive research is untested, some papers identified in this review have been overtaken by this proposal and, hence, are not reported here. Again, monitoring the success of the Culyer reforms in eliminating research that is not worth supporting is recommended.

Relevant to this recommendation is the survey of 487 planned studies by Easterbrook and Matthews;[26] of 100 studies that never started only 40 lacked funding, and of 58 studies that were abandoned only six did so because their funding was withdrawn. If the Culyer reforms were successful, a future survey might look for an increase in the proportion of sound proposals and a reduction in the proportion of flawed proposals that gained funding.

Most of the remaining literature on trial costs discusses American studies. Several authors have compared the effectiveness of different methods of recruiting patients – usually including mail and the media – simultaneously within the same trial (see *Table 20*). Many of these authors also studied cost-effectiveness.[13,16,31,36] Although these studies, together with two others that focused on cost-effectiveness,[17,21] are the only comparative studies reported within this chapter, no consistent findings emerged. This can be attributed to the wide range of topics and settings, and the general lack of rigour in the comparisons, none of which were experimental.

Of the remaining papers, two were keen to reduce trial costs: Winn by reimbursing only those tests relevant to the trial's major objectives,[44] and Ho by using volunteer staff.[29] In contrast, Mansour was concerned that doctors are not reimbursed for all the time spent recruiting patients.[33] Steward considered that the cost of quality assurance in trials, and regular site visits in particular, can be justified by the resulting improvements in quality.[38]

Conclusion and recommendations

It may, in due course, be possible to assess many of the proposals reported in this chapter by reference to sophisticated models of the payback from health technology assessment, such as that described by Buxton.[18] In the meantime, the authors note that, on the basis of selected case studies, Detsky[20] and Cohen and colleagues[19] claimed that trials are more cost-effective than established treatment or prevention programmes.

All but two of the following recommendations are derived from at least two methodological or theoretical papers, systematic reviews or competent questionnaire surveys. Although recommendation 12 is derived mainly from a single questionnaire survey, and recommendation 14 from only a single non-systematic review, each is consistent with the basic philosophy of the Culyer report. The first ten recommendations relate to trials that seek to influence healthcare decision making in the short term.

Recommendations

1. Consider whether the primary aim of the trial should not be pragmatic,[1] in particular:
   (a) whether it should not faithfully reflect normal clinical conditions
   (b) whether it should not optimise the placebo effect of each treatment, notably by avoiding the additional costs incurred by a double-blind, placebo-controlled trial[2]
   (c) whether it should not be analysed using ITT

(d) whether the primary criterion should not be a single comprehensive criterion that reflects as wide a range as possible of the benefits and costs of the treatments being compared.

2. More specifically, treatment cost should be included as an end-point.[3–10]
3. Clinicians who seek to implement recommendation 2 should work closely with health economists.[6,7,11,12]
4. Treatments should be compared in the form in which they are likely to be considered by future decision-makers.[1,2,11,24]
5. In sample size calculations, both economically important as well as clinically important effect sizes should be considered.[5,11,12,24]
6. All trial costs should be excluded when estimating treatment costs.[3,11,24,35]
7. In reporting treatment costs, the quantities of resources consumed should be separated from the prices paid.[6,9,14,24]
8. Patient outcome measures should be chosen rigorously and include clinical measures, disease-specific measures of quality of life and generic measures of patient utility.[4,8,10,24]
9. Findings about resources and outcomes should be reported in the form of net direct costs (i.e. direct costs minus direct financial benefits) and net gain in quality-adjusted life-years.[4,7,8,24]
10. The extrapolation of trial findings to other settings should be based on economic modelling that takes account of differences between trial centres.[4,9,24,37]
11. The success of the Culyer reforms, in overcoming financial barriers to trials and in eliminating research that is not worth supporting, should be monitored.[25,37]
12. Unproven treatments should only be funded for patients in peer-reviewed trials.[37,40]
13. Serious consideration should be given to reimbursing trial patients for the full cost of taking part.[27,28,39,42,44]
14. Only those tests relevant to the trial's major objectives should be reimbursed.[44]

References

1. Schwartz D, Lellouch D. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;**20**:637–45.
2. Hansson L, Hedner T, Dahlof B. Prospective randomized open blinded end-point (PROBE) study. A novel design for intervention trials. *Blood Press* 1992;**1**:113–19.
3. Fetter MS, Feetham SL, d'Apolito K, Chaze BA, Fink A, Frink BB, *et al.* Randomized clinical trials: issues for researchers. *Nurs Res* 1989;**38**:117–20.
4. Bonsel GJ, Rutten FF, Uyl-de Groot CA. Economic evaluation alongside cancer trials: methodological and practical aspects. *Eur J Cancer* 1993;**29A** (suppl 7):S10–14.
5. Drummond M, O'Brien B. Clinical importance, statistical significance and the assessment of economic and quality-of-life outcomes. *Health Econ* 1993;**2**:205–12.
6. Bennett CL, Armitage JL, LeSage S, Gulati SC, Armitage JO, Gorin NC. Economic analyses of clinical trials in cancer: are they helpful to policy makers [review]? *Stem Cells* 1994;**12**:424–9.
7. Drummond MF. Economic analysis alongside clinical trials: an introduction for clinical researchers. London: Department of Health; 1994.
8. Tannock IF. New perspectives in combined radiotherapy and chemotherapy treatment [review]. *Lung Cancer* 1994;**10**(suppl 1):S29–51.
9. Drummond M. Economic analysis alongside clinical trials: problems and potential [review]. *J Rheumatol* 1995;**22**:1403–7.
10. Hall JC, Mills B, Nguyen H, Hall JL. Methodologic standards in surgical trials. *Surgery* 1996;**119**:466–72.
11. Schwartz JS. Prostate cancer intervention versus observation trial: economic analysis in study design and conditions of uncertainty. Oxford: Oxford University Press; National Cancer Institute, Monograph no.19; 1995. p.73–5.
12. Torgerson DJ, Ryan M, Ratcliffe J. Economics in sample size determination for clinical trials. *QJM* 1995;**88**:517–21.
13. Anderson LA, Fogler J, Dedrick RF. Recruiting from the community: lessons learned from the diabetes care for older adults project. *Gerontologist* 1995;**35**:395–401.
14. Bennett CL, Armitage JL, Buchner D, Gulati S. Economic analysis in Phase III clinical cancer trials. *Cancer Invest* 1994;**12**:336–42.
15. Bigorra J, Banos JE. Weight of financial reward in the decision by medical students and experienced healthy volunteers to participate in clinical trials. *Eur J Clin Pharmacol* 1990;**38**:443–6.
16. BjornsonBenson WM, Stibolt TB, Manske KA, Zavela KJ, Youtsey DJ, Buist AS. Monitoring recruitment effectiveness and cost in a clinical trial. *Control Clin Trials* 1993;**14**:52–67S.
17. Borhani NO, Tonascia J, Schlundt DG, Prineas RJ, Jefferys JL. Recruitment in the hypertension prevention trial (Hypertension Prevention Trial Research Group). *Control Clin Trials* 1989; **10**(3 suppl):30–9S.

18. Buxton M, Hannay S. How can payback from health services research be assessed? *J Health Services Res Policy* 1996;**1**:35–43.

19. Cohen MC, McKenna C, Lewis SM, Muller JE. Requirements for controlled clinical trials of preoperative cardiovascular risk reduction. *Control Clin Trials* 1995;**16**:89–95.24. Drummond MF, Davies L. Economic analysis alongside clinical trials: revisiting the methodological issues. *Int J Technol Assess Health Care* 1991;**7**:561–73.

20. Detsky AS. Are clinical trials a cost-effective investment? *JAMA* 1989;**262**:1795–800.

21. Diekmann JM, Smith JM. Strategies for accessment and recruitment of subjects for nursing research. *West J Nurs Res* 1989;**11**:418–30.

22. Penman DT, Holland JC, Bahna GF, Morrow G, Schmale AH, Derogatis LR, *et al.* Informed consent for investigational chemotherapy: patient and physician perspectives. *J Clin Oncol* 1984;**2**:849–55.

23. Baines CJ. Impediments to recruitment in the Canadian breast screening study: response and resolution. *Control Clin Trials* 1984;**5**:129–40.

24. Drummond MF, Davies L. Economic analysis alongside clinical trials: revisiting the methodological issues. *Int J Technol Assess Health Care* 1991;**7**:561–73.

25. Drummond MF, Crump BJ, Little VA. Funding research and development in the NHS. *Lancet* 1992;**339**:230–1.

26. Easterbrook PJ, Matthews DR. Fate of research studies. *J R Soc Med* 1992;**85**:71–6.

27. Fleming ID. Barriers to clinical trials: Part I. Reimbursement problems [review]. *Cancer* 1994;**74**(9 suppl):2662–5.

28. Foley JF, Moertel CG. Improving accrual into cancer clinical trials. *J Cancer Educ* 1991;**6**:165–73.

29. Ho RC. The future direction of clinical trials [review]. *Cancer* 1994;**74**(9 suppl):2739–44.

30. Johansen MA, Mayer DK, Hoover HC Jr. Obstacles to implementing cancer clinical trials. *Semin Oncol Nurs* 1991;**7**:260–7.

31. King AC, Harris RB, Haskell WL. Effect of recruitment strategy on types of subjects entered into a primary prevention clinical trial. *Ann Epidemiol* 1994;**4**:312–20.

32. Lawrence W Jr, Antman K, Freeman HP, Huber SL, Kaufman D, Lantos J, *et al.* The impact of clinical trial protocols on patient care systems. *Cancer* 1993;**72**(9 suppl):2839–41.

33. Mansour EG. Barriers to clinical trials: Part III. Knowledge and attitudes of health care providers. *Cancer* 1994;**74**:2672–5.

34. NHS Research and Development Task Force. Supporting research and development in the NHS. London: HMSO; 1994.

35. Oddone E, Weinberger M, Hurder A, Henderson W, Simel D. Measuring activities in clinical trials using random work sampling: implications for cost-effectiveness analysis and measurement of the intervention. *J Clin Epidemiol* 1995;**48**:1011–18.

36. Silagy CA, Campion K, McNeil JJ, Worsam B, Donnan GA, Tonkin AM. Comparison of recruitment strategies for a large-scale clinical trial in the elderly. *J Clin Epidemiol* 1991;**44**:1105–14.

37. Smyth JF, Mossman J, Hall R, Hepburn S, Pinkerton R, Richards M, *et al.* Conducting clinical research in the new NHS: the model of cancer (UK Coordinating Committee on Cancer Research). *BMJ* 1994;**309**:457–61.

38. Steward WP, Vantongelen K, Verweij J, Thomas D, van Oosterom AT. Chemotherapy administration and data collection in an EORTC collaborative group – can we trust the results? *Eur J Cancer* 1993;**29A**:943–7.

39. Swanson GM, Ward AJ. Recruiting minorities into clinical trials: toward a participant-friendly system. *J Natl Cancer Inst* 1995;**87**:1747–59.

