



SUPPLEMENTARY GUIDANCE FOR AUTHORS

UNDERTAKING REVIEWS WITH THE COCHRANE CONSUMERS AND COMMUNICATION REVIEW GROUP

Welcome to the Cochrane Consumers and Communication Review Group (CC&CRG) supplementary guidance for review authors on study design, study quality and analysis. Send feedback or queries about this document to cochrane@latrobe.edu.au

We present this guidance in three documents:

- 1: Study design guide
- 2: Study quality guide
- 3: Study analysis guide [*available in draft only from Cochrane@latrobe.edu.au*]

These documents are available online at

<http://www.latrobe.edu.au/chcp/cochrane/resources.html>

This guidance is *not* intended to replace the information published by The Cochrane Collaboration. Rather, it attempts to supplement the information for authors and to address in more detail some of the issues specific to reviews of complex interventions covered by the CC&CRG scope. For more information about the CC&CRG visit <http://www.latrobe.edu.au/chcp/cochrane/index.html> or read about us on *The Cochrane Library*.

We have cross-referenced each document with other materials produced by The Cochrane Collaboration. It is essential that review authors also consult these materials throughout the review process, as this additional guidance is not intended as a substitute guide for the review process.

COCHRANE CONSUMERS & COMMUNICATION REVIEW GROUP

STUDY DESIGN GUIDE

FOR REVIEW AUTHORS

CONTENTS

<i>Overview: Common Problems – checklist for review authors</i>	4
<i>1.0 Introduction</i>	5
<i>1.1 What is a 'study design'?</i>	6
<i>1.2 Study design and Cochrane reviews</i>	8
<i>1.3 Study design and reviews of complex interventions</i>	10
1.3.1 General study design issues.....	10
1.3.2 Randomised controlled trials (RCTs).....	14
1.3.2.1 Randomisation methods.....	15
1.3.2.2 Concealing the random sequence: allocation concealment.....	18
1.3.2.3 Approaches to randomisation.....	17
1.3.2.4 Summary: Key points for authors to consider about RCTs.....	20
1.3.3 Cluster randomised controlled trials (cluster RCT, C-RCT).....	22
1.3.3.1 Summary: Key points for authors to consider about cluster RCTs.....	23
1.3.4 Non-randomised studies.....	24
1.3.4.1 Quasi-randomised controlled studies.....	25
1.3.4.2 A note about Controlled Clinical Trials (CCTs).....	26
1.3.4.3 Controlled before-and-after studies (CBAs).....	26
1.3.4.4 Interrupted time series (ITS) designs.....	28
<i>1.4 Study design and study quality</i>	29
<i>1.5 Quantitative and qualitative study designs</i>	32
<i>Appendix A – Study design terms</i>	35
<i>Appendix B – Research design and methods terms</i>	40
<i>Appendix C – Additional sources of information for review authors</i>	46
<i>Appendix D – Additional information about study design</i>	49
D.1 Descriptive and comparative study designs.....	49
D.1.1 Descriptive study designs.....	49
D.1.2 Comparative study designs.....	50
D.1.2.1 Observational studies.....	50
D.1.2.2 Experimental studies.....	51
D.1.3 Summary table of study designs.....	52
D.1.4 A note on timing and study design.....	53
<i>Appendix E – Qualitative research and study design</i>	54
Qualitative research designs.....	54
Qualitative research methods.....	55
<i>References</i>	56

This guide was prepared by Rebecca Ryan, using a first draft prepared by Dominique Broclain, and edited by Dominique Broclain, Dell Horey, Sandy Oliver and Megan Prictor. Additional feedback from editors of the Consumers and Communication Review Group was also incorporated into the paper at various stages. Additional figures on design included in text were adapted from a paper on study designs prepared by Alex Broom and Neil Thalagala for the Cochrane Consumers and Communication Review Group. This paper was approved by Sophie Hill, Coordinating Editor of the Consumers and Communication Review Group on 5 March 2007. **Last updated: 30 July 2009**

How to cite this Guide:

Ryan R, Hill S, Broclain D, Horey D, Oliver S, Prictor M; Cochrane Consumers and Communication Review Group. *Study Design Guide*. July 2009.

<http://www.latrobe.edu.au/chcp/cochrane/resources.html> (accessed *DATE*).

People using this guide should check the website

<http://www.latrobe.edu.au/chcp/cochrane/resources.html> regularly to ensure they have the latest version.

OVERVIEW: COMMON ISSUES – CHECKLIST FOR REVIEW AUTHORS*Particularly at protocol stage:*

- **Issue:** *What types of studies should I include in the review?*
Action: Assess what type of question your review is attempting to address. Check appropriateness of different study designs for answering this type of question. Justify your decision to include study types in addition to randomised controlled trials (RCTs) alone if necessary. A rationale for this decision must be made at the protocol stage of the review.
- **Issue:** *Terminology – Why can't I include Controlled Clinical Trials (CCTs)?*
CCT is an indexing term. Trials allocating participants to intervention and control groups using quasi-random methods should be described as quasi-randomised controlled trials rather than CCTs.
- **Issue:** *Should qualitative research be included in the review?*
Action: Consider what the purpose of including qualitative research in the review might be. If it can inform the review or add depth to it, provide this as a justification for considering this research in addition to quantitative research, in your review.

Particularly at review stage:

- **Issue:** *Lack of reporting of details of the study design and execution in the published study.*
Action: Contact authors of all included studies and borderline excluded studies for further information about their study. Include requests for specific additional details, such as details of the randomisation process, numbers of participants etc. (Ask the Managing Editor for a sample letter).
- **Issue:** *Was the study really an RCT?*
Action: Check the [list of approaches considered to be truly random](#) and assess the study based on these. If you need more information about the study, contact the study authors.
- **Issue:** *Was allocation to intervention and control groups adequately concealed until participants are assigned to groups?*
Action: Check the [list of approaches considered adequate to conceal allocation](#) and make a judgement based on these approaches. If you need more information about the study, contact the authors.
- **Issue:** *Various study types were included in my review. How should I deal with the results of different study designs?*
Action: Consider separately the findings from each type of study, then systematically compare the results across study types, identifying key similarities and differences. Consider how differences in study designs may have influenced the overall results of your review.

1.0 INTRODUCTION

The primary purposes of this guide are:

- To give review authors and referees a general understanding and supplementary information about study design issues, relevant to preparing or assessing a Cochrane systematic review;
- To assist review authors to determine which study designs to include in their review.

This guide is essential reading for review authors at the title development stage, and also at the protocol stage.

Structure

Section 1.1 is a general introduction to study designs and is intended primarily for authors without formal training in epidemiology and/or research.

Section 1.2 describes particular study designs that may be included in Cochrane systematic reviews.

Section 1.3 provides detailed advice on study designs eligible for inclusion in systematic reviews of complex interventions, such as those published by the Cochrane Consumers & Communication Review Group (CC&CRG). This includes a detailed outline of both randomised and non-randomised study designs.

Section 1.4 describes study design and study quality and the relation between them. For more information see the **Study Quality Guide**.

Section 1.5 describes some of the ways in which qualitative research can be used in systematic reviews.

- Appendices A and B to this document contain glossaries of useful terms for study design and the characterisation of study design, adapted primarily from the *Cochrane Handbook for Systematic Reviews of Interventions, Glossary of terms* (see www.cochrane.org/resources/glossary.htm), with additional material from other sources, as referenced.
- Appendix C contains additional sources of information for authors, with brief descriptions of the key points of the sources.
- Appendix D contains additional background information on understanding research and study design and may be useful for those seeking additional contextual information about research design.
- Appendix E contains detailed information about the aims and methods of qualitative research.

This document should be read in conjunction with the [Cochrane Handbook for Systematic Reviews of Interventions](#) (Higgins 2008), which gives extensive guidance on the conduct of systematic reviews.

Authors must read the Handbook before embarking on their review and should refer to it throughout the review process.

Readers are also referred to The Cochrane Collaboration *Open Learning Material for Reviewers*, Version 1.1, November 2002, which provides additional information on methodological issues in conducting a systematic review (see www.cochrane-net.org/openlearning/index.htm).

1.1 WHAT IS A 'STUDY DESIGN'?

- See the [Cochrane Handbook for Systematic Reviews of Interventions](#), Section 5.5
- See the [Cochrane Open Learning materials](#), Modules 1 and 5

The study design is the **overall** structure of a research plan. In other words, it is the general strategy adopted in order to answer a specific research question. In a scientific publication, the methods section includes a description of the study design, that is, the specific methods or techniques that are used to collect, process and analyse the data.

The particular study design adopted should be closely related to the type of research question(s) that it aims to answer. As Harris has written:

'Research designs are plans that guide investigators in the process of collecting, analyzing and interpreting observations. The design of a study is determined by the purposes of the authors, the nature of their questions, the resources and skills available to them. When designing a study, decisions are made relating to sampling and case definitions, data collection strategies, data management, analysis, interpretation and reporting' (Harris, 2002).

Designing a research strategy therefore involves making several decisions, including the following:

- Identifying the target population and how they should be sampled, if necessary;
- Deciding whether data are needed on what is already happening to the target population (that is, at baseline);
- Determining what might happen if some sort of intervention, or change, is introduced; and
- Deciding whether a control group is needed to assess these changes.