40. Tannock IF. The recruitment of patients into clinical trials [editorial; comment]. *Br J Cancer* 1995;**71**:1134–5.

41. Tilley BC, Peterson EL, Kleerekoper M, Phillips E, Nelson DA, Shorck, MA. Designing clinical trials of treatment for osteoporosis: recruitment and follow-up [editorial]. *Calcif Tissue Int* 1990;**47**:327–31.

42. Welsh KA, Ballard E, Nash F, Raiford K, Harrell L. Issues affecting minority participation in research studies of Alzheimer disease [review]. *Alzheimer Dis Assoc Disord* 1994;**8**(suppl 4):38–48.

43. Wineman NM, Durand E. Incentives and rewards for subjects in nursing research. *West J Nurs Res* 1992;**14**:526–31.

44. Winn RJ. Obstacles to the accrual of patients to clinical trials in the community setting. *Semin Oncol* 1994;**21**(suppl 7):112–17.

Chapter 9

Recommendations for practice

All the recommendations from earlier chapters are brought together here. The structure of this report is such that some issues have arisen in several chapters; hence, there may be overlap between the recommendations. In such situations, the initial proposals have either been combined into composite recommendations or been restated, as appropriate. The basis for each recommendation is restated and each recommendation is cross-referenced to where it occurred earlier.

Most of the recommendations are directed at trialists but some are directed primarily at other groups, such as journal editors. The structure followed in this chapter reflects the target audience for the recommendations, and the original recommendations have sometimes been reworded to reflect this.

Recommendations for trialists

The research question

1. The questions addressed by RCTs should be important enough to clinicians for them to be comfortable with the research role and be willing to take part and comply with protocol requirements (chapter 4, pages 29 and 31). (Basis: anecdotal evidence and logical argument.)
2. The choice of trial treatments should take into account of findings from previous trials, with a critical assessment of whether the choice will allow the main objectives of the trial to be achieved (chapter 3, page 14). (Basis: logical argument.)

Clinicians

3. The value of RCTs as a 'risk minimising' strategy, when there is uncertainty, should be emphasised to clinicians (chapter 4, page 29). (Basis: judgement of the authors.)
4. Clinicians should be recruited who understand and are in agreement with the proposed research protocol (chapter 4, page 33). (Basis: logical argument.)
5. Clinicians need preparation and appropriate support staff if they are to participate in RCTs (chapter 4, page 29). (Basis: anecdotal evidence).
6. Clinicians should be asked to do the minimum required from them, with other activities, such as follow-up, performed by specifically funded and employed staff (chapter 4, page 29). (Basis: logical argument.)
7. Clinicians should be rewarded appropriately and adequately for taking part in RCTs. The rewards need not be financial but should include positive feedback and support (chapter 4, page 29). (Basis: anecdotal evidence.)
8. Contributions to RCTs should be credited, for example, in all publications (chapter 4, page 29). (Basis: judgement of the authors.)

Trialists

9. There is a need for improvement in the statistical training of trialists (chapter 6, page 87). (Basis: logical reasoning.)
10. Inexperienced trialists should be supported by experienced trialists (chapter 5, page 76). (Basis: anecdotal evidence.)

Planning

11. Protocols should be designed to avoid incompatibility with normal practice (chapter 4, pages 29 and 33). (Basis: anecdotal evidence and logical argument.)
12. The parallel group design with a fixed sample size and fixed treatment schedules remains the most commonly used RCT design. Inexperienced trialists should not, however, automatically adopt this design, as other designs can sometimes produce major benefits. The views of experienced trialists should be sought whenever possible (chapter 3, page 15). (Basis: logical reasoning.)
13. Entry criteria should be as simple and clear as possible so that the study will accommodate all relevant patients (chapter 4, page 35). (Basis: anecdotal evidence.)
14. Permissive inclusion/exclusion criteria are generally preferred, to allow both faster trial accrual and a more representative trial population. Exceptions, when there may be an advantage in using a more homogeneous population, are for expensive or toxic treatments, or for hazardous investigations which might only be justified in high-risk groups (chapter 3, page 12). (Basis: logical argument.)

15. A trial should be designed to recruit a representative group of patients. This should be monitored carefully to ensure later generalisability of the study findings (chapter 4, page 57). (Basis: logical argument.)
16. Clinicians entering patients into multicentre trials should be chosen to give representative patient populations, subject to them having the relevant skills and resources to administer the trial treatments and procedures, and having an adequate throughput of appropriate patients (chapter 3, page 12). (Basis: logical argument.)
17. Multicentre trials should not be restricted to expert academic centres (chapter 5, page 69). (Basis: case studies of two cancer trials' networks.)
18. Good communication is essential especially in multicentre trials (chapter 5, page 76). (Basis: anecdotal evidence.)
19. Clinicians are likely to find pragmatic RCT protocols easier to use because they test interventions as used in 'everyday' care (chapter 4, page 33). (Basis: logical argument.)
20. Pragmatic RCTs are likely to be more acceptable to clinicians because this type of design permits more clinician freedom (chapter 4, page 33). (Basis: judgement of the authors.)

Run-in periods/compliance

21. A run-in period prior to the post-randomisation phase of an RCT should always be given consideration. Run-in periods have logistic and resource implications, and address more explanatory than pragmatic considerations; these issues should be taken into account (chapter 3, page 19). (Basis: judgement of the authors.)
22. Use of a run-in period to exclude non-adherers is particularly likely to be helpful when an appreciable proportion of participants are expected to be treatment-intolerant or to fail to comply well enough to achieve appreciable treatment benefit (chapter 3, page 19). (Basis: theoretical considerations and limited case studies.)
23. In explanatory trials, poor compliers should be identified before randomisation so that they can be excluded (chapter 5, page 72). (Basis: anecdotal evidence.)
24. A run-in period may more often be of value to ensure stability of disease in participants rather than to detect non-adherence (chapter 3, page 19). (Basis: one small systematic review.)
25. When run-in periods are employed to ensure stability, compliance should be assessed in order to provide data for possible enhancement of treatment estimates using compliance rate as a baseline covariate (chapter 3, page 19). (Basis: judgement of the authors.)
26. Measurement of compliance during a trial can help to explain a trial's result when a treatment appears not to work but complicated assessments may increase the workload unnecessarily (chapter 5, page 72). (Basis: anecdotal evidence.)
27. Compliance with medication can be improved by patient education, simplifying regimens and using reminder charts (chapter 5, page 72). (Basis: systematic review of small poor quality RCTs.)

Sample size

28. Full details of sample size calculations should always be reported (chapter 3, page 18). (Basis: logical argument.)
29. When small trials are necessary, they should be reported as hypothesis forming (chapter 3, page 18). (Basis: logical argument.)
30. Sample size calculations should consider a sensitivity analysis and should give 'ballpark' estimates rather than unrealistically precise numbers (chapter 3, page 17). (Basis: logical argument.)
31. The possibility of sub-optimal compliance should be taken into account in planning sample sizes of pragmatic trials (chapter 5, page 72). (Basis: empirical evidence.)

Interim analysis

32. Interim analyses should be carefully planned (chapter 5, page 70). (Basis: anecdotal and empirical evidence.)
33. Interim analyses should not be released to study investigators unless absolutely necessary (chapter 5, page 77). (Basis: one small non-random comparative study.)
34. Trials should not be stopped on statistical grounds alone (chapter 5, page 70). (Basis: anecdotal evidence.)
35. An *ad hoc* data monitoring committee should be established in small trials if interim analyses suggest that the trial should be stopped early or the design altered (chapter 5, page 77). (Basis: anecdotal evidence.)

Subgroup analysis

36. Any subgroup analyses which are to be considered as hypothesis testing should

be specified in the trial protocol. These should be limited to subgroups where there is an *a priori* reason to expect a subgroup by treatment interaction (chapter 6, page 86). (Basis: logical argument and widespread expert endorsement.)

37. All other subgroup analyses should be regarded as hypothesis-generating and their interpretation in this more limited role should depend on the number of subgroups examined (chapter 6, page 86). (Basis: logical argument and widespread expert endorsement.)
38. Subgroup analyses should be approached through a comparison of the estimates of treatment effects across the levels of a subgroup variable. The practice of carrying out significance tests within levels of a subgroup variable should be forcefully discouraged. Any significance tests should be based on subgroup by treatment interactions (chapter 6, page 86). (Basis: logical argument and widespread expert endorsement.)

Recruitment

39. Multiple recruitment strategies should be used with the aim of screening at least twice the planned sample size (chapter 5, page 69). (Basis: single survey of RCTs.)
40. Recruitment should be closely monitored and there should be contingency plans for fewer patients being randomised than expected (chapter 5, page 69). (Basis: anecdotal evidence.)
41. Staggering recruitment at centres may help to prevent falling recruitment over time (chapter 5, page 69). (Basis: anecdotal evidence.)
42. Before starting most RCTs, pilot studies are required to refine important aspects of their design and conduct, and to check that recruitment strategies are adequate (chapter 5, pages 69 and 78). (Basis: anecdotal evidence.)
43. There should not be a requirement in large trials to register details of those who are not randomised; however, recruitment logs can be useful and should be encouraged whenever feasible, especially for pilot trials or trials in rare conditions (chapter 5, page 69). (Basis: anecdotal evidence.)

Consent

44. The purpose and requirement for obtaining informed consent from a patient should be stressed in the study protocol. The procedures for giving information and obtaining consent should be designed to ensure that potential participants are given appropriate information in a way that does not interrupt clinical care. This may require trained and dedicated staff to be available to provide information and to obtain consent from patients (chapter 4, page 39). (Basis: logical argument.)
45. The demands of a study on patients should be kept to the minimum consistent with the scientific purpose of the study. In addition, the extent and purpose of the investigations should be clearly explained at the start of the study. Patients should be financially compensated for any travel costs incurred. An appeal to altruism may also be effective if patients are genuinely uncertain whether or not to take part (chapter 4, page 40). (Basis: logical argument, anecdotal evidence.)
46. Trial design should seek to minimise the burden for patients and reassure clinicians about the demands on patients (chapter 4, page 29). (Basis: logical argument.)
47. Potential participants 'must be adequately informed' about the study in accordance with the Declaration of Helsinki. Information should be carefully prepared so that patients of all levels of education have sufficient understanding of the study and what is being asked of them. Information should be given in oral as well as written form, with special provision for those whose first language is not English or who find reading difficult (chapter 4, page 52). (Basis: logical argument.)
48. The consent procedure should be carefully designed so that patients give consent freely, having had the study described to them as fully as they wish. The consent process should be as simple and straightforward as possible, and patients should be given a written copy of the information and consent forms. Staff should receive training in seeking informed consent (chapter 4, page 56). (Basis: empirical evidence and logical argument.)
49. The possible benefits and adverse effects of the treatment options should be described to patients in a balanced way, together with the rationale for random allocation when the best approach is not known. No coercion should be used, however, to persuade patients to participate and the arrangements for care of those who choose not to participate should be part of the study protocol. The trial design should take into account the reasons for some patients' refusal to participate: for example, changes in medication should be minimised, the use of placebo must be justifiable scientifically and ethically, and

the process of randomisation should be presented as an extension of standard medical practice (chapter 4, page 40). (Basis: logical argument.)

50. Patients deserve consideration and sympathy as they decide whether or not to take part in an RCT. They should not be pressurised into taking part. In the longer term, there is a need for ongoing public education about the appropriateness of RCTs when there is clinical uncertainty (chapter 4, page 46). (Basis: logical argument.)

Randomisation

51. There should always be a clear protocol for the preparation of (a) the sequence generation and (b) the concealment up to irrevocable trial entry of the randomisation; the operation of the system should not involve any staff involved in determining entry of patients to the trial (chapter 3, page 10). (Basis: logical argument, anecdotal evidence.)
52. A telephone- or computer-based randomisation scheme can provide secure treatment allocation and allow systematic checking of entry criteria (chapter 3, page 11). (Basis: logical argument and one major systematic review.)

Outcome and follow-up

53. Trialists should collect a small number of relevant and feasible outcome measures (chapter 5, page 75). (Basis: anecdotal evidence.)
54. Data collection should be kept to the minimum consistent with the scientific purpose of the study (chapter 4, page 35). (Basis: anecdotal evidence.)
55. Outcome measures should ideally be objective, valid, reliable, sensitive to change and clinically relevant. In clinical areas where there are 'standard' outcome measures, these should be included so that the results from different RCTs may be combined meaningfully (chapter 3, page 21, and chapter 5, page 75). (Basis: logical argument.)
56. Those assessing outcomes should be properly trained (chapter 5, page 75). (Basis: anecdotal evidence.)
57. Outcome assessments should be blinded if possible (chapter 5, page 74). (Basis: two case studies.)
58. Follow-up should be attempted on all patients who were randomised (chapter 5, page 73). (Basis: empirical evidence.)
59. Most of the issues relating to timescale will be specific to the clinical area in which an RCT is performed; hence, our recommendations are limited to one which is general. Duration of treatment and follow-up should be sufficiently long for effects on clinically important outcomes to be identified (chapter 3, page 22). (Basis: systematic reviews indicating that treatment periods are commonly too short.)
60. Attendance at follow-up can be improved by mailed or telephoned reminders (Basis: systematic review of seven small RCTs.); by providing information about the appointment (Basis: systematic review of three small RCTs.); or by contracting with patients (Basis: systematic review of two small RCTs.) (chapter 5, page 72).

Data quality

61. Double data entry is preferable but single data entry with consistency checks may be appropriate (chapter 5, page 75). (Basis: one RCT and one non-random comparison.)
62. There should be explicit mechanisms to monitor the quality of trial procedures and data (chapter 5, page 76). (Basis: anecdotal evidence.)

Analysis/reporting

63. An ITT analysis will be the method of choice when an unbiased estimate of treatment effects is required, and certainly when the objectives of the trial are pragmatic in terms of treatment effectiveness. There are, however, situations in which the aims of a trial are more explanatory than pragmatic and, in these circumstances, a biased estimate of treatment effectiveness may be acceptable; however, the rules for defining a 'per protocol' analysis need to precede the analysis (chapter 6, page 84). (Basis: case studies and logical argument.)
64. Results should be presented as both absolute and relative treatment effects and should include both benefits and side-effects or hazards (chapter 7, page 97). (Basis: two randomised and two non-random comparisons of the impact of reporting absolute versus relative treatment effects.)
65. Two-tailed tests and corresponding two-tailed CIs should be used unless the trialist is so convinced that differences can only occur in one pre-specified direction, that differences in the opposite direction, however large or unlikely, would irrevocably be interpreted as owing to chance (chapter 6, page 87). (Basis: philosophical argument.)
66. Trialists should submit the results of all RCTs for publication (chapter 7, page 100).

(Basis: two systematic reviews of surveys of final publication status of RCTs in general and those initially published in conference proceedings.)

67. Trialists should make their results publicly available (chapter 7, page 100). (Basis: two systematic reviews of surveys of final publication status of RCTs in general and those initially published in conference proceedings.)
68. Efforts to establish prospective registration of RCTs should continue, to allow unpublished results to be traced (chapter 3, page 18, and chapter 7, page 100). (Basis: two systematic reviews of surveys of final publication status of RCTs in general and those initially published in conference proceedings.)

Administration

69. Multicentre and important single trials need a Steering Group, an independent DMC, and a trial coordinator (chapter 5, page 77). (Basis: anecdotal evidence.)
70. Important commercially sponsored trials should minimise the concerns about potential bias by having independent Steering Groups and DMCs, and independent data management and analysis (chapter 5, page 77). (Basis: anecdotal evidence.)
71. DMCs should be multidisciplinary and have explicit terms of reference (chapter 5, page 77). (Basis: anecdotal evidence.)

Refereeing

72. The status of refereeing should be elevated so that it is of positive value in career development (chapter 6, page 87). (Basis: judgement of the authors.)

Economic issues

(These recommendations are all from chapter 8 (page 109) and are based on at least two methodological or theoretical papers, systematic reviews or competent questionnaire surveys.)

73. Consideration should be given to whether the primary criterion should be a single comprehensive criterion that reflects as wide a range as possible of the benefits and costs of the treatments being compared.
74. Working closely with health economists, treatment cost should be included as an end-point.
75. Treatments should be compared in the form in which they are likely to be considered by future decision makers.
76. In sample size calculations, economically as well as clinically important effect sizes should be considered.
77. All trial costs should be excluded when estimating treatment costs.
78. In reporting treatment costs, the quantities of resources consumed should be separated from the prices paid.
79. Patient outcome measures should be chosen rigorously, to include clinical measures, disease-specific measures of quality of life, and generic measures of patient utility.
80. Findings on resources and outcomes should be reported in the form of net direct costs (i.e. direct costs minus direct financial benefits) and net gain in quality-adjusted life-years.
81. Extrapolation of trial findings to other settings should be based on economic modelling that takes account of differences between trial centres.
82. Serious consideration should be given to reimbursing trial patients for the full cost of taking part.

Recommendations for funders, ethics committees and regulatory bodies

83. They should demand provision of sample size considerations (chapter 3, page 18). (Basis: logical argument.)
84. They should require identification of a primary outcome variable, which may be supplemented by a limited number of secondary outcome variables plus safety variables (chapter 6, page 85). (Basis: logical argument.)
85. They should expect a statistical analysis plan. This will deal in detail with such aspects of the data as repeated observations and how these will be analysed to avoid problems of multiple testing (chapter 6, page 85). (Basis: logical argument.)

Recommendations for journal editors

86. Authors and editors should use the CONSORT guidelines to improve the reporting of RCTs, although certain items may not be necessary; details of data quality checks should also be included (chapter 7, page 99). (Basis: evidence of poor reporting from many surveys of trial reports.)
87. The CONSORT guidelines should be regularly updated in the light of valid criticism and new

research. (chapter 7, page 99). (Basis: anecdotal evidence of problems with CONSORT guidelines.)

88. More journals should adopt a fast-track publication policy for important RCTs (chapter 7, page 100). (Basis: anecdotal evidence.)
89. Reports should make explicit the role of any sponsor in the design, conduct, data management, analysis, and reporting of RCTs, and any financial inducements to participation (chapter 7, page 98). (Basis: one survey of the design and reporting of sponsored trials and further anecdotal evidence.)
90. Multiple publications of the same results should be avoided and journal editors should insist that any previous publications of the same trial are clearly referenced (chapter 7, page 100). (Basis: anecdotal evidence.)
91. Journal editors should be encouraged to implement the earlier recommendations on subgroup analysis in their journal policy (chapter 6, page 86). (Basis: case-study evidence that even journals with a positive attitude to statistical refereeing occasionally allow possibly misleading subgroup analyses to feature strongly in abstracts of papers.)
92. Conclusions should be supported by the data and include a balanced discussion of all relevant evidence, preferably by quoting up-to-date systematic reviews (chapter 7, page 98). (Basis: three surveys of trial reports.)

Recommendations for NHS trusts in the UK

93. Participation in clinical trials should be encouraged as a component of the core activity of clinicians (chapter 4, page 28). (Basis: judgement of the authors.)
94. All trials within the UK should have access to a central randomisation service. Centres offering this service should be established around the country with some offering 24-hour access (chapter 5, page 69). (Basis: anecdotal evidence.)
95. Only those tests relevant to a trial's major objectives should be reimbursed (chapter 8, page 110). (Basis: one non-systematic review.)
96. Unproven treatments should be funded only for patients in peer-reviewed trials (chapter 8, page 110). (Basis: one questionnaire survey.)

Chapter 10

Recommendations for further research

All the research recommendations from earlier sections are brought together in this chapter. Apart from inconsequential reformatting, the recommendations have been taken verbatim from the original chapters, to which each has been cross-referenced. Those areas which are considered to be priorities are italicised.

Conduct and structure of RCTs

- Prospective surveys are needed of the number of planned RCTs that fail to start, the reason(s) for failure and the quality and importance of those trials that fail (chapter 5, page 65).
- *Further assessment is required of different recruitment strategies for community- and hospital-based RCTs, and of the role of specific recruitment coordinators. For example, in multicentre trials, different centres could be randomised to have a recruitment drive or not, or to have a recruitment coordinator or not (chapter 5, page 69).*
- The motives for taking part in clinical trials should be further examined with a view to designing protocols that are more acceptable to patients (chapter 4, page 40).
- Further research is required to find the most appropriate methods to describe trials, particularly randomisation, to patients (chapter 4, page 46).
- *Further research is needed to clarify our understanding of the way in which patients decide whether to enter clinical trials, and to identify the best methods of providing information; in particular, to overcome patient worry about any uncertainties involved in taking part in RCTs (chapter 4, pages 46 and 53).*
- Further research, particularly using an RCT design, is needed to clarify the best ways of achieving informed consent (chapter 4, page 56).
- Research is needed to identify those trial designs which interfere least with the clinician–patient relationship (chapter 4, page 29).
- Research is required to identify and modify those aspects of trial design which are not consistent with normal practice (chapter 4, page 33).
- Further well-designed evaluation studies of interventions designed to improve the conduct and quality of RCTs are required (chapter 5, page 78).
- All existing RCTs of methods to improve compliance should be reviewed by extending the existing Cochrane systematic review (chapter 5, page 72).
- Randomised trials to assess methods of improving health professionals' and patients' compliance with RCT protocols are required (chapter 5, page 72).
- The most feasible methods of assessing compliance in RCTs and the impact of measuring compliance on the conduct (especially recruitment) and results of RCTs need to be evaluated (chapter 5, page 72).
- Both theoretical and empirical research should be directed at methods of using information on compliance during the run-in phase (and potentially during the post-randomisation phase) of trials to improve estimates of treatment effect (chapter 3, page 19).
- Further evaluation of the impact of measuring the success of blinding on the conduct and results of RCTs is required (chapter 5, page 74).
- Further study is required of the best way to assess repeated subjective outcome measures (chapter 5, page 75).
- *Different intensities of quality control should be compared to see how they affect the conduct and quality of the trial. For example, in multicentre trials, centres could be randomised to different levels of quality control (chapter 5, page 76).*
- *Further research is required on:*
  (a) the specific aspects of trial structure, staffing and organisation that improve the quality and progress of large and small trials
  (b) whether small trials need formal structures such as steering groups and DMCs
  (c) the selection of members for steering committees and DMCs (chapter 5, page 77).
- *Further research is required on the risk of bias in commercially sponsored trials compared with unsponsored trials, so that the need for independent control of commercially sponsored trials can be assessed (chapter 5, page 77).*
- Research is needed to identify the most appropriate form of reward for clinicians (chapter 4, page 29).
- Research is required into the effect of offering financial rewards to investigators