It is important to note that different study designs are appropriate to answering different types of research questions. Consider which study designs will best answer your specific review question. The choice of which types of studies you will include has important implications for subsequent stages of review. It will influence the design of the search strategy; decisions about the inclusion and exclusion criteria for studies; the choice of appropriate quality assessment criteria; and the analysis of results. **It is essential that you choose the types of studies to include in your review on the basis of the review question, and the methodological appropriateness of particular research designs for answering these specific questions - not the other way around.**

Because of this, it is important to have an idea about the purpose of different research designs and their appropriateness for addressing the particular research question asked by a systematic review.

Certain study designs are better than others for answering particular types of research questions, so considering the type of question that you are trying to answer will help you to decide which studies should be included (see table 1). For example, randomised controlled trials (RCTs) are considered to be the best type of study for addressing questions about the effectiveness of therapy (or of interventions). Other research questions are best addressed by other study designs. The Cochrane Collaboration produces reviews of the effects of interventions and therefore most reviews done under its auspice include principally RCTs. There are, however, many reviews which include controlled trials^Ω, ie, randomised and quasi randomised controlled trials for instance.

Table 1: A description of the purpose of different study designs.

PURPOSE OF THE STUDY (RESEARCH QUESTION)	MOST APPROPRIATE STUDY DESIGN
Effectiveness	Controlled trial ^Ω
Harm*	Case Control Cohort
Prognosis	Cohort
Aetiology/ Risk factors*	Case control Cohort Cross-sectional survey
Prevalence	Cross-sectional survey
Meaning	Qualitative data collection and analysis

For guidance on how to develop clear, well-formulated questions to be addressed by systematic reviews, see the [Cochrane Handbook for Systematic Reviews of Interventions](#), Chapter 5.

^Ω The term **controlled trial** here is used to generally mean studies that include a control group; that is, both randomised studies and non-randomised studies. The differences between these types of study design are discussed in more detail in later sections.

* In these cases, a **cohort study** may be most appropriate when the exposure is rare but the outcome of interest is common; whereas a **case-control study** design may be most appropriate where the exposure is common but the outcome of interest is rare.

1.2 STUDY DESIGN AND COCHRANE REVIEWS

- See the [Cochrane Handbook for Systematic Reviews of Interventions](#), Chapter 5.5
- See the [Cochrane Collaboration Open Learning materials](#), Module 1

As discussed above, the type of research question determines the particular study design that is most appropriate for answering that question. A systematic review of the effects of an intervention (such as a Cochrane review) is therefore a summary of all of the studies that have tested a particular intervention, to see what effect(s) it has.

Most Cochrane reviews are reviews of RCTs. There are both epistemological and practical reasons for this. First, in the epistemology of experimental studies, RCTs are considered to provide the strongest measure of whether an intervention has an effect. This is because there are particular aspects of the RCT design which help to minimise **bias**¹. (See **Section 1.3.1** and the Group's **Study Quality Guides** for more information about issues of study design and bias).

Practically, there is now international consensus that it is possible to identify a majority of RCTs from the world's databases. This is important, as systematic reviews attempt to identify **all** known research on a particular research question. Having a substantial proportion of RCTs catalogued on databases worldwide can allow identification of many of the relevant studies via systematic searching. Substantial work has now been performed to develop ways of consistently identifying RCTs, but this has not yet been matched in the systematic identification of studies of other designs².

Furthermore, as much work has been conducted in the area, a degree of consensus now exists on how to appraise the quality of RCTs, and how to combine the results of different RCTs into a summary or pooled effect estimate using meta-analysis. These are both important elements of systematic reviews. A comparable level of consensus on these issues for studies other than RCTs does not yet exist, nor is there clear agreement about how the combination of results could practically be done. Cochrane reviews are therefore reviews of the effects of an intervention, and in the main, the analysis contained in these reviews is generally an analysis of data extracted from RCTs.

¹ **Bias** refers to any error that causes results to differ systematically from the 'true' results. In studies of the effects of healthcare interventions there are a number of major sources of bias. These are covered in more detail in the **Study Quality Guide**; see also the [glossary of terms in Appendix B](#).

² Review authors should note however, that while identification of all relevant published studies is essential to produce a high-quality systematic review, that this alone is not sufficient to be sure that all relevant research has been identified and included. Many studies remain unpublished, particularly those with negative results, and reviews that fail to search for relevant unpublished studies therefore tend to overestimate the effects of an intervention (publication bias). This concept is covered in more detail in the following **Study Quality Guide, Section 2.3.5**.

It is possible, however, to prepare a systematic review that includes a slightly broader range of experimental study designs than RCTs alone. This might be appropriate, for example, in situations where few RCTs have been conducted or where it is not possible (for ethical or practical reasons) to randomise people to receive an intervention. It is also possible to extend the method of systematic reviewing to prepare a systematic review of observational studies. However, consensus on how to perform these kinds of systematic reviews is only now being developed.

Researchers often ask: why don't Cochrane reviews contain qualitative research? This question is particularly important for people working in the area of communicating with and involving consumers because qualitative research can help us to understand the communication experience so well. Some Cochrane reviews now do contain qualitative research, but it is not necessarily included in the same way that data from RCTs is included in reviews. For instance, qualitative research can assist in developing the Background section for the review so that there is an in-depth understanding of the context for the intervention being investigated. Qualitative research can help to identify important outcomes that should be measured in an experimental study of an intervention. It can also help us to interpret the results of RCTs. There is now a small number of Cochrane reviews in preparation that aim to include qualitative evaluations in the analysis section of the review. However, this is truly experimental! There is, as yet, little consensus on how to prepare a review in this way and little consistent guidance is currently available.

1.3 STUDY DESIGN AND REVIEWS OF COMPLEX INTERVENTIONS

- See the [Cochrane Handbook for Systematic Reviews of Interventions](#), Chapter 5 and 13
- See the [CRD Paper](#), especially section 1.3.4 on quality assessment.

1.3.1 GENERAL STUDY DESIGN ISSUES

There are many different types of study design and ways of classifying them, which can lead to confusion. The following sections give a basic outline of study designs that may be eligible for inclusion in systematic reviews of complex interventions, while the additional sources of information referred to throughout may also be of interest to authors. *Appendix D* of this document contains further information on study and research design.

Within The Cochrane Collaboration, most review groups advise review authors that systematic reviews should only include RCTs. As discussed earlier, there are several reasons for this recommendation. In practice, depending on the study question and the research available, review authors may decide to include both randomised and non-randomised studies in their review. The decision about the types of study designs to include in a review is determined primarily by the review question. For example, many complex interventions, such as those to do with communicating with consumers or those examining organisational change in healthcare delivery³ or public health interventions may not be easily or appropriately addressed by RCTs. In these cases, it may be appropriate to include additional study design types in a review. This is discussed further below.

In cases where reviews include both RCTs and other study designs, analysis of the effects of the intervention, and of the quality of the individual studies, are examined separately for the different study types⁴. This approach takes into account the fact that RCTs provide a stronger or higher **level of evidence** than other designs when examining the effectiveness of healthcare interventions.

The strength or level of evidence that study designs provide can be ranked hierarchically, depending on the question that is being asked. For example, studies that best answer questions of effectiveness (the effects of interventions) differ from those that are best for answering questions about accuracy of diagnostic tests.

³ For example, the systematic reviews published by the Cochrane Effective Practice and Organisation of Care (EPOC) Review Group.

⁴ Examining different study types separately within a review can also allow clear comparison to be made in later sections of the review. For example, it can be useful to compare the findings of RCTs and non-RCTs, to determine whether they give conflicting or consistent information about the effects of an intervention. Such a comparison may help to add depth to the review by helping to clarify some potential reasons for differences (or similarities) if they exist across different study designs.

Table 2 gives an indication of the best type of study design (the strongest level of evidence) to answer different types of healthcare questions.

Table 2: Type of study design appropriate to answering different research questions.

TYPE OF QUESTION	TYPE OF STUDY DESIGN
Intervention (therapy)	RCT
Diagnosis/ screening: Accuracy of the test Effect of test on health outcomes	Cohort study RCT
Prognosis	Cohort study (longitudinal)
Aetiology/ risk factors	RCT Cohort study ^r Case-control study ^s

At present, Cochrane reviews focus on assessing questions about the effects of healthcare interventions. Questions about effectiveness tend to focus on **comparisons**⁵; for example, how one intervention performs compared with another intervention; or how an intervention compares with no intervention. To answer these questions, comparative studies that **minimise bias** provide the strongest level of evidence. For more detailed information about the differences between comparative and descriptive⁶ study designs, see **Appendix D section D.1.1 and D.1.2**.

RCTs are highly regarded because they are designed specifically to minimise bias. By randomly assigning participants to groups, for example, the likelihood that the groups will differ on important baseline characteristics (whether the investigator knows them or not) is minimised. By chance it is likely that the two groups will be equivalent on important characteristics at the start of the study. Any differences between groups measured at the end of the study (after the intervention has been delivered) can therefore be attributed to the effects of the intervention being examined – rather than to differences existing between the groups that may be independent of the intervention, or that may have already existed before the intervention was given. The main issue, therefore, with allocating

^r **Cohort studies** can be useful for examining rare exposures with common outcomes (see [Appendix A](#)).

^s **Case-control studies** can be useful for studying rare outcomes with common exposure (see [Appendix A](#)).

⁵ Study designs aiming to examine causal relationships (whether a change in one variable causes change in another) are based on comparisons – for example, comparing the same participants before and after an intervention; or comparing two or more groups that have received different treatments. This is in contrast to descriptive studies, which record observations to look for associations between variables. **Comparative study** designs can be either observational or experimental.

⁶ **Descriptive studies** record observations (such as opinions, behaviours, clinical measurements, and so on) in a defined population. These study designs can therefore look for associations (for example, the relationship of a disease with a particular characteristic in a certain group of people). They do **not** aim to examine causal relationships (whether a change in one variable, the intervention, caused changes in another variable, the outcome).

people to the different groups in an experimental study is the potential for bias. If the two groups are different before the experiment even really begins, (before they are given the intervention that we want to measure the effects of) there may be problems in interpreting the effect of the intervention. For example, you might be interested in examining whether giving women more information about their choices before a caesarean delivery affected certain outcomes (such as their knowledge and understanding, satisfaction with care, or clinical outcomes). If the allocation is not random and women who attended public health clinics were assigned to one group (eg the intervention group) and women who employed private obstetricians were assigned to the other group (the control group), it would be difficult to tell whether the intervention alone affected the results (outcomes), or whether some of other factor was responsible for the changes. Participants in the intervention and control groups may differ substantially in terms of educational level, age, number of previous children, and many other factors. Such differences in characteristics of the women in the two groups at baseline could therefore contribute to the outcomes measured. This can produce misleading results.

In comparison, the methods of allocating participants in other study designs may not be as 'fair,' because they do not rely on chance alone. The study groups that are formed using non-random methods of allocating participants - such as alternating group assignment - may differ substantially on important characteristics at baseline. This can then distort the results of the study.

Cochrane systematic reviews assess the effects of healthcare interventions. Study designs examining the effects of interventions are typically based on comparisons. However, comparative study designs can themselves be either observational⁷ or experimental⁸ and many different ways of classifying or ranking studies based on their design have now been developed.

Figure 1 (next page) shows a broad classification scheme or hierarchy of studies in terms of the suitability of their design for answering questions of effectiveness⁹ - and hence how suitable they are for inclusion in a Cochrane systematic review.

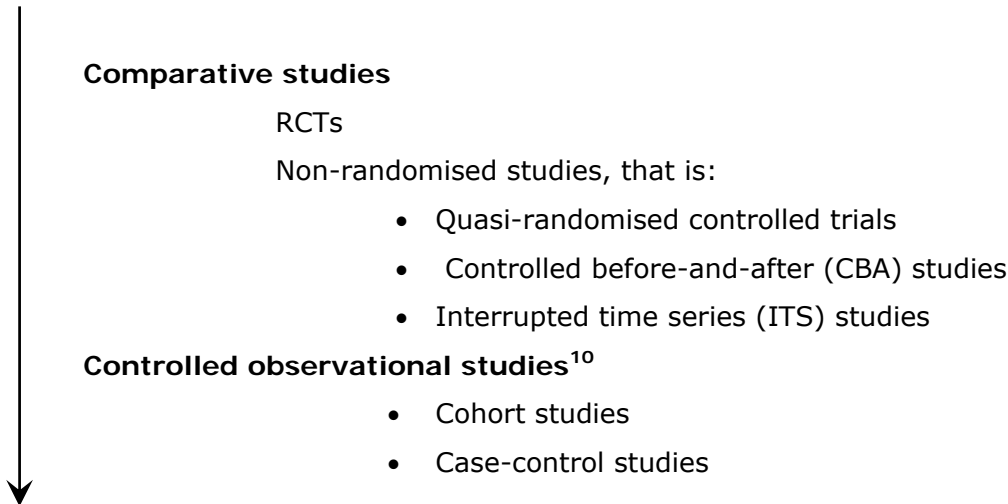
⁷ **Observational studies:** A study is considered to be observational when nature is allowed to take its course; that is, where the investigator does *not* actively manipulate or control what happens to the groups that are being studied. In these studies, variations that naturally exist in terms of interventions or exposure to a particular factor are studied for their effects on health outcomes. For example, this kind of study might record whether people exposed to a particular chemical died or not. In these studies, the investigator does not actively control who gets the intervention and who does not; instead the investigator looks for associations between a particular outcome and who got the intervention. Observational studies are not confined only to those situations where nature takes its course; they can be about 'usual practice;' such as comparing outcomes among pupils in two different schools. It is important to note that in observational studies the experimenter does not actively manipulate the participants' exposure; they simply observe the occurrence of two or more factors. (See also [Appendix A](#)).

⁸ **Experimental studies:** A study is considered to be experimental when the investigator controls some of the conditions of the study - for example, how people are allocated to the different study groups that are to be compared. Randomised and non-randomised controlled trials are examples of experimental study designs.

⁹ It should be noted that any hierarchy of study designs depends on how well the studies are conducted, and should be treated **as a guide only**. For example, if there are serious flaws or limitations in the design or conduct of an RCT, it may be considered of lower quality (ie. results may be less likely to reflect the 'truth,' and so, should be regarded as less reliable).

Reviews within the scope of the Cochrane Consumers and Communication Review Group include interventions for which an RCT may not exist. Sometimes interventions have not or cannot be subjected to an RCT or even to a quasi-RCT for practical or ethical reasons. For example, if the course of a particular disease is uniformly life-threatening without treatment, it would not be ethical to intentionally assign people randomly to receive no intervention (for instance, a placebo pill).

FIGURE 1: GENERAL HIERARCHY OF STUDY DESIGNS TO ANSWER QUESTIONS OF EFFECTIVENESS



The Consumers and Communication Review Group will follow the advice developed by the Cochrane Effective Practice and Organisation of Care (EPOC) Review Group (see also <http://www.epoc.cochrane.org/> or email amayhe2@uottawa.ca)¹¹. This advice allows review authors to decide whether to include a limited range of experimental study designs **in addition to RCTs** in their review.

These guidelines state that the following study designs are eligible for consideration for inclusion in systematic reviews of complex interventions:

- RCTs (including cluster RCTs)
- Non-randomised studies
 - Quasi-randomised controlled trials
 - Controlled before-and-after (CBA) studies
 - Interrupted time series (ITS) studies

¹⁰ **Controlled observational studies:** such as cohort and case-control studies, include a control group, as well as a group which has received the intervention or exposure that is of interest. For more information refer to [Appendix A](#).

¹¹ EPOC reviews examine the effects of complex organisational healthcare interventions. The scope of the Consumers and Communication Review Group also often includes complex interventions, but these tend to focus on the individual consumer.

For some research questions, RCTs can provide a rigorous assessment of a particular intervention, and there are RCTs available to be systematically reviewed. There are also many instances however, where RCTs may not be possible; may not be ethical; or may simply not exist.

In addition, there are also many kinds of research question that may benefit from the inclusion of non-randomised studies. These include many public health and health promotion interventions, interventions covered by the scope of the EPOC Group and many of those of interest to the Consumers and Communication Group.

The decision to include non-randomised studies in a Cochrane review must be made on the basis of whether the study designs are appropriate for answering the review's particular research question. It should be decided in advance of conducting the review, and a rationale for this decision must be provided at the protocol stage.

In providing a rationale, you should consider the following issue amongst others. The formulation of review questions concerned with complex interventions are a balance of research with pragmatic issues. For instance, if you are aware that many trials have been conducted, then you might wish to include only RCTs in your review. If, however, few trials have yet to be conducted, then you may wish to broaden inclusion to less rigorous study designs, such as non-randomised studies. You may wish to provide people in the field with some guidance, ie the best available evidence from experimental studies, or evidence regarding the promise of new interventions. It is possible to revise inclusion criteria to make them narrower, if your search identifies many RCTs. This step just requires approval and then reporting in the review.

The following sections provide definitions and an outline of each type of study design eligible for inclusion in Cochrane Consumers and Communication Group systematic reviews. Review authors are also directed to key points to consider when making decisions about the study types to include in their review, and to relevant supplementary materials that may be helpful when making these decisions.

1.3.2 RANDOMISED CONTROLLED TRIALS (RCTS)

- See [Cochrane Handbook for Systematic Reviews of Interventions](#), especially Ch 8.
- See [Jadad \(1998\)](#): Chapters 1 and 2

In an RCT, individual participants are randomly assigned to either an **intervention group** or to a **control group**¹². Randomisation ensures that participants in each group should, at least theoretically, differ only in their exposure to the intervention – all other measurable characteristics (such as gender, age, educational level, smoking status, etc) and those that can not be easily measured (such as attitude, personal beliefs, etc) should, by chance, be distributed equally between the intervention and control groups.

Similarly, randomisation should produce groups of participants that are equivalent on characteristics that are known about (for example, age, sex, education) as well as on key characteristics that are not known, or **confounders**¹³. That is, both the known and unknown determinants of the outcome will (theoretically) be equally distributed across the groups – rather than being systematically different between the groups. In theory, the only difference between groups in an RCT will therefore be whether they received the intervention or not.

1.3.2.1 RANDOMISATION METHODS

There are several methods that can be used to randomly assign participants to groups to be studied in an RCT. **Randomisation is the process of producing the sequence of random numbers.** This sequence of random numbers, once generated, is used to assign participants to either the intervention or control group (the groups to be studied and compared).

Some randomisation methods are considered to generate a truly random sequence of numbers, while other approaches are considered inadequate. This is based on a judgement that:

- Each participant had an *equal chance* of being assigned to the control group or the intervention group;
- The investigators *can not predict* which group each participant will be assigned to (that is, the random number sequence is not predictable).