for participating in trials and guidelines relating to the level of reimbursement in commercially sponsored trials are needed (chapter 5, page 76).
- Studies are needed of the number of RCTs that terminate early, the reasons for early termination and whether the decision to stop was appropriate (chapter 5, page 70).
- *Prospective surveys of trials are needed to identify the main problems with conduct (especially problems other than recruitment), together with potential solutions (chapter 5, page 78).*
- Further systematic reviews are required to investigate any effect that treatment effect sizes may have on the method of randomisation (chapter 3, page 11).
- Further surveys of trialists are needed to document how frequently randomisation is subverted (chapter 5, page 69).
- *There is need for a systematic review of the use of run-in periods to assess the benefit of this design aspect of RCTs over a range of clinical conditions and various modes of intervention (chapter 3, page 19).*
- Individual clinical areas need to establish a consensus on core variables that should regularly be recorded, in order to allow combined analyses of trials in the area (chapter 3, page 21).
- Systematic reviews are needed in additional areas to document trials using surrogates, unreliable or invalid outcome measures (chapter 3, page 21).
- *Systematic consideration needs to be given to how often results from subgroup analyses have been misleading (chapter 6, page 86).*

Reporting of clinical trials

- The impact of presenting absolute versus relative treatment effects needs to be assessed on more representative samples of doctors and surgeons, and for a wider variety of trials than at present (chapter 7, page 98).
- Randomised trials are needed to assess whether educating clinicians about the use of absolute risks can alter their interpretation of trial results and clinical practice (chapter 7, page 98).
- The adequacy of reporting of the role of sponsors in RCTs needs to be studied and further assessments made of the presence of bias in reports of sponsored trials (chapter 7, page 98).
- *The effect of the CONSORT guidelines on the quality of reporting of RCTs needs to be assessed (chapter 7, page 99).*
- The compliance of journals and authors with the CONSORT guidelines should be monitored (chapter 7, page 99).
- The relationship between the quality of reporting and the quality of the design and conduct of RCTs needs further research (chapter 7, page 91).
- *Further studies are required of the frequency with which the conclusions of RCTs (both commercially sponsored and unsponsored) are supported by the data presented (chapter 7, page 98).*

Training

- The impact and feasibility of having experienced trialists helping inexperienced trialists should be assessed (chapter 5, page 76).
- There is a need for the effects of statistical training on the adequacy of the statistical reporting to be assessed (chapter 6, page 87).

Miscellaneous

- Guidelines should be developed for the evaluation of new technologies before they come into widespread clinical practice (chapter 5, page 65).
- The success of the Culyer reforms in overcoming financial barriers to trials and in eliminating research that is not worth supporting should be monitored (chapter 8, page 110).

Updating this review

Given the breadth of this review, it should probably not be revisited in its entirety. The authors found that the workload was excessive, and even though an update would not repeat all that has been done here it would still be a large undertaking. However, it could be split, for example, into design/conduct, barriers, analysis/reporting. There appears to be an upsurge in interest in the conduct of trials and, if a new review is commissioned in this area, then it should update relevant sections of the report when current research is complete. Realistically, this may not be until 2005 or 2006. The year 2006 has much to recommend it as a possible updating point, since it will be 10 years after the end of the search period considered in the present review. It might be opportune to revisit reporting a little sooner but only when the CONSORT guidelines have had an opportunity to take effect, for example, in 2002 or 2003. In the fast-changing environment in which RCTs are taking place, barriers might profitably be revisited sooner.

Chapter 11

Epilogue

Since its inception, research has been commissioned under the NHS Health Technology Assessment programme in five clinical fields – acute sector, diagnostics and imaging, pharmaceuticals, population screening, and primary and community care – and in the methodological domain. In clinical fields, the aim has been to commission both primary research, typically in the form of randomised trials, and secondary research, in the form of systematic literature reviews. These reviews have covered broad topics like laparoscopic cholecystectomy (acute sector), screening for high blood pressure (population screening), and near patient testing in general practice (primary and community care).[1]

Because these clinical reviews cover such broad topics their authors generally adopt the inclusive approach favoured by the NHS Centre for Reviews and Dissemination,[2] rather than the exclusive approach used by the Cochrane Collaboration.[3] Typically they seek to increase the generalisability of their findings by including studies with a wide range of research designs, rather than ensuring the validity of those findings by restricting attention to randomised trials meeting specified criteria. Thus, the authors constantly face decisions about which studies to include. They seek to ensure the reliability of this process by using more than one reviewer to assess each paper for inclusion and, subsequently, to extract data from papers thus included. As a further check on the validity of this process, authors may subject findings to 'meta-regression' and test for homogeneity of findings between studies.[2]

In creating a methodological domain alongside the five clinical fields, the NHS Standing Group on Health Technology had three basic aims. First, it wanted to acknowledge from the beginning that licence to challenge the *status quo* in clinical technology also carried a responsibility to scrutinise the methodological technologies by which those clinical technologies were assessed. Second, it was keen to meet what it saw as an urgent need for methodological training in an activity that was still in its infancy both in the UK and across the world. Third, it needed to make a long-term commitment to achieving a high methodological standard within the primary programme of technology assessment to be commissioned in the five clinical fields. The Methodology Group soon concluded that all three aims were best served by a comprehensive initial portfolio of systematic literature reviews, and that the need for primary methodological research could not be assessed until reports of these reviews were available.

In tackling this review, the authors soon realised that reviewing methodological literature was at least as far from the inclusive reviews of effectiveness favoured by the NHS Centre for Reviews and Dissemination as those were from the exclusive reviews required by the Cochrane Collaboration. In particular, methodological reviews have at least **three** features that distinguish them from reviews of effectiveness or efficacy. **First**, theory needs to be given **relatively** more weight than empirical evidence: for example, to argue from physiological theory that laparoscopic cholecystectomy is more effective than classical surgery is generally regarded as insufficient to justify a major shift in the type of surgery used to treat gallstones. In contrast, to argue from epidemiological theory that the randomised trial is the research design of choice for assessing most health technologies has been accepted by the Cochrane Collaboration and the editors of major clinical journals.

In saying this we are not, of course, arguing that empirical evidence has no part to play in guiding the choice of design in health technology assessment. Indeed, the authors acknowledge that the findings of a parallel methodological review (as yet unpublished) suggest that the findings of a well-conducted quasi-experiment can be as valid as those of a randomised trial. Nevertheless, there are few fields in which such a comparison can be made, even using non-experimental methods. Furthermore, there are very few examples indeed of experimental comparisons of randomised trials and quasi-experimental methods. This observation is an illustration of the **second** distinctive feature of methodological reviews: empirical evidence is rare and rigorous empirical evidence is even rarer.

In this review, the important question of the best method for recruiting patients to randomised trials

provides an apt example. Whether the criterion of effectiveness was applied (page 66) or that of cost-effectiveness (page 109), no clear recommendations based on comparative studies were possible. This can be attributed both to the general lack of rigour in the published comparisons, none of which was experimental, and to the wide range of topics and contexts in which these comparisons were set. Nevertheless, it was possible in chapter 5 to make six recommendations for recruitment practice, all derived from surveys, case-studies or anecdotal evidence.

That weak evidence should have given rise to stronger conclusions may perhaps be permitted, given the experience of the five principal investigators, all of whom are trialists (in five different fields) and reviewers, either with the Cochrane Collaboration or in the more diffuse form of effectiveness review favoured by the NHS Centre for Reviews and Dissemination. Thus, the greater role for theory in this domain and the shortage of rigorous empirical evidence leads us to suggest a **third** distinctive feature of methodological reviews, at least in the initial learning phase. The authors consider that there is more potential than in traditional reviews for the exercise of logical reasoning and critical judgement in choosing between alternative theories and extrapolating from limited evidence. That said, we are quick to acknowledge the need to validate subjective processes like these and to make methodological reviews more systematic than this prototype review has been.

Our proposal is that, like contemporary reviews of effectiveness, future methodological reviews should be systematic in four basic respects – search strategy, inclusion criteria, data extraction and data synthesis.[2] Chapter 2 demonstrates that our search strategy was indeed explicit and well-defined. However, because it sought to be comprehensive only over the past 10 years, it fell between two stools: it identified many published papers of poor quality while missing many classic methodological papers. With the benefit of hindsight, more of our limited resources should have been devoted to four other methods of searching – a survey of experienced trialists to identify the best methodological papers in this field, scanning the reference lists of these and other key papers, scanning the lists of publications that cite these key papers, and handsearching key journals, notably *Controlled Clinical Trials*. Such a combination of methods of searching would have relied less on the five reviewers, certainly on their memories and probably on their reasoning and judgement.

Hindsight also suggests that we should have tried to make our inclusion criteria more explicit and listed all those papers that were rejected; for example, any theoretical paper that has been endorsed by (say) three peer-reviewed publications might be declared worthy of inclusion, and the list of excluded theoretical papers might specify the very few publications that had endorsed them. Again, although chapters 4, 5, 7 and 8 are all based on structured tables that summarise the main characteristics and conclusions of the papers included in the review, it is hoped that future methodological reviewers will be able to structure such tables more rigorously. Finally, in our judgement, there is also scope for making synthesis more systematic; for example, recommendations might be graded – first according to the rigour of their experimental basis; second, according to the strength of their theoretical basis; and third, according to the number and diversity of the individual trials that had (apparently) adopted them successfully.

It is hoped that these reflections will help readers to appraise our recommendations as they appear in the individual chapters. Just as important, we hope they will be of use to those responsible for reviewing the initial portfolio of methodological literature reviews commissioned under the NHS Health Technology Assessment programme. In particular, we would exhort them to lay the foundations for a protocol for such reviews that will stand comparison with those already in place, both for reviews sponsored by the Cochrane Collaboration[3] and for more diffuse reviews of effectiveness.[2]

That said, our recommendations are commended to those responsible for commissioning, conducting or reviewing health technology assessments based on randomised trials. We consider that the combination of theoretical foundation, empirical judgement and critical judgement that underpins most of these recommendations provides a useful basis for current practice and future development.