What is so important about the process of randomisation is that each participant in the study has an equal chance of being assigned to either the intervention group or to the control group. True randomisation means that allocation is not decided by the clinician or the participant, and that assignment to one group or another is not predictable.

¹² In trials comparing at least one intervention, people in the control group can receive no intervention, usual care, placebo, or another form of care, depending on what the intervention is. For example, in an experimental trial investigating the effects of a particular drug, the control group might receive a placebo pill to take into account (control for) the intervention (the active pill). In comparison, a trial investigating the effects of psychological counselling (the intervention), might use a control group that receives 'usual care.' When considering comparative studies (randomised or non-randomised), it is important to note what sort of control group has been used, and whether it is appropriate for the particular intervention under investigation.

¹³ A **confounder** is a variable that can be due to chance or bias, and can cause, prevent or influence the outcome of interest. It is not an intermediate variable associated with the interventions and outcomes being investigated, and unless it is controlled for (for example, by randomly assigning participants to treatment groups), it cannot be distinguished from the effects of the intervention being studied. This may introduce bias. See also [Appendix B](#).

This process of allocation is a major distinction between RCTs and non-randomised study designs. While the process of allocating participants to groups is only one step in designing and conducting a study, the way that participants are allocated to groups has important implications for the study, its strengths and weaknesses overall, and ultimately the study's potential for bias (this is discussed in more detail in the *Study Quality Guide*).

Adequate randomisation approaches include:

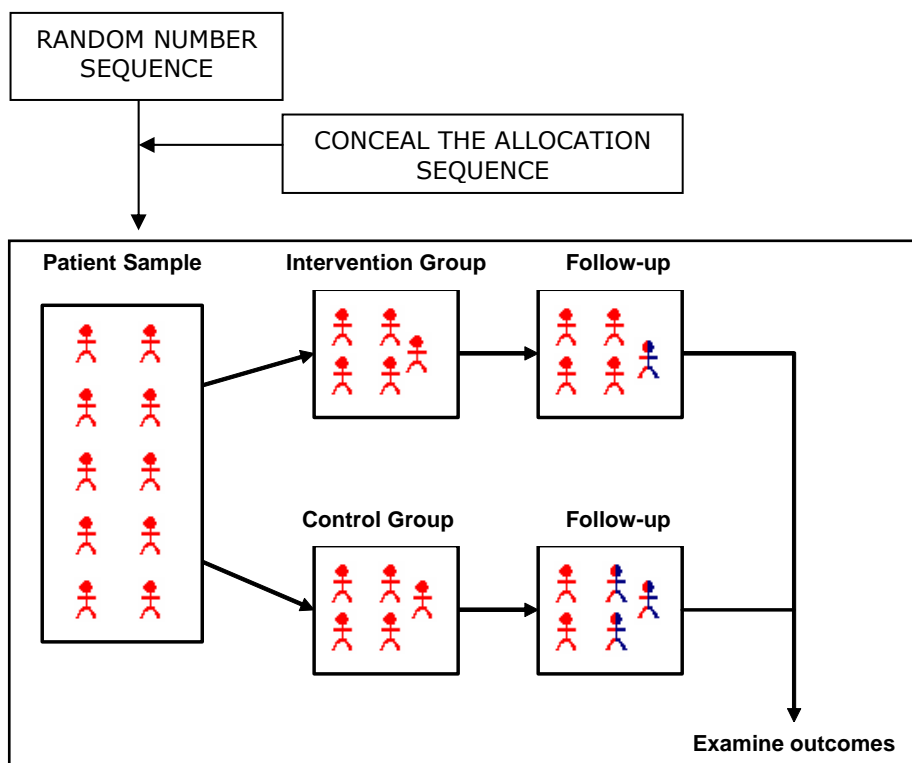
- Computer-generated random numbers
- Random numbers tables
- Coin toss (for a trial with two groups) or die toss (where there are more than 2 groups).

Refer to the [Cochrane Handbook](#) chapter 8.9.2.1 for more information.

Note: The Cochrane Collaboration considers trials with adequate randomisation to be randomised controlled trials (RCTs).

Review authors should also note, however, that **not all studies that describe themselves as RCTs are considered to truly be RCTs**, based on the criteria specified above: review authors must independently decide whether the methods used by a study to randomise participants is truly random or not.

Figure 2: The randomised controlled trial.



Trials that attempt to randomly assign participants to groups but use an inadequate approach to generate the random sequence are designated as **quasi-randomised controlled trials**. Such trials do attempt to randomly allocate participants with the intent of producing equivalent groups, but the randomisation methods used are not adequate because in practice they are relatively easy to manipulate or predict.

This can introduce bias if, for example, clinicians conducting the trial have certain patients that they think would particularly benefit from the intervention to be tested. In such a case, they might try to ensure that these select patients are allocated to receive the intervention. If the random number sequence used to assign participants to groups really is random, the clinician would not be able to predict to which group each participant would be assigned. On the other hand, if alternation or some other technique that is predictable (not truly random) is used to assign participants to groups, doctors might be able to manipulate which patient is assigned to the different groups. This might mean, for example, that patients who are sicker, or those who are less ill, or those who are more motivated to comply with treatment, are assigned to the treatment (intervention) group. Any manipulation of the randomisation process can therefore introduce systematic differences in the characteristics of participants in the different study groups at the beginning of the study: this can introduce bias.

Inadequate randomisation approaches:

- Alternation
- Case record numbers
- Birth dates
- Week days or month of the year

Note that trials with inadequate randomisation are considered to be quasi-RCTs.

1.3.2.2 APPROACHES TO RANDOMISATION

- See [Jadad \(1998\)](#), Chapter 1 for further information

There are different ways of generating random number sequences for the purpose of allocating participants to different treatment groups. The discussion above refers to '**simple randomisation**,' where a single sequence of random numbers is generated and used to assign individual participants to

either the control or intervention groups. If the sample is large, simple randomisation can be adequate to produce groups that are basically comparable on both known and unknown prognostic factors¹⁴.

There are also other approaches to randomisation, such as **restricted (or block)** and **stratified randomisation** which can be used to control the randomisation procedure to ensure that the groups formed by randomisation are actually balanced (for example, in terms of group size (numbers) or with respect to particular participant characteristics).

- **Restricted randomisation ('blocking')**: Can be used to ensure that groups of approximately equal size are formed. It involves creating 'blocks' of sequences, which ensures that roughly the same number of participants is assigned to each of the groups within each of these blocks.
- **Stratified randomisation**: Can be used to help to balance the participants' characteristics in each of the study groups. It involves using a separate randomisation procedure within subsets of participants, in order to match groups approximately on key characteristics (for example, age, sex, disease severity). To perform stratified randomisation, the researchers must first identify key factors, or strata, that are known to be related to the outcome(s) of the study. Separate block randomisation schemes are then formed for each of these strata, so that resulting groups are balanced within each.

One other approach, known as **weighted or unequal randomisation**, can be used to allocate unequal numbers of participants to study groups. This can be useful, for example, in studies where there might be a high probability of adverse effects associated with the intervention. In such a case, investigators might wish to expose fewer participants to the intervention, for example, due to a lack of knowledge about adverse effects. In such cases, the investigators can use a weighted randomisation schedule to preserve the homogeneity of the participants' characteristics across groups. For example, in a weighted randomisation schedule where little is known about possible adverse effects of a drug, one third of participants might be allocated to received the active drug, with the remaining two thirds assigned to the control group.

1.3.2.3 CONCEALING THE RANDOM SEQUENCE: ALLOCATION CONCEALMENT

- See the [Cochrane Handbook for Systematic Reviews of Interventions](#), Chapter 8.10.

¹⁴ **Prognostic factors** are demographic or other characteristics (such as disease-related characteristics) that are associated strongly enough with a condition's outcomes to predict the development of those outcomes. These are different to **risk factors** (factors (such as smoking, for example) that are associated with an increased likelihood of developing a condition or disease). It is important to note however that although both prognostic and risk factors describe associations (that is, that certain factors can co-occur), these do *not* necessarily imply any kind of causal relationship (that is, that the occurrence of one factor causes another to happen).

Randomly assigning participants to groups helps to protect the study against **selection bias**¹⁵. However, to truly prevent selection bias, the process of allocating participants to groups (that is, actually placing participants into the groups to which they have been assigned using the random number sequence) must be concealed from those involved in the trial. This process is known as **allocation concealment** (see Figure 2).

Allocation concealment is important as it prevents the person enrolling participants into the RCT from knowing prior to the moment of allocation which group the participant has been assigned to. If the random number sequence is known, it is open to manipulation, and may result in systematic differences in the characteristics of participants allocated to the different arms (treatment groups) of the trial. If a clinician believes that a particular patient would especially benefit from receiving the intervention, they might manipulate the random number sequence to ensure that this patient was allocated to the intervention group (eg, selecting the intervention they want for specific enrolled patients). The allocation of participants to study groups would therefore no longer be truly randomised; and this can introduce bias. Empirical studies have now demonstrated that studies in which allocation concealment is inadequate tend to overestimate the effects of an intervention or treatment (Schultz *et al*, 1995).