References

1. NHS Health Technology Assessment programme. Report 1996. London: NHS Executive; 1996.
2. NHS Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness. CRD Report 4. York: University of York; 1996.
3. The Cochrane Library. The Cochrane Collaboration. Oxford: Update Publications; 1998.

Acknowledgements and authorship

This study and report, which was commissioned under the NHS Health Technology Assessment programme, was conducted under the guidance of a Steering Committee comprising:

Dr RJ Prescott,
Medical Statistics Unit, University of Edinburgh

Dr CE Counsell,
Department of Clinical Neurosciences,
University of Edinburgh

Professor WJ Gillespie,
Department of Orthopaedic Surgery,
University of Edinburgh

Professor AM Grant,
Health Services Research Unit,
University of Aberdeen

Professor IT Russell,
Department of Health Sciences, University of York.

The research associates were S Kiauka and IT Colthart, who were responsible for designing the search strategy, running searches, developing the coding system, reading and making the initial assessment of the relevance of the articles, and coding.

The chapters in this report have been individually authored as listed below. The Steering Committee gratefully acknowledges the contributions of Sue Ross, Sue Shepherd, Daphne Russell and Sandra Kiauka in the writing of this report.

We would also like to acknowledge with thanks the contributions of David Stephen and Neil Scott, who assisted in reviewing the references in chapter 4, and Neil Drummond, who made helpful comments on drafts of this chapter. Our thanks are also due to Susan Kracinska, for secretarial help with chapters 5 and 7, Wilma Warwick, for secretarial help particularly with chapters 3 and 6, and special thanks are also due to Thorsten Förster who, together with Wilma Warwick, has overcome many problems in assembling the separate contributions and standardising their format so that the final report could be produced.

Finally our thanks are also due to the referees for their perseverance in reading the report and for the quality of their comments.

Authorship of individual chapters

Chapter	Authors
1 Introduction	WJ Gillespie, AM Grant
2 Methodology of the review	S Kiauka, CE Counsell, WJ Gillespie, RJ Prescott
3 Design issues	RJ Prescott, SM Shepherd
4 Barriers to participation in clinical trials	S Ross, AM Grant
5 Limiting factors related to the conduct and structure of RCTs	CE Counsell
6 The analysis of RCTs	RJ Prescott
7 Limiting factors relating to the reporting of RCTs	CE Counsell
8 Costs	IT Russell, D Russell
Appendix 1	S Kiauka, IR Colthart
Appendix 2	WJ Gillespie CE Counsell

Appendix 1

Search strategies and coding of papers

Database-specific electronic search strategies

TABLE 29 *Cochrane search strategy for controlled trials in MEDLINE*

Set	Search
1	randomized controlled trial.pt.
2	randomized controlled trials/
3	random allocation/
4	double-blind method/
5	single-blind method/
6	1 or 2 or 3 or 4 or 5
7	(animal not (human and animal)).ti,ab,sh.
8	6 not 7
9	clinical trial.pt.
10	exp clinical trials/
11	(clin$ adj5 trial$).ti,ab.
12	((singl$ or doubl$ or trebl$ or tripl$) adj5 (blind$ or mask$)).ti,ab.
13	placebos/
14	placebo$.ti,ab.
15	random$.ti,ab.
16	research design/
17	9 or 10 or 11 or 12 or 13 or 14 or 15 or 16
18	(animal not (human and animal)).ti,ab,sh.
19	17 not 18
20	comparative study/
21	exp evaluation studies/
22	follow-up studies/
23	prospective studies/
24	(control$ or prospectiv$ or volunteer$).ti,ab.
25	20 or 23 or 22 or 23 or 24
26	(animal not (human and animal)).ti,ab,sh.
27	25 not 26
28	8 or 19 or 27

TABLE 30 *Ovid MEDLINE search for 1992–April 1996 (dated 24 April 1996)*

Set	Search	Results
001	randomized controlled trials/	3536
002	random allocation/	4611
003	double-blind method/	12,510
004	single-blind method/	1675
005	exp clinical trials/	10,796
006	placebos/	1864
007	1 or 2 or 3 or 4 or 5 or 6	32,503
008	(clin$ adj5 trial$).ti,ab.	11,606
009	((singl$ or doubl$ or trebl$ or tripl$) adj5 (blind$ or mask$)).ti,ab.	11,329
010	placebo$.ti,ab.	12,714
011	random$.ti,ab.	43,365
012	randomized controlled trial.pt.	31,062
013	controlled clinical trial.pt.	5116
014	clinical trial.pt.	55,479
015	8 or 9 or 10 or 11 or 12 or 13 or 14	85,428
016	7 or 15	97,508
017	(animal not (human and animal)).ti,ab,sh.	303,425
018	16 not 17	89,204
019	*clinical protocols/	435
020	*planning techniques/	71
021	*research design/	833
022	methods/	3274
023	*patient selection/	237
024	*sample size/	15
025	*eligibility determination/	67
026	*selection bias/	62
027	*patient participation/	557
028	*patient education/	2441
029	*patient satisfaction/	1062
030	*informed consent/	1075
031	*refusal to treat/	200
032	*treatment refusal/	493
033	*patient compliance//	1064
034	*patient dropouts/	174
035	*physician's role/	1628
036	exp *professional-patient relations/	3525
037	*patient acceptance of health care/	809
038	*meta-analysis/	283
039	*data collection/	497
040	*research support/	931
041	*health care reform/	3127
042	*insurance, liability/	155
043	*quality assurance health care	2513
044	*quality control/	162
045	*health services research/	758
046	*bias (epidemiology)/	222
047	*publication bias/	33
048	*ethics committees/	231
049	19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48	22,384
050	49 and 18	2407
051	(gcp? or good clinical practice$).ti,ab.	107
052	((difficult$ or problem$ or obstacle$ or issue$) adj3 (trial$ or stud$ or research$ or random$)).ti.	365
053	((difficult$ or problem$ or obstacle$ or issue$) adj3 (protocol$)).ti,ab.	118
054	((difficult$ or problem$ or obstacle$ or issue$) adj3 (design$)).ti,ab.	469
055	((difficult$ or problem$ or obstacle$ or issue$or lack$) adj5 (funding or funds or fund)).ti,ab.	136
056	((difficult$ or problem$ or obstacle$ or issue$) adj5 (accru$ or recruit$ or enrol$ or participa$)).ti,ab.	433
057	((difficult$ or problem$ or obstacle$ or issue$) adj5 (informed consent$)).ti,ab.	70
058	((difficult$ or problem$ or obstacle$ or issue$ or flaw$) adj5 (methodolog$)).ti,ab.	1369
059	((difficult$ or problem$ or obstacle$ or issue$) adj5 (bias$)).ti,ab.	189
060	((difficult$ or problem$ or obstacle$ or issue$) adj2 (conduct$)).ti,ab.	268
061	((difficult$ or problem$ or obstacle$ or issue$ or refus$) adj3 (random$)).ti,ab.	200
062	((treat$ adj3 prefer adj5 (random$ or alloc$ or assign$ or arm or active)).ti,ab.	7
063	((difficult$ or problem$ or obstacle$ or issue$) adj5 (blind$)).ti,ab.	116
064	((difficult$ or problem$ or obstacle$ or issue$ or inappropriat$ or inadequat$) adj3 (tim?scale$ or timing or time?frame)).ti,ab.)	71
065	(quality adj3 (trial$ or stud$ or result$ or method$ or science or research$) adj3 (poor or bad or low or problem$)).ti,ab.	95
066	intent$ to treat$ analys$.ti,ab.	151
067	((research$ or stud$ or trial$) adj5 (nhs$ or national health service$)).ti,ab.	108
068	((reform$ or new$ or change$) adj5 (nhs$ or national health service$)).ti,ab.	242
069	51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68	3818
070	(protocol$ adj3 (trial$ or stud$ or research$ or random$)).ti.	117
071	(design$ adj3 (trial$ or stud$ or research$ or random$)).ti.	448
072	((funding or funds or fund) adj3 (trial$ or stud$ or research$ or random$)).ti.	121
073	((accru$ or recruit$ or enrol$ or participa$) adj3 (trial$ or stud$ or research$ or random$)).ti.	229
074	(consent$ adj3 (trial$ or stud$ or research$ or random$)).ti.	81
075	(methodolog$ adj3 (trial$ or stud$ or research$ or random$)).ti.	275

continued

TABLE 30 contd *Ovid MEDLINE search for 1992–April 1996 (dated 24 April 1996)*

Set	Search	Results
continued		
076	(conduct$ adj3 (trial$ or stud$ or research$ or random$)).ti.	198
077	(randomi#ation$ adj3 (trial$ or stud$ or research$ or random$)).ti.	22
078	(committee$ adj3 (trial$ or stud$ or research$ or random$)).ti.	337
079	(quality adj3 (issue$ or trial$ or stud$ or research$ or random$)).ti.	368
080	((complian$ or comply) adj3 (trial$ or stud$ or research$ or random$)).ti.	42
081	(bias$ adj3 (report$ or publication$ or select$ or trial$ or stud$ or research$ or random$)).ti.	62
082	(sample size$ adj3 (trial$ or stud$ or research$ or random$)).ti.	38
083	(statistic$ adj3 (trial$ or stud$ or research$ or random$)).ti.	164
084	((indemni$ or insur$ or liab$) adj3 (trial$ or stud$ or research$ or random$)).ti.	17
085	(blinding$ adj3 (trial$ or stud$ or research$ or random$)).ti.	365
086	((terminat$ or stop$ or discontin$ or halt$) adj3 (trial$ or stud$ or research$ or random$)).ti.	49
087	((drop?out$ or withdraw$) adj3 (patient$ or trial$ or stud$ or research$ or random$)).ti.	177
088	70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87	2867
089	(69 or 88) and 18	6341
090	50 or 89	**7747**

TABLE 31 *BIDS EMBASE search for 1986–April 1996*

Set	Hits	
1	805,794	(**clinical trial**)@**EX**,(**controlled study**)@**EX**,(**medical research**)@**EX**,(**clinical study**)@**KW**, (**randomization**)@**KW**,(**questionnaire**)@**KW**,(**survey***)@**TI**
2	45	((**method*+compar***)+(**trial***,**research***,**random***))@TI +((clinical trial)@EX,(controlled study)@EX,(medical research)@EX,(clinical study)@KW, (randomization)@KW,(questionnaire)@KW, (survey*)@TI
3	97	(**informed consent**)@KW+(((trial*,research*))@TI)+((consent*,compar*)@TI) +((clinical trial)@EX,(controlled study)@EX,(medical research)@EX,(clinicalstudy)@KW, (randomization)@KW,(questionnaire)@KW,(survey*)@TI)
4	4	(**randomized consent***)@(TI,AB,KWDS)
5	241	((**recruit***,**enrol***,**accru***,**particip***)@TI)+(((trial*,research*))@TI)
6	233	(**patient selection**)@KW+(((trial*,research*))@TI) +((clinical trial)@EX, (controlled study)@EX,(medical research)@EX, (clinical study)@KW, (randomization)@KW,(questionnaire)@KW,(survey*)@TI)
7	72	(**patient education**)@KW+(((trial*,research*))@TI) +((clinical trial)@EX, (controlled study)@EX,(medical research)@EX, (clinical study)@KW, (randomization)@KW,(questionnaire)@KW,(survey*)@TI)
8	58	(**patient attitude**)@KW+(((trial*,research*))@TI) +((clinical trial)@EX,(controlled study)@EX,(medical research)@EX,(clinical study)@KW, (randomization)@KW, (questionnaire)@KW,(survey*)@TI)
9	64	(**physician attitude**)@KW+(((study,studies,trial*,research*))@TI) +((clinical trial)@EX ,(controlled study)@EX,(medical research)@EX ,(clinical study)@KW, (randomization)@KW,(questionnaire)@KW,(survey*)@TI)
10	66	(**doctor patient relation**)@KW+(((trial*, research*))@TI) +((clinical trial)@EX,(controlled study)@EX,(medical research)@EX,(clinical study)@KW, (randomization)@KW,(questionnaire)@KW,(survey*)@TI)
11	55	(**patient compliance**)@KW+(((study,studies,trial*,research*))@TI)+(compar*)@TI +((clinical trial)@EX ,(controlled study)@EX,(medical research)@EX ,(clinical study)@KW, (randomization)@KW,(questionnaire)@KW,(survey*)@TI)
12	15	(**dropout***)@TI+(((study,studies,trial*,research*))@TI)
13	10	(**blinding***)@TI+(((study,studies,trial*,research*))@TI)
	863	**TOTAL**

TABLE 32 *CINAHL search for 1986–April 1996*

Set	Search	Results
1	exp clinical trials/	1571
2	exp random sample/	1443
3	comparative studies/	3486
4	placebos/	98
5	(clin$ adj3 trial$).tw.	607
6	surveys/	7705
7	questionnaires/	5465
8	1 or 2 or 3 or 4 or 5 or 6 or 7	17,498
9	consent/	1084
10	"consent(research)"/	128
11	informed consent.ti.	269
12	9 or 10 or 11	1214
13	12 and 8	119
14	patient selection/	492
15	14 and 8	43
16	(trial$ or research$ or random$).ti.	7410
17	(problem$ or obstacle$ or difficult$ or barrier$).ti.	7786
18	selection bias/	9
19	"bias (research)"/	76
20	18 or 19	85
21	20 and 16	18
22	sample size	118
23	22 and 16	30
24	(recruit$ or enrol$ or accru$ or participa$).ti.	1279
25	24 and 16	70
26	consumer participation/	589
27	26 and 8	54
28	patient satisfaction/	1025
29	28 and 16	28
30	patient attitudes/	630
31	30 and 16	23
32	patient compliance/	1391
33	(comply or complian$).ti.	541
34	32 or 33	1551
35	34 and 16	60
36	research dropouts/	24
37	36 and (16 or 8)	18
38	drop?out$.tw.	44
39	38 and 16	5
40	exp professional-patient relations/	4841
41	40 and 16	76
42	exp research methodology/	33,165
43	42 and 13	106
44	methodolog$.ti.	358
45	44 and 16 and 17	30
46	(compar$ and method$ and (trial$ or research$)).ti.	5
47	"conflict of interest"/	53
48	47 and 16	6
49	ethics committees/	172
50	ethic$ committee$.tw.	114
51	49 or 50	201
52	51 and (8 or 16)	33
53	(gcp$ or good clinical practice or guideline$).tw.	3021
54	53 and 16	130
55	indemni$.tw.	15
56	(conduct$ and trial$).ti.	8
57	(trial$ or medical research or random$ or phase$).ti.	1092
58	57 and 17	29
59	or 13,15,21,23,25,27,29,31,35,37,39,41, 43,45-46,48,52,54-56,58	735
60	limit 59 to (yr=1986 or yr=1987 or yr=1988 or yr=1989 or yr=1990 or yr=1991 or yr=1992 or yr=1993 or yr=1994 or yr=1995 or yr=1996)	705

Recording form

Title .....

Author/s .....

Journal .....

Reviewed by .....

Key article ☐

Relevant issues/factors	Guidelines/ recommendations	Problems	Methodological	New design of analysis in specialised area	Other
☐	☐	☐	☐	☐	☐

Design	Analysis	Costing	Other obstacles	Clinician participation	Patient participation	Reporting	Conduct
☐	☐	☐	☐	☐ ☐	☐	☐	☐

Subject area

Orthopaedics	☐	Cancer	☐
Stroke	☐	General	☐
Perinatal	☐	Other	☐
		(specify)	.....

Setting

General practice ☐ Hospital ☐ Community ☐ Unclear ☐ N/A ☐

Methodology

Comparison of trial design	☐
Other evaluation of trial design	☐
Meta-analysis	☐
Systematic review	☐
Review	☐
Survey	☐
Case study	☐
Primary research	☐
Education paper	☐

Article type

Article	☐
Abstract	☐
Editorial	☐
Comment	☐
Commentary	☐
Letter	☐
News	☐

Type of intervention

Drug intervention	☐	Therapy	☐	Surgery	☐
Complex interventions	☐	Other	☐	N/A	☐

Country .....

Report	☐	**Bibliography**	☐		
Generalisable	☐	**Not generalisable**	☐	**Unsure**	☐

Keywords:

-
-
-
-
-
-
-
-
-
-

-
-
-
-
-
-
-
-
-
-

Other keywords/notes:

Keywords for coding

TABLE 33 *Keyword definitions*

Index word	Description
Accessibility	to trial centre, hospital, etc
Administration	administration of trial
AIDS	trials in the fields HIV/AIDS
Ambition	insufficient ambition on behalf of clinicians
Attitude	attitude of investigator to trial
Awareness	lack of patient awareness of balance of risk (of being in a trial)
Biased losses	biased losses to follow-up
Bibliography	article may only be useful as bibliographic reference
Blinding	
Characteristics	patient characteristics, e.g. demographics
Compliance	patient compliance in trial
Confidence interval failure	failure to report confidence intervals
Context failure	failure to put research in the context of other research
Data management	e.g. data collection, entry
Description failure	failure to describe procedures e.g. randomisation, blinding
Double standards	in consent procedures inside and outside a trial
Drop-outs	drop-outs from trials
Duplication	unnecessary duplication, inadequate prior review of previous research
Economics	relates to economic analysis alongside trials
Eligibility criteria	too restrictive criteria leading to poor generalisability
Emergency medicine	trials in emergency situations, e.g. resuscitation research
Empirical	contains empirical evidence (primarily used for letters)
Ethics	ethics of trials
Ethics committees	the role of ethics committees in trials
Experience	experience of clinicians and nurses in conducting trials
Failure to complete	failure to complete trial
Failure to stop	failure to stop a trial that is no longer useful or necessary
Follow-up	follow up of patients
Form design	design of informed consent form/ease of comprehension
Fraud	fraudulent activities, e.g. falsifying patient data
Funding	issues related to the funding of trials
Generalisable	is article generalisable to other countries/settings?
Health services research	non-therapeutic trials, e.g. counselling services
Impact	impact of trials on practice, clinician's behaviour, etc.
Inappropriate comparison	inappropriate comparison groups
Inappropriate conclusions	inappropriate conclusions for data presented
Inappropriate end-points	
Inappropriate outcomes	inappropriate outcome measures
Inappropriate tests	inappropriate statistical tests
Indemnity	
Info presentation	the way information about trials is presented to patients
Informed consent	
Interference	outside interference in trials, e.g. from drug companies
Interim analyses	poorly conducted or inappropriate interim analyses
Investigation costs	investigative costs of trials
ITT analysis	intention-to-treat analysis
Language	issues related to other than English language trials
Legal	medico-legal problems
Link	indicates a useful piece of information in article
Location	of trial centre, hospital, etc.

continued

TABLE 33 contd *Keyword definitions*

Index word	Description
Mega-trial	
Meta-analysis	issues relating to meta-analyses
Minorities	trial participation of women, non-white patients, minors, etc.
Monitoring	specifically the monitoring of the trial by DMC
Motivation	motivation of patients in participating, e.g. altruism
Multiple analysis same endpoint	
Multiple comparisons	
Multiple end-points	
Multiple publication	multiple publication of trial
NHS change	impact of NHS reforms
Not generalisable	article is not generalisable, typically outside USA
Nursing	nursing involvement in trials
Organisation	of trials
Outcome assessment	poor assessment of outcome
Over-elaborate	of trial design
Patient relations	impact of trial on doctor/nurse–patient relationship
Phase I	Phase I trial of drug
Phase II	Phase II trial of drug
Phase III	Phase III trial of drug
Phase IV	Phase IV trial of drug
Placebo	
Preferences	preferences of doctor or patient for a particular treatment
Preparation	inadequate piloting, lack of preparation for trial
Pressure	pressure from, for example, drug companies, consumer groups, activists to conduct trials
Prevention trial	article describes prevention trial
Protocol	issues/problems related to trial protocol
Publication bias	
QoL	quality of life – all areas
Quality	quality of trial
Quality assessment tool	tool or method for assessing the quality of trials
Quality assurance	procedures to ensure/maintain quality during conduct of trial
Randomisation	
Recruitment	issues related to patient recruitment
Recruitment cost	cost of recruiting subjects
References	article contains useful references for follow-up
Regulation	regulation of trials
Report	article should be included in final report (only suggestion)
Resources	resources required for trial/impact of trial on existing resources
Result generalisability	whether trial results are generalisable to wider population
Rewards	rewards (financial, recognition) related to trial participation (doctor and patient)
Rigid design	too rigid trial design
Run-in	run-in period of trial – general
Sample size	
Selection bias	
Service costs	
Significance level	inappropriate choice of significance level
Stopping rules	rules employed to guide decision to halt trial
Subgroup analyses	inappropriate (or over-emphasis on) subgroup analysis
Technology	role of technology in trials
Time	pressures of time for clinicians
Timescale	inadequate timescale for trial
Treatment	treatment costs of trial
Trial type	specifically addresses trial type/design
Uncertainty	patient unwillingness or inability to cope with uncertainty
Wash-out	use of a wash-out period

Appendix 2

Summary data from two Cochrane Collaboration trial registers

Introduction

Data were examined from the registers of trials held by the Cochrane Collaboration Stroke Group,[1] and from two systematic reviews,[2,3] carried out by members of the Cochrane Collaboration Musculo-skeletal Group, to establish the reasons for exclusion of trials from systematic reviews, and the extent to which researchers have focused on clinically relevant outcomes.