While both random allocation and allocation concealment help to prevent selection bias, they are distinct aspects of the randomisation process. Assessing whether allocation was truly random (eg, table of random numbers) requires a judgement about whether each participant in the study had a truly equal chance of being assigned to any group in the trial. In comparison, assessing the adequacy of allocation concealment requires an evaluation of whether the sequence (generated by randomisation) was protected until the point of allocation (that is, so that it was not possible for anyone to interfere with the allocation sequence and the placement of participants into the groups to which they had been assigned by using the random number sequence).

It is also important to understand the difference between allocation concealment and blinding. **Allocation concealment is different to blinding**¹⁶. Allocation concealment is achieved when the randomisation sequence is concealed *before* and up until the point at which people are allocated to groups. This means that no-one should know who has been assigned to the different groups before it actually occurs. In comparison, blinding (whether of participants, providers or outcome assessors) refers to measures that are taken *after* people have been assigned to groups. This means that no-one knows which participant belongs to the different groups throughout the course of the study. An

¹⁵ **Selection bias**, as the name suggests, is bias that occurs at the point of selection, where comparison groups are systematically different (for example older, less sick, more educated, more compliant) in ways that can affect the study results. (See also [Appendix B](#)).

¹⁶ **Blinding** prevents the participants, providers and/or outcome assessors from knowing which participants are receiving the intervention(s) in the study. (See also [Appendix B](#)).

important difference between the two is that allocation concealment can *always* be implemented, whereas blinding cannot always be successfully achieved.

Approaches considered to adequately conceal allocation:

- Centralised allocation – such as allocation by a central office unaware of participants' characteristics; or pharmacy-controlled randomisation.
- Pre-numbered or coded identical containers which are administered serially (sequentially) to participants.
- On-site computer system, combined with allocations kept in a locked unreadable computer file that can be accessed only after the characteristics of an enrolled participant have been entered.
- Sequentially numbered, sealed, opaque envelopes.

Other approaches that are adequate to conceal allocation might be similar to those above, together with reassurance that the person who generated the randomisation sequence did not administer it. Some concealment schemes may be novel and not fit exactly with any of the above approaches. These may provide adequate concealment and must be independently judged by review authors. See the [Cochrane Handbook](#), chapter 8.10.2.1 and table 8.10.a for further guidance.

A sequence can only be properly concealed if it is not predictable. Therefore any method of generating the sequence that is not really random **cannot** be adequately concealed. These include:

- Alternation;
- Use of case record numbers, dates of birth or day of the week.

Any process that is entirely transparent before allocation, such as an open list of random numbers, is also considered inadequate (in terms of allocation concealment).

The Cochrane Collaboration recommends that when studies do not report their concealment approach, adequacy of concealment of allocation should be considered unclear.

For example, if a study only reports that a list or table was used; only specifies that sealed envelopes were used; or reports an apparently adequate concealment scheme together with other information that leads to some doubt over the adequacy of the approach; any of these approaches should be considered less than completely transparent and so are most accurately labelled as unclear. Review authors should contact triallists where possible to obtain further information about allocation concealment.

1.3.2.4 SUMMARY: KEY POINTS FOR AUTHORS TO CONSIDER ABOUT RCTS

1. Is the randomisation method really random?

Consider:

- Did each participant have an **equal chance** of being assigned to any group?

- Was the assignment **predictable**?

'Truly' random methods of generating the randomisation sequence (ie. methods that produce a non-predictable assignment pattern):

- Computer-generated random numbers
- Random number tables
- Coin toss (for a trial with two groups) or die toss (where there are more than two groups)

Note: Trials employing a truly random sequence generation method are designated as RCTs.

Inadequate approaches (methods that produce a predictable assignment pattern):

- Alternation
- Case record numbers
- Birth dates
- Week days

Note: Trials employing such sequence generation methods are designated as quasi-RCTs.

2. Was randomisation concealed until allocation?

Consider:

- Was the random number sequence (generated by randomisation) protected (unknown) until the point at which participants were actually put into their allocated groups?
- Was it possible for the random number sequence to be subverted? Was it possible for anyone to have influenced or tampered with the random number sequence, in order to change how participants were allocated to the intervention or control groups?

Adequate approaches (sequence for allocating participants is truly hidden from investigators)

- Centralised allocation – such as allocation by a central office unaware of participants' characteristics; or pharmacy-controlled randomisation
- Pre-numbered or coded identical containers which are administered serially (sequentially) to participants.
- On-site computer system, combined with allocations kept in a locked unreadable computer file that can be accessed only after the characteristics of an enrolled participant have been entered.
- Sequentially numbered, sealed, opaque envelopes.

Inadequate approaches (sequence may be accessed or predicted by investigators before allocation to groups has occurred)

- **Alternation;**
- **Use of case record numbers, dates of birth or day of the week; and**
- **Any process that is entirely transparent before allocation, such as an open list of random numbers.**

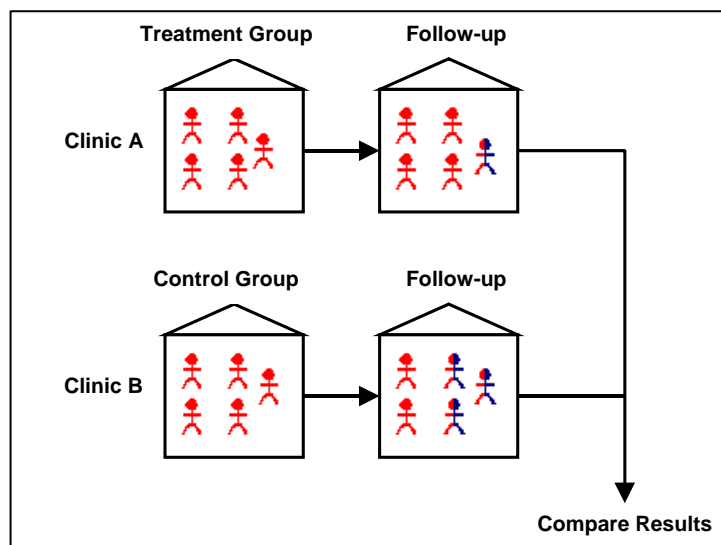
1.3.3 CLUSTER RANDOMISED CONTROLLED TRIALS (CLUSTER RCT, C-RCT)

- See the [Cochrane Handbook for Systematic Reviews of Interventions](#), Chapter 16.3.
- See the [Cochrane Open Learning materials](#) Appendix A2
- See [Jadad \(1998\)](#), Chapter 1

Using randomisation to allocate participants to intervention and control groups can be applied not only to individuals, as described above, but also to groups or 'clusters' of participants. **Cluster RCTs**, (sometimes referred to as C-RCTs), are trials in which clusters of individuals, rather than individuals themselves, are randomised to different arms of the trial. These clusters, or groups, have a natural affinity. For example participants within a cluster may attend the same school, medical practice or hospital, be from the same village, geographic region or family, or be patients of the same doctor. Cluster randomisation can be used in cases where it is not possible or not appropriate to randomise individual participants. For example, to evaluate the effects of interventions that involve organisational change, public education or media promotions, it may be appropriate to randomise clusters of participants at the level of the organization, at the level of the health professional, or at the community level.

Cluster RCTs can help to avoid **contamination** in trials.¹⁷ Some interventions can be difficult to deliver to one participant and not another within communities. For example, when studying the effects of health education interventions, it can be difficult to ensure that the intervention is received by participants in the intervention group, while at the same time ensuring that it is not received by those in the control group. Likewise, if the intervention relies on a health professional giving out particular information, it may be difficult for them to only give it to those in the intervention group. This may underestimate of the true effect of the intervention strategy, and can be overcome by randomising professionals, or groups of professionals, to different interventions. Cluster randomisation is also appropriate when the intervention is intended to be applied to a group, such as public education campaigns, or it is delivered to groups such as classrooms, families or community groups.

¹⁷ **Contamination** in clinical trials occurs where people in the control group inadvertently receive the intervention; or where people in the intervention group fail to receive the intervention.

Figure 3: The cluster randomised controlled trial

While cluster randomisation can help to avoid contamination, it also has limitations as a technique. People within any one cluster are more likely to be more alike than those across communities (for example, they may be more likely to respond in a similar manner). It can therefore be difficult to determine whether the results of cluster RCTs are due to baseline differences between clusters, rather than to the intervention itself.

Additionally, individuals within a cluster are not wholly independent of one another (as is the case in an individually randomised trial). A cluster RCT involving a certain number of participants is therefore less powerful¹⁸ than an individual RCT with the same number of participants. Statistical adjustment must be made when a cluster randomised design is used, as the participants within a given cluster are not truly independent of one another. Larger sample sizes are required to adjust for the clustering effect, and the analysis should be undertaken at the cluster level or using specific analytic techniques.

1.3.3.1 SUMMARY: KEY POINTS FOR AUTHORS TO CONSIDER ABOUT CLUSTER RCTs

- Cluster randomisation uses the randomisation of groups, rather than individuals, and is used to overcome problems associated with contamination.
- Specific statistical techniques are required to incorporate cluster RCTs into a meta-analysis.

¹⁸ **Power** in clinical trials is the probability that the study will be able to detect a statistically significant difference between groups if one actually exists. For example, if a clinical trial has a power of 0.80 (sometimes referred to as 80%), and the intervention does cause an effect, then if the trial was repeated 100 times, a statistically significant effect of the intervention would be found in 80 of the 100 times. Ideally, the higher the power of a study, the better it is. Power is affected by the sample size of the study (the number of participants in the trial), so that for the same effect size, the more participants in the study, the higher its power.

1.3.4 NON-RANDOMISED STUDIES

- See the [Cochrane Handbook for Systematic Reviews of Interventions](#), Chapter 13.
- See [Jadad \(1998\)](#), Chapter 1
- See [Deeks et al \(2003\)](#)

The distinction between randomised and non-randomised controlled studies is not a clear one, and several classification schemes have been developed. Differences in terminology between these different classification schemes can, however, lead to confusion and inconsistencies in assessing study design.

As a general distinction from RCTs, non-randomised studies can be said to include the following:

- Quasi-randomised controlled studies (studies that attempt but do not achieve true randomisation);
- Controlled before-and-after (CBA) studies;
- Interrupted time series (ITS) studies;
- Cohort studies;
- Case-control studies.

For our purposes, which is to discuss the types of study suitable for inclusion in Cochrane systematic reviews, we will use the term ‘non-randomised studies’ to include controlled *experimental* studies: those without truly randomised assignment (quasi-randomised controlled studies), CBA studies and ITS studies.

Cohort and case-control studies are also non-randomised studies but they are *observational* study designs. These are less relevant to decisions about which study designs eligible for inclusion in Cochrane systematic reviews. While these types of studies can provide valuable additional information to a review, particularly in relation to the Background and Discussion sections, their design is such that they are unable to assess the effects of interventions (See **Appendix D** for further discussion of descriptive and comparative study designs and associated design issues).

As mentioned earlier, techniques are now being developed to allow systematic reviews of observational studies to be prepared. At present however there is little consensus in this area. We therefore restrict the use of the term ‘non-randomised studies’ throughout this document to mean ‘non-randomised experimental studies’.

While non-randomised studies can provide valuable information about the effects of interventions when RCTs are not possible, ethical or appropriate, they are more prone to selection bias than RCTs. This is because they do not employ randomisation to allocate participants to intervention and control groups. As the methods of allocating participants may be less fair than randomisation, they may produce groups that are systematically and substantially different at baseline.

For example, one group may contain participants who are older, or sicker, less educated or less motivated to comply with treatment. This might happen, for example, because the investigators think that certain patients are particularly likely to benefit from receiving the intervention; or that they are too ill to risk not receiving the active treatment; or alternatively, that they are too ill to be involved in a trial. When such systematic differences occur between groups in a study, it is likely to distort the true effect of the intervention under examination – what ends up being measured is a combination of the differences that existed at baseline, together with any differences that the intervention has caused. In studies of non-randomised design, it is therefore especially important to assess the baseline characteristics of the study groups – those characteristics of groups before the intervention is given. It is important to consider whether the study groups are comparable; and to identify any systematic differences between groups at the outset that are likely to influence the effects of the intervention.

1.3.4.1 QUASI-RANDOMISED CONTROLLED STUDIES

Quasi-randomised studies are those studies where participants were allocated to either the intervention or control groups by investigators by using a random allocation sequence, but where the actual method of randomisation was inadequate and not really random. Such methods include allocation by alternation between groups, the use of birth dates, weekdays or case or medical record numbers.

These approaches are not considered to be truly random, as they do not give each of the participants in the study the same chance of being allocated to any of the study groups. They are also predictable.

The distinction between quasi-randomised studies and other non-randomised study designs is that other non-randomised designs do not attempt to randomly assign participants to treatment groups (whereas quasi-randomised studies have attempted to do so but the methods used fall short of being truly random).

1.3.4.1.1 SUMMARY: KEY POINTS FOR AUTHORS TO CONSIDER ABOUT QUASI-RCTS

- Quasi-RCTs attempt to allocate participants randomly to study groups; however, the methods that are used fall short of true randomisation.

- While quasi-RCTs attempt to randomly assign participants to groups, other types of non-randomised study design do not attempt randomisation but instead rely on other methods of allocating participants.
-

1.3.4.2 A NOTE ABOUT CONTROLLED CLINICAL TRIALS (CCTs)

Review authors should not use the term Controlled Clinical Trial (CCT) in their reviews except as part of a search strategy, as it is an indexing term for databases rather than a well-defined study design. Trials allocating participants to intervention and control groups using quasi-random methods should be described as quasi-randomised controlled trials rather than CCTs.

1.3.4.3 CONTROLLED BEFORE-AND-AFTER STUDIES (CBAs)

- See EPOC resources, available at <http://www.epoc.cochrane.org> or email amayhe2@uottawa.ca

A CBA study is a type of non-randomised study. Two groups are identified: the intervention group which receives the intervention and the control group which does not. The control group can be an appropriate site or activity or ward. The control group should be comparable with the intervention group on key characteristics, eg size, socio-demographic features, patient or professional type. The effect of an intervention is tested by collecting data both **before** and **after** the intervention is introduced, and the control and intervention groups are then compared (see **Figure 4**). The outcomes have to be measured at the same time for comparable time periods for both the before and after measurements, and for both groups. Any systematic differences between groups should be identified at the outset, as this may influence the effects of the intervention, or the outcomes over time.

These studies may be prone to bias, as there may be unidentified differences between the control and intervention groups that can influence or contribute to the measured results.

Some '**before-and-after studies**' do not contain a separate control group. Such studies rely on the participants acting as their own controls; that is, they are compared to themselves before and after the intervention (for example, their blood pressure might be measured prior to the intervention, then following the intervention, and the two measurements compared). However, if there is no control group, this does not allow for, or control for, the effects of time. Any changes or effects that are measured at the end of the study may have occurred anyway, without the intervention, and so cannot be dissociated from those effects due to the intervention itself. The inclusion of a separate control group in a before-and-after study avoids this problem, and means that the effect of the intervention

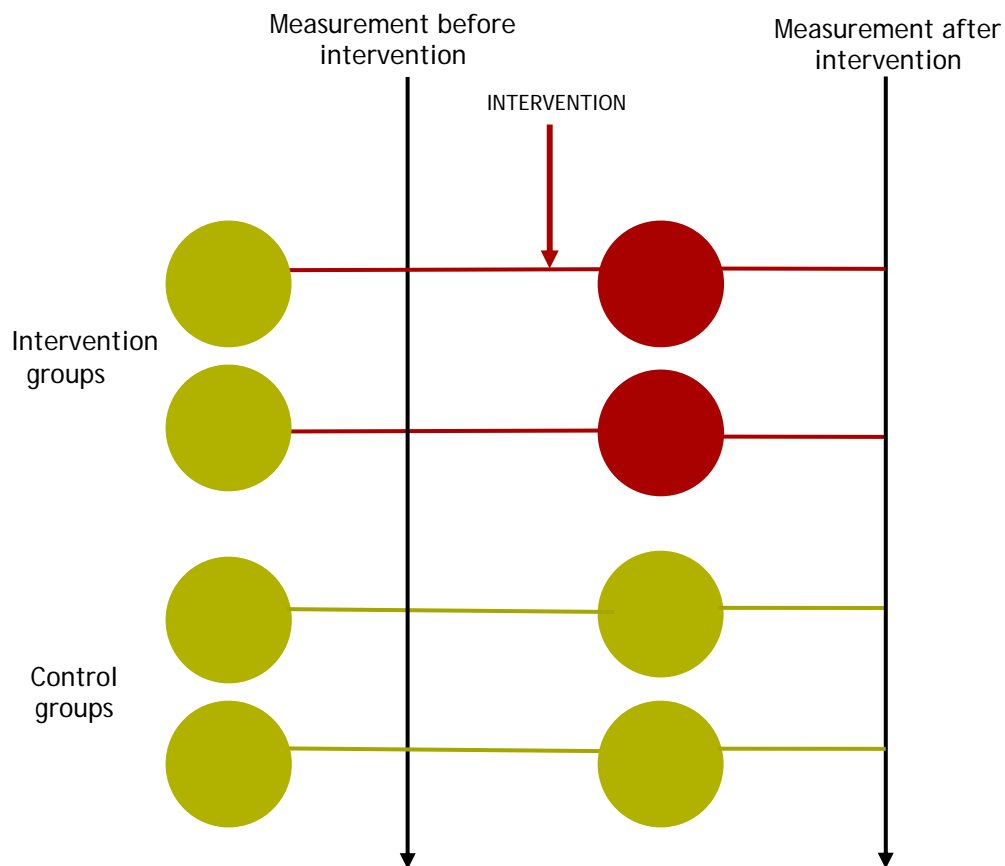
can be determined separately from those effects which occur over time anyway (that is, those that are not related to the intervention under investigation).

The EPOC Review Group specifies that, to be included in a systematic review, a CBA study must meet the following criteria.

- There must be at least two intervention sites and two control sites (note, this is a new criterion added in 2009).
- The timing of the periods for study for the control and intervention groups should be comparable (that is, the pre- and post- intervention periods of measurement for the control and intervention groups should be the same); and
- The intervention and control groups should be comparable on key characteristics.

CBA studies that do not meet these criteria should be excluded from a Cochrane review.

Figure 4: The controlled before-and-after study



1.3.4.4 INTERRUPTED TIME SERIES (ITS) DESIGNS

- See *EPOC Methods Paper: Including Interrupted Times Series (ITS) designs in an EPOC review* (available at <http://epoc.cochrane.org/> or email amayhe2@uottawa.ca)

In ITS studies, a single group of participants is measured repeatedly, both before and after the intervention; that is, the measurements are 'interrupted,' usually by the intervention being examined. These designs can therefore allow measurement of the effect of an intervention when randomisation or the identification of a control group are impractical or impossible, although a control group may also be included if it is possible to do so.