Reasons for exclusion of trials from systematic reviews

The 128 trials initially considered eligible in 21 completed reviews of the management of acute stroke completed by the Cochrane Stroke Group and the 104 trials which met the inclusion criteria in a review of hormone replacement therapy in the prevention and treatment of osteoporosis were examined. The reasons for exclusion of the trials which met the initial inclusion criteria are listed in *Table 34*. Only 58% of identified trials were included in stroke reviews and only 32% in the hormone replacement therapy review.

Surrogate outcomes

"For phase three trials, the primary end-point should be a clinical event relevant to the patient, that is, the event of which the patient is aware and wants to avoid".[4] However, the use of intermediate or surrogate end-points has been widespread, because they are likely to reduce the duration and cost of trials and provide quicker academic reward for trialists.

There is mounting evidence, recently reviewed by Fleming and DeMets,[4] that in a number of fields, including cardiology, cancer research, AIDS and osteoporosis, trials using biologically plausible intermediate measurements as end-points have given results interpreted as beneficial but which have not been confirmed in long-term studies of clinically important outcomes. New technologies which measure such surrogate end-points may be enthusiastically taken up by the research community – the measurement of bone mineral density is an example. For some time thereafter, great emphasis may be placed on its value as a surrogate that is assumed to predict the outcome of clinical interest.[5]

TABLE 34 *Reasons for exclusion of trials from systematic reviews*

Reason for exclusion	Management of acute stoke	Hormone replacement therapy for osteoporosis/ fracture prevention
Satisfied initial inclusion criteria	128	104
Classified as non-randomised during quality assessment	12	23
Additional information sought: no response	1	8
No extractable data		6
Results reported elsewhere		18
Only active treatment comparisons used		15
Absence of appropriate participants, interventions or outcomes	13	1
Large loss to follow-up (ascertainment bias)	3	
Confounding by other treatment factor	24	
Too small to be useful	1	
Studies included	74	33

TABLE 35 *Trials included in reviews: clinically relevant and intermediate outcomes*

Theme	RCTs identified	Clinically important outcomes alone (participants)	Intermediate and clinically important outcomes (participants)	Intermediate outcomes alone (participants)
Hormone replacement therapy in osteoporosis/ fracture prevention	33 (to end 1993)	None	A 0 (0) B 1 (70)	32 (1390)
Vitamin D analogues in fracture prevention in the elderly	25 (to end 1995)*	A 2 (176) B 3 (239)	A 4 (6020) (Intermediate 576) B 8 (360)	8 (2144)

* *14 trials met the inclusion criteria for the review*

Two reviews completed by members of the Cochrane Musculo-skeletal Group and 21 by the Cochrane Stroke Group were examined to assess whether the trials had measured clinically important outcomes. In each review, the reviewers had excluded (see *Table 34*):

- studies which were classified as non-randomised during quality assessment
- studies in which although necessary additional information had been sought from the trialist, it was not provided
- studies without extractable data for any outcome
- studies reporting data duplicated in included studies
- studies with only active treatment comparisons.

The two reviews by the Cochrane Musculo-skeletal Group explored the possibilities for prevention of fractures in the elderly by the administration of hormone replacement therapy to postmenopausal women, and by the administration of vitamin D or a Vitamin D analogue. The outcome of clinical interest was a fracture event of which the individual patient was aware (Category A) or a radiologically identified vertebral fracture (Category B); the intermediate outcome was bone mass (measured as bone mineral density or bone mineral concentration by a variety of methods). The distribution of trials between those recording clinically important outcomes and those recording intermediate outcomes only is shown in *Table 35*. Only one hormone replacement therapy trial reported outcomes of clinical importance. For vitamin D and its analogues, clinical outcomes were more frequently recorded. Most studies of hormone replacement therapy were carried out in women in the early postmenopausal period, in whom clinical fracture events would be infrequent in the short and medium term. In contrast, the majority of participants in the vitamin D studies were elderly and at higher risk of sustaining a fracture during a feasible study period.

In the stroke reviews, the outcomes of clinical interest were death and a measure of disability in survivors. In the 74 trials included in acute stroke reviews, all reported death rates but only 43 (58%) reported any measure of disability.

The data from these systematic reviews confirm the widespread use of surrogate outcome measures. Anxieties have been expressed about the predictive value of such intermediate outcomes.[4]

References

1. Warlow C, Sandercock P. Carotid endarterectomy vs. carotid angioplasty. *Lancet* 1997;**349**:880.
2. O'Connell D, Robertson J, Henry D, Gillespie W. A systematic review of the skeletal effects of estrogen therapy in post-menopausal women. II: An assessment of treatment effects. *Climacteric* 1998;**1**:112–23.
3. Gillespie WJ, Henry DA, O'Connell DL, Lau EM, Robertson J. Vitamin D analogues in the prevention of fractures. In: Cochrane Library. Issue 3, 1996. (Musculoskeletal Injuries module). Oxford: Update Software.
4. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;**125**:605–13.
5. O'Connell D, Robertson J, Henry D, Gillespie W. Estrogens and osteoporosis: investigators concerned with technology rather than clinical science [abstract]. 3rd annual Cochrane Colloquium, 4–8 October 1995;V-II.

Health Technology Assessment panel membership

This report was identified as a priority by the Methodology Group.

Acute Sector Panel

Current members

Chair:
Professor Francis H Creed
University of Manchester

Professor Clifford Bailey
University of Leeds

Ms Tracy Bury
Chartered Society of Physiotherapy

Professor Collette Clifford
University of Birmingham

Dr Katherine Darton
M.I.N.D.

Mr John Dunning
Papworth Hospital, Cambridge

Mr Jonathan Earnshaw
Gloucester Royal Hospital

Mr Leonard Fenwick
Freeman Group of Hospitals, Newcastle-upon-Tyne

Professor David Field
Leicester Royal Infirmary

Ms Grace Gibbs
West Middlesex University Hospital NHS Trust

Dr Neville Goodman
Southmead Hospital Services Trust, Bristol

Professor Mark Haggard
MRC Institute of Hearing Research, University of Nottingham

Professor Robert Hawkins
University of Manchester

Dr Duncan Keeley
General Practitioner, Thame

Dr Rajan Madhok
East Riding Health Authority

Dr John Pounsford
Frenchay Hospital, Bristol

Dr Mark Sculpher
University of York

Dr Iqbal Sram
NHS Executive, North West Region

Mrs Joan Webster
Consumer member

Past members

Professor John Farndon*
University of Bristol

Professor Senga Bond
University of Newcastle-upon-Tyne

Professor Ian Cameron
Southeast Thames Regional Health Authority

Ms Lynne Clemence
Mid-Kent Health Care Trust

Professor Cam Donaldson
University of Aberdeen

Professor Richard Ellis
St James's University Hospital, Leeds

Mr Ian Hammond
Bedford & Shires Health & Care NHS Trust

Professor Adrian Harris
Churchill Hospital, Oxford

Dr Gwyneth Lewis
Department of Health

Mrs Wilma MacPherson
St Thomas's & Guy's Hospitals, London

Dr Chris McCall
General Practitioner, Dorset

Professor Alan McGregor
St Thomas's Hospital, London

Professor Jon Nicholl
University of Sheffield

Professor John Norman
University of Southampton

Professor Michael Sheppard
Queen Elizabeth Hospital, Birmingham

Professor Gordon Stirrat
St Michael's Hospital, Bristol

Dr William Tarnow-Mordi
University of Dundee

Professor Kenneth Taylor
Hammersmith Hospital, London

* Previous Chair

continued

continued

Diagnostics and Imaging Panel

Current members

Chair:
Professor Mike Smith
University of Leeds

Dr Philip J Ayres
Leeds Teaching Hospitals NHS Trust

Dr Paul Collinson
St George's Hospital, London

Dr Barry Cookson
Public Health Laboratory Service, Colindale

Professor David C Cumberland
University of Sheffield

Professor Adrian Dixon
University of Cambridge

Mr Steve Ebdon-Jackson
Department of Health

Mrs Maggie Fitchett
Association of Cytogeneticists, Oxford

Dr Peter Howlett
Portsmouth Hospitals NHS Trust

Professor Alistair McGuire
City University, London

Dr Andrew Moore
Editor, *Bandolier*

Dr Peter Moore
Science Writer, Ashtead

Professor Chris Price
London Hospital Medical School

Dr William Rosenberg
University of Southampton

Mr Tony Tester
South Bedfordshire Community Health Council

Dr Gillian Vivian
Royal Cornwall Hospitals Trust

Dr Greg Warner
General Practitioner, Hampshire

Past members

Professor Michael Maisey*
Guy's & St Thomas's Hospitals, London

Professor Andrew Adam
Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Dr Pat Cooke
RDRD, Trent Regional Health Authority

Ms Julia Davison
St Bartholomew's Hospital, London

Professor MA Ferguson-Smith
University of Cambridge

Dr Mansel Haeney
University of Manchester

Professor Sean Hilton
St George's Hospital Medical School, London

Mr John Hutton
MEDTAP International Inc., London

Professor Donald Jeffries
St Bartholomew's Hospital, London

Dr Ian Reynolds
Nottingham Health Authority

Professor Colin Roberts
University of Wales College of Medicine

Miss Annette Sergeant
Chase Farm Hospital, Enfield

Professor John Stuart
University of Birmingham

Dr Ala Szczepura
University of Warwick

Mr Stephen Thornton
Cambridge & Huntingdon Health Commission

Dr Jo Walsworth-Bell
South Staffordshire Health Authority

Methodology Group

Current members

Chair:
Professor Martin Buxton
Health Economics Research Group, Brunel University

Professor Doug Altman
ICRF/NHS Centre for Statistics in Medicine, University of Oxford

Dr David Armstrong
Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Professor Nicholas Black
London School of Hygiene & Tropical Medicine

Professor Ann Bowling
University College London Medical School

Dr Mike Clarke
UK Cochrane Centre, Oxford

Professor Paul Dieppe
MRC Health Services Research Collaboration, University of Bristol

Professor Mike Drummond
Centre for Health Economics, University of York

Dr Vikki Entwistle
University of Aberdeen

Professor Ewan Ferlie
Imperial College, London

Professor Ray Fitzpatrick
University of Oxford

Mrs Jenny Griffin
Department of Health

Professor Jeremy Grimshaw
University of Aberdeen

Dr Stephen Harrison
University of Leeds

Mr John Henderson
Department of Health

Professor Richard Lilford
R&D, West Midlands

Professor Theresa Marteau
Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Dr Henry McQuay
University of Oxford