The EPOC Review Group specifies that, to be included in a systematic review, two criteria of the ITS study must be met. These are that:

- There must be a clearly defined point in time at which the intervention occurred, and this should be specified by the researchers.
- There should be collection of at least three data points before and three data points after the intervention was introduced.

ITS studies that do not meet these criteria should be excluded from a Cochrane review.

1.4 STUDY DESIGN AND RISK OF BIAS

- See the [Cochrane Handbook for Systematic Reviews of Interventions](#), chapter 8 and 13.5,

While randomisation and the process of properly concealing randomisation are important elements of study design, they do not represent the only important parts of an experimental research design.

'Experimental design involves the scheduling of observations, the choice of interventions and comparisons, the selection of measured control variables, and the manner of assigning units to interventions. One can have experiments where the questions, samples, interventions comparisons, and measures are chosen poorly or well. Random assignment deals with only the last of these issues, the manner of assignment.... Random assignment does not make causal inference infallible. It is still dependent on accepting assumptions eg., that random assignment resulted in initial comparability ... or systematic attrition did not occur, and so forth' (Cook and Campbell (1979), p. 405).

This description introduces the concept of **risk of bias** and the idea that there are various distinct but interrelated components of study quality that need to be assessed when performing a systematic review.

In the context of a systematic review, study quality is the extent to which the design and conduct of a study avoids **bias**. Bias is any systematic error that systematically distorts the results or inferences of a study, whether it is recognised or not.

Random assignment of participants to groups may, for example, produce basically equivalent groups at baseline. However, if there are uneven attrition rates from groups over the course of the study (that is, participants in the control and intervention groups are lost or drop out of the study at different rates), this may influence the result. For example, participants in the intervention group who are receiving a particular drug may experience more side-effects and so be more likely to drop out of the study than those in the control group who are receiving a placebo. This could underestimate the severity of the side-effects of the intervention. Alternatively, people assigned to an the intervention group where they receive a complete package of care upon discharge from hospital might be less likely to withdraw from the study because of dissatisfaction with the care they are receiving, compared with people receiving routine care. This however might lead to an underestimation of the effects of the intervention, as fewer of the people in the routine care group who are really dissatisfied with their care will be measured at the end of the study.

Likewise, other forms of bias can be introduced and so distort the study's results if outcomes are not systematically and objectively measured. These are only a few of the many possible factors that can affect the quality of a study, or the extent to which readers can rely on the findings of a study to

represent a real effect, rather than the influence of other factors. The Group's **Study Quality Guide** deals in more detail with different forms of bias that can influence the results of a study, and how to minimise the effects of these different sources of bias.

The issues that are important for study quality vary according to the study's design. While design and quality are different issues, they are also related and together serve to determine the **strength of the evidence** that can be drawn from a particular study.

As discussed earlier, RCTs are regarded as the best type of study with which to evaluate the effects of interventions; that is, they sit at the top of the 'evidence hierarchy'. There is an important distinction to be made, however, between the design and the quality of an RCT (or any other type of study design). While the RCT avoids many sources of bias, there can also be limitations in the actual design or execution of an RCT that affects its quality. RCTs, or any other study design, can therefore be of variable quality. It is possible to have both high-quality RCTs (that is, their design and execution are rigorous and so avoid potential sources of bias), as well as lower quality RCTs (where there are limitations in either or both of the design and execution of the study). Limitations in the design or execution of an RCT that affects its quality affects the strength of the evidence, or strength of the inferences, that can be drawn from it. These issues are discussed further in the guide on study quality.

It is important to appreciate that when including a study in a systematic review, **no matter its design**, the study needs to be **critically** and **systematically** evaluated. This is true whether the study is of randomised or non-randomised design. It is common for people to regard RCTs highly and to mistakenly assume that they do not need to be evaluated or critically appraised. However, an RCT, just like any other type of study design, can be well or poorly designed and executed.

Considering how well a particular study included in a systematic review has managed to avoid systematic errors is one important task for a review author. This involves considering the study quality (or how likely it is that the study results reflect what is true, or **internal validity**). Review authors must also consider the applicability of the study findings (whether the results can be generalised beyond the study to the population, the **external validity** of the study). These issues are covered in more detail in the **Study Quality Guide**.

It is important for review authors to note that the quality of a study is not the same as the quality of reporting of a study. Often, for various reasons, study authors will neglect to provide all relevant details about their trials for publication. For example, they may say that participants were 'randomly assigned' to groups, without providing details about how the random number sequence was generated. They may not specifically mention allocation concealment, blinding or numbers of participants followed over the course of the study. Just because the study authors did not specifically mention the details of their study does not mean that they did not perform the necessary steps – they

may have limited word space to describe their study in the publication; or they may not be aware that certain details are especially important for readers of their study and researchers conducting secondary analysis of their work.

Whatever the reason, it is rare to find studies, when assessing them for inclusion in a systematic review, that are so completely reported as to answer all questions. This can make it difficult for review authors to assess and to accurately appraise the studies included in their review.

It is for these reasons that review authors must attempt to contact the authors of all of the included studies, for additional information.

Additional information provided by study authors can have dramatic effects on a systematic review, markedly changing the quality rating of an included study, or altering the decision to include a study in a review. For example, a trial publication might report that participants were assigned randomly to groups. This study would therefore seem to be eligible for inclusion in a systematic review of RCTs. Contact with the authors of the study, however, might reveal that the randomisation process was not actually random – perhaps participants were assigned to groups by alternate allocation. In this case, the study would no longer be eligible for inclusion in a review of RCTs, as it is now rated as a quasi-randomised study, rather than a true RCT. Review authors must therefore attempt to contact all authors of studies included in their review to request these types of information.

1.5 QUANTITATIVE AND QUALITATIVE STUDY DESIGNS

- See the [Cochrane Handbook of Systematic Reviews of Interventions](#), chapter 20.

Qualitative and quantitative studies often differ markedly in their aims. While qualitative studies are commonly used for exploration or description, they may also be used to explain phenomena or to generate hypotheses. The role of experimental studies, in comparison, is usually to test such hypotheses¹⁹.

See **Appendix E** for further details on the specific aims and methods of qualitative research.

More and more, researchers are realising the benefits of combining qualitative and quantitative approaches to research. This can be achieved, for example, by using a qualitative study to inform the design of a quantitative study, or vice versa; or can involve the conduct of both qualitative and quantitative studies simultaneously in order to address a particular question. For further discussion on ways of integrating qualitative and quantitative data, see Dixon-Woods *et al.*, 2006; and Oliver *et al.*, 2005.

Because qualitative research uniquely focuses on meaning, context and depth, including such studies in systematic reviews raises different issues than those associated with the inclusion of quantitative studies. While some Cochrane reviews do now contain qualitative research, such research is not necessarily included in reviews the same way that data from RCTs is included. For instance, qualitative research can assist in developing the Background section for the review so that there is an in-depth understanding of the context for the intervention being investigated. Qualitative research can help to identify important outcomes that should be measured in an experimental study of an intervention. It can also help us to interpret the results of RCTs.

For many of the interventions relevant to the Cochrane Consumers and Communication Review Group, qualitative research can add valuable depth to the information obtained from experimental research designs, so aiding in the interpretation of results. It can also help set the scene for a systematic review by informing the context and by generating hypotheses for future research. While guidance on appraising qualitative research and routinely including it in systematic reviews is still in its infancy, review authors are encouraged to consider qualitative research and the additional information that this type of research can provide.

¹⁹ In fact, the purpose of experimental studies is to answer specific, discrete research questions – such as whether a particular intervention will have the hypothesised effect on a particular population of people.

Note however that while authors are encouraged to consider using qualitative research to help inform their systematic review, the Consumers and Communication Review Group does not at present accept systematic reviews of purely qualitative research.

1.5.1 SUMMARY: KEY POINTS FOR AUTHORS TO CONSIDER ABOUT USING QUALITATIVE RESEARCH IN COCHRANE REVIEWS

Ways that qualitative data can be used or included in systematic reviews can include the following:

- Providing a context in which to interpret the study's results.
 - Identification of outcomes that are important, especially patient-centred outcomes.
 - Providing in-depth description of the intervention and its important characteristics.
 - Hypothesis generation for future research.
-

APPENDICES

APPENDIX A – STUDY DESIGN TERMS

Except where indicated, taken from the [Cochrane Handbook for Systematic Reviews of Interventions Glossary](#).

CASE SERIES

An uncontrolled observational study involving an intervention and outcome for more than one person.

CASE STUDY (SYNONYMS: ANECDOTE, CASE HISTORY, SINGLE CASE REPORT)

An uncontrolled observational study involving an intervention and outcome for a single person (or other unit).

CASE-CONTROL STUDY (SYNONYMS: CASE REFERENT STUDY, RETROSPECTIVE STUDY)

A study that starts with identification of people with the disease or outcome of interest (cases) and a suitable control group without the disease or outcome. The relationship of an attribute (intervention, exposure or risk factor) to the outcome of interest is examined by comparing the frequency or level of the attribute in the cases and controls. For example, to determine whether thalidomide caused birth defects, a group of children with birth defects (cases) could be compared to a group of children without birth defects (controls). The groups would then be compared with respect to the proportion exposed to thalidomide through their mothers taking the tablets. Case-control studies are sometimes described as being retrospective as they are always performed looking back in time.