Dr Nick Payne
University of Sheffield

Professor Maggie Pearson
NHS Executive North West

Dr David Spiegelhalter
Institute of Public Health, Cambridge

Professor Joy Townsend
University of Hertfordshire

Ms Caroline Woodroffe
Standing Group on Consumers in NHS Research

Past members

Professor Anthony Culyer*
University of York

Professor Michael Baum
Royal Marsden Hospital

Dr Rory Collins
University of Oxford

Professor George Davey Smith
University of Bristol

Professor Stephen Frankel
University of Bristol

Mr Philip Hewitson
Leeds FHSA

Mr Nick Mays
King's Fund, London

Professor Ian Russell
University of York

Professor David Sackett
Centre for Evidence Based Medicine, Oxford

Dr Peter Sandercock
University of Edinburgh

Dr Maurice Slevin
St Bartholomew's Hospital, London

Professor Charles Warlow
Western General Hospital, Edinburgh

* Previous Chair

Pharmaceutical Panel

Current members

Chair:
Professor Tom Walley
University of Liverpool

Dr Felicity Gabbay
Transcrip Ltd

Dr Peter Golightly
Drug Information Services,
NHS Executive Trent

Dr Alastair Gray
Health Economics
Research Centre,
University of Oxford

Professor Rod Griffiths
NHS Executive
West Midlands

Mrs Jeanette Howe
Department of Health

Professor Trevor Jones
ABPI, London

Ms Sally Knight
Lister Hospital, Stevenage

Dr Andrew Mortimore
Southampton & SW Hants
Health Authority

Mr Nigel Offen
NHS Executive Eastern

Dr John Reynolds
The Oxford Radcliffe Hospital

Mrs Marianne Rigge
The College of Health,
London

Mr Simon Robbins
Camden & Islington
Health Authority, London

Dr Frances Rotblat
Medicines Control Agency

Dr Eamonn Sheridan
St James's University Hospital,
Leeds

Mrs Katrina Simister
National Prescribing Centre,
Liverpool

Dr Ross Taylor
University of Aberdeen

Past members

Professor Michael Rawlins*
University of Newcastle-
upon-Tyne

Dr Colin Bradley
University of Birmingham

Professor Alasdair
Breckenridge
RDRD, Northwest Regional
Health Authority

Ms Christine Clark
Hope Hospital, Salford

Mrs Julie Dent
Ealing, Hammersmith &
Hounslow Health Authority,
London

Mr Barrie Dowdeswell
Royal Victoria Infirmary,
Newcastle-upon-Tyne

Dr Tim Elliott
Department of Health

Dr Desmond Fitzgerald
Mere, Bucklow Hill, Cheshire

Professor Keith Gull
University of Manchester

Dr Keith Jones
Medicines Control Agency

Dr John Posnett
University of York

Dr Tim van Zwanenberg
Northern Regional
Health Authority

Dr Kent Woods
RDRD, Trent RO,
Sheffield

Population Screening Panel

Current members

Chair:
Professor Sir John
Grimley Evans
Radcliffe Infirmary, Oxford

Mrs Stella Burnside
Altnagelvin Hospitals Trust,
Londonderry

Mr John Cairns
University of Aberdeen

Professor Howard Cuckle
University of Leeds

Dr Carol Dezateux
Institute of Child Health,
London

Mrs Anne Dixon-Brown
NHS Executive Eastern

Professor Dian Donnai
St Mary's Hospital,
Manchester

Dr Tom Fahey
University of Bristol

Mrs Gillian Fletcher
National Childbirth Trust

Dr JA Muir Gray
National Screening Committee,
NHS Executive Oxford

Professor Alexander Markham
St James's University Hospital,
Leeds

Dr Ann McPherson
General Practitioner,
Oxford

Dr Susan Moss
Institute of Cancer Research

Mr John Nettleton
Consumer member

Mrs Julietta Patnick
NHS Cervical
Screening Programme,
Sheffield

Dr Sarah Stewart-Brown
Health Service Research Unit,
University of Oxford

Past members

Dr Sheila Adam*
Department of Health

Professor George Freeman
Charing Cross & Westminster
Medical School, London

Dr Mike Gill
Brent & Harrow
Health Authority

Dr Anne Ludbrook
University of Aberdeen

Professor Theresa Marteau
Guy's, King's &
St Thomas's School of
Medicine & Dentistry,
London

Professor Catherine Peckham
Institute of Child Health,
London

Dr Connie Smith
Parkside NHS Trust,
London

Ms Polly Toynbee
Journalist

Professor Nick Wald
University of London

Professor Ciaran Woodman
Centre for
Cancer Epidemiology,
Manchester

* Previous Chair

continued

continued

Primary and Community Care Panel

Current members

Chair:
Dr John Tripp
Royal Devon & Exeter Healthcare NHS Trust

Mr Kevin Barton
East London & City Health Authority

Professor John Bond
University of Newcastle-upon-Tyne

Dr John Brazier
University of Sheffield

Ms Judith Brodie
Cancer BACUP

Mr Shaun Brogan
Ridgeway Primary Care Group, Aylesbury

Mr Joe Corkill
National Association for Patient Participation

Dr Nicky Cullum
University of York

Professor Pam Enderby
University of Sheffield

Dr Andrew Farmer
Institute of Health Sciences, Oxford

Dr Jim Ford
Department of Health

Professor Richard Hobbs
University of Birmingham

Professor Allen Hutchinson
University of Sheffield

Dr Aidan MacFarlane
Independent Consultant

Professor David Mant
Institute of Health Sciences, Oxford

Dr Chris McCall
General Practitioner, Dorset

Dr Robert Peveler
University of Southampton

Professor Jennie Popay
University of Salford

Dr Ken Stein
North & East Devon Health Authority

Past members

Professor Angela Coulter*
King's Fund, London

Professor Martin Roland*
University of Manchester

Dr Simon Allison
University of Nottingham

Professor Shah Ebrahim
Royal Free Hospital, London

Ms Cathy Gritzner
King's Fund, London

Professor Andrew Haines
RDRD, North Thames Regional Health Authority

Dr Nicholas Hicks
Oxfordshire Health Authority

Mr Edward Jones
Rochdale FHSA

Professor Roger Jones
Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Mr Lionel Joyce
Chief Executive, Newcastle City Health NHS Trust

Professor Martin Knapp
London School of Economics & Political Science

Dr Phillip Leech
Department of Health

Professor Karen Luker
University of Liverpool

Dr Fiona Moss
Thames Postgraduate Medical & Dental Education

Professor Dianne Newham
King's College London

Professor Gillian Parker
University of Leicester

Dr Mary Renfrew
University of Oxford

Ms Hilary Scott
Tower Hamlets Healthcare NHS Trust, London

* Previous Chair

National Coordinating Centre for Health Technology Assessment, Advisory Group

Current members

Chair:
Professor John Gabbay
Wessex Institute for Health Research & Development

Dr Sheila Adam
Department of Health

Professor Nicholas Black
London School of Hygiene and Tropical Medicine

Professor Martin Buxton
Health Economics Research Group, Brunel University

Mr Harry Cayton
Alzheimer's Disease Society

Professor Angela Coulter
The King's Fund, London

Professor Paul Dieppe
MRC Health Services Research Collaboration, University of Bristol

Professor Mike Drummond
Centre for Health Economics, University of York

Professor Shah Ebrahim
MRC Health Services Research Collaboration, University of Bristol

Ms Lynn Kerridge
Wessex Institute for Health Research & Development

Professor Jos Kleijnen
NHS Centre for Reviews and Dissemination, University of York

Dr Ruairidh Milne
Wessex Institute for Health Research & Development

Ms Kay Pattison
Research & Development Directorate, NHS Executive

Professor James Raftery
Health Economics Unit, University of Birmingham

Professor Ian Russell
Department of Health Sciences & Clinical Evaluation, University of York

Dr Ken Stein
North & East Devon Health Authority

Professor Andrew Stevens
Department of Public Health & Epidemiology, University of Birmingham

Professor Kent Woods
Department of Medicine & Therapeutics, University of Leicester

Past member

Dr Paul Roderick
Wessex Institute for Health Research & Development

HTA Commissioning Board

Current members

Chair:
Professor Shah Ebrahim
Professor of Epidemiology of Ageing, University of Bristol

Professor Doug Altman
Director, ICRF Medical Statistics Group, Centre for Statistics in Medicine, University of Oxford

Professor John Bond
Director, Centre for Health Services Research, University of Newcastle-upon-Tyne

Mr Peter Bower
General Manager and Independent Health Advisor, Thames Valley Primary Care Agency

Ms Christine Clark
Honorary Research Pharmacist, Hope Hospital, Salford

Professor Martin Eccles
Professor of Clinical Effectiveness, University of Newcastle-upon-Tyne

Dr Mike Gill
Regional Director of Public Health, NHS Executive South East

Dr Alastair Gray
Director, Health Economics Research Centre, University of Oxford

Professor Mark Haggard
Director, MRC Institute of Hearing Research, University of Nottingham

Dr Jenny Hewison
Senior Lecturer, Department of Psychology, University of Leeds

Professor Alison Kitson
Director, Royal College of Nursing Institute

Dr Donna Lamping
Senior Lecturer, Department of Public Health, London School of Hygiene & Tropical Medicine

Professor Alan Maynard
Joint Director, York Health Policy Group, University of York

Professor David Neal
Joint Director, York Health Policy Group, University of York

Professor Jon Nicholl
Director, Medical Care Research Unit, University of Sheffield

Professor Gillian Parker
Nuffield Professor of Community Care, University of Leicester

Dr Tim Peters
Reader in Medical Statistics, Department of Social Medicine, University of Bristol

Professor Martin Severs
Professor in Elderly Health Care, University of Portsmouth

Dr Sarah Stewart-Brown
Health Service Research Unit, University of Oxford

Professor Ala Szczepura
Director, Centre for Health Services Studies, University of Warwick

Dr Gillian Vivian
Consultant, Royal Cornwall Hospitals Trust

Professor Graham Watt
Department of General Practice, University of Glasgow

Professor Kent Woods
Professor of Therapeutics, University of Leicester

Dr Jeremy Wyatt
Senior Fellow, Health Knowledge Management Centre, University College London

Past members

Professor Ian Russell*
Department of Health Sciences & Clinical Evaluation, University of York

Professor Charles Florey*
Department of Epidemiology & Public Health, Ninewells Hospital & Medical School, University of Dundee

Professor David Cohen
Professor of Health Economics, University of Glamorgan

Mr Barrie Dowdeswell
Chief Executive, Royal Victoria Infirmary, Newcastle-upon-Tyne

Dr Michael Horlington
Head of Corporate Licensing, Smith & Nephew Group Research Centre

Professor Sir Miles Irving
Professor of Surgery, University of Manchester, Hope Hospital, Salford

Professor Martin Knapp
Director, Personal Social Services Research Unit, London School of Economics & Political Science

Professor Theresa Marteau
Director, Psychology & Genetics Research Group, Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Professor Sally McIntyre
MRC Medical Sociology Unit, Glasgow

Professor David Sackett
Centre for Evidence Based Medicine, Oxford

Dr David Spiegelhalter
MRC Biostatistics Unit, Institute of Public Health, Cambridge

Professor David Williams
Department of Clinical Engineering, University of Liverpool

Dr Mark Williams
Public Health Physician, Bristol

* Previous Chair

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.hta.nhsweb.nhs.uk

ISSN 1366-5278