CLUSTER RANDOMISED TRIAL

A trial in which clusters of individuals (eg. clinics, families, geographical areas), rather than individuals themselves, are randomised to different treatments, or **arms** of the trial.

COHORT STUDY (SYNONYM: FOLLOW-UP, INCIDENCE, LONGITUDINAL, PROSPECTIVE STUDY)

An observational study in which a defined group of people (the cohort) is followed over time. The outcomes of people in subsets of this cohort are compared, to examine for example people who were exposed or not exposed (or exposed at different levels) to a particular intervention or other factor of interest. A cohort can be assembled in the present and followed into the future (this would be a prospective study or a "concurrent cohort study"), or the cohort could be identified from past records and followed from the time of those records to the present (this would be a retrospective study or a "historical cohort study"). Because random allocation is not used, matching or statistical adjustment at the analysis stage must be used to minimise the influence of factors other than the intervention or factor of interest.

CONTROLLED BEFORE AND AFTER STUDY (CBA)

A non-randomised study where a control population of similar characteristics and performance as the intervention group is identified. Data is collected from both the control and intervention groups before and after the intervention.

The EPOC Group states that, to be included in a review, a CBA study must meet these criteria:

- There must be at least two intervention sites and two control sites (note, this is a new criterion added in 2009).
- The timing of the periods for study for the control and intervention groups should be comparable.
- The intervention and control groups should be comparable on key characteristics.

Adapted from EPOC resources, available at <http://www.epoc.cochrane.org> or by emailing amayhe2@uottawa.ca

CONTROLLED (CLINICAL) TRIAL (CCT)

A trial that has a control group. Such trials are not necessarily randomised. Note that while not all controlled studies are randomised, all randomised studies are controlled. Note also that CCT is an indexing term used in Medline and CENTRAL. In CENTRAL it refers to trials using quasi-randomisation, or trials where double blinding was used but randomisation was not mentioned.

CROSS-SECTIONAL STUDY (SYNONYM: PREVALENCE STUDY)

A study that examines the relationship between diseases (or other health related characteristics) and other variables of interest as they exist in a defined population at one particular time. The temporal sequence of cause and effect cannot necessarily be determined in a cross-sectional study.

CROSS-OVER TRIAL

A type of clinical trial comparing two or more interventions in which the participants, upon completion of the course of one treatment are switched to another. For example, for a comparison of treatments A and B, half the participants are randomly allocated to receive them in the order A, B and half to receive them in the order B, A. A problem with this design is that the effects of the first treatment may carry over into the period when the second is given.

FACTORIAL DESIGN

Most trials only consider a single factor, where an intervention is compared with one or more alternatives, or a placebo. In a trial using a 2 x 2 factorial design, participants are allocated to one of four possible combinations. For example in a 2 x 2 factorial RCT of nicotine replacement and counselling, participants would be allocated to: nicotine replacement alone, counselling alone, both, or neither. In this way it is possible to test the independent effect of each intervention on smoking cessation and the combined effect of (interaction between) the two interventions.

INTERRUPTED TIME SERIES (ITS)

A study design that collects observations at multiple time points before and then after an intervention (ie. the intervention 'interrupts' the collection of data). This design aims to detect whether the effect of the intervention is significantly greater than the underlying trend occurring over time.

The EPOC Review Group specifies that, to be included in a systematic review, two criteria of the ITS study must be met:

- There must be a clearly defined point in time at which the intervention occurred, and this should be specified by the researchers.
- There should be collection of at least three data points before and three data points after the intervention was introduced.

Adapted from EPOC Methods Paper: Including Interrupted Times Series (ITS) designs in an EPOC review (available at <http://www.epoc.cochrane.org/> or email amayhe2@uottawa.ca).

N OF 1 RANDOMISED TRIAL

A randomised trial in an individual. N of 1 trials can be used in medical practice to determine the optimum treatment for an individual patient. There are many ways of conducting N of 1 randomised trials, one approach is:

1. A clinician and patient agree to test an intervention (the "experimental therapy") for its ability to improve or control the symptoms, signs, or other manifestations (the "treatment targets") of the patient's health problem.
2. The patient then undergoes "pairs" of treatment "periods" organized so that one period of each pair applies the experimental therapy and the other period applies an alternative intervention or placebo. The order of these two periods within each pair is randomized by a coin toss or other method that ensures that patient is equally likely to receive the experimental or control intervention during any period.
3. Whenever possible, both clinician and patient are blind to which intervention the patient is receiving.
4. The clinician monitors the treatment targets, often through a patient diary, to document the effect of the intervention currently being applied.
5. Pairs of treatment periods are replicated until the clinician and patient are convinced that the experimental therapy is effective, is harmful, or has no effect on the treatment targets. This usually requires 3 pairs.

OBSERVATIONAL STUDY (SYNONYM: NON-EXPERIMENTAL STUDY)

A study in which nature is allowed to take its course. Changes or differences in one characteristic (eg. whether or not people received the intervention of interest) are studied in relation to changes or differences in other(s) (eg. whether or not they died), without action by the investigator. There is a greater risk of selection bias than in experimental studies (randomised controlled trials).

PAIRED DESIGN

A study in which participants or groups of participants are matched (eg. based on prognostic factors) and one member of each pair is allocated to the experimental (intervention) group and the other to the control group.

PARALLEL GROUP TRIAL (SYNONYM: INDEPENDENT GROUP DESIGN)

A trial that compares two groups of people, one of which receives the intervention of interest and one of which is a control group. Some parallel trials have more than two comparison groups and some compare different interventions without including a no-intervention control group.

PREFERENCE TRIAL

A preference trial is a trial in which at least one group is included in which eligible participants are able to choose their preferred option from the several choices offered. Such trials therefore take into account the preferences of individuals eligible to participate, whether they actually participate in the trial or not. Such trials attempt to take into account the influence of holding a strong preference (for example, a strong preference for a particular intervention, a strong preference not to receive a placebo), compared with those individuals who do not hold a strong preference. There are at least three types of preference trials: those with Zelen's design, a comprehensive cohort design, or Wennberg's design (Jadad, 1998 p 23; Torgerson and Sibbald 1998).

PROSPECTIVE STUDY

In evaluations of the effects of healthcare interventions, a study in which people are divided into groups that are exposed or not exposed to the intervention(s) of interest before the outcomes have occurred. Randomised controlled trials are always prospective studies and case control studies never are. Concurrent cohort studies are prospective studies, whereas historical cohort studies are not (see cohort study), although in epidemiology a prospective study is sometimes used as a synonym for cohort study. See [Retrospective study](#).

QUASI RANDOMISED/ RANDOMIZED TRIAL

A trial using a quasi-random method of allocating participants to different forms of care. There is a greater risk of selection bias in quasi-random trials where allocation is not adequately concealed compared with randomised controlled trials with adequate allocation concealment.

RANDOMISED/ RANDOMIZED CONTROLLED TRIAL (RCT) (SYNONYM: RANDOMISED CLINICAL TRIAL)

An experiment in which investigators randomly allocate eligible people into intervention groups to receive or not to receive one or more interventions that are being compared. The results are assessed by comparing outcomes in the treatment and control groups. NOTE: when using randomised controlled trial as a search term (publication type) in MEDLINE, the US spelling (randomized) must be used.

RETROSPECTIVE STUDY

A study in which the outcomes have occurred to the participants before the study commenced. Case control studies are always retrospective, cohort studies sometimes are, randomised controlled trials never are. See [Prospective study](#).

SEQUENTIAL TRIAL

A trial in which the data are analysed after each participant's results become available, and the trial continues until a clear benefit is seen in one of the comparison groups, or it is unlikely that any difference will emerge. The main advantage of sequential trials is that they will be shorter than fixed length trials when there is a large difference in the effectiveness of the interventions being compared. Their use is restricted to conditions where the outcome of interest is known relatively quickly.

UNIT OF ALLOCATION

The unit that is assigned to the alternative interventions being investigated in a trial. Most commonly, the unit will be an individual person but, in some trials, people will be assigned in groups to one or other of the interventions. This is done to avoid contamination or for convenience and the units might be, for example, hospitals or communities. In other trials, different parts of a person (such as the left or right eye) might be assigned to receive different interventions. See unit of analysis error.

WASHOUT PERIOD

The stage in a cross-over trial when treatment is withdrawn before the second treatment is given. Washout periods are usually necessary because of the possibility that the intervention administered first can affect the outcome variable for some time after treatment ceases. A run-in period before a trial starts is sometimes called a washout period if treatments that participants were using before entering the trial are discontinued.

APPENDIX B – RESEARCH DESIGN AND METHODS TERMS

Except where indicated, taken directly from the *Glossary of Cochrane Collaboration and research terms*, available at <http://www.cochrane.org/resources/glossary.htm>

BIAS

A systematic error or deviation in results or inferences. In studies of the effects of healthcare, bias can arise from systematic differences in the groups that are compared (selection bias), the care that is provided, or exposure to other factors apart from the intervention of interest (performance bias), withdrawals or exclusions of people entered into the study (attrition bias) or how outcomes are assessed (detection bias). Bias does not necessarily carry an imputation of prejudice, such as the investigators' desire for particular results. This differs from conventional use of the word in which bias refers to a partisan point of view.

BIAS, SELECTION

1. In assessments of the validity of studies of healthcare interventions, selection bias refers to systematic differences between comparison groups in prognosis or responsiveness to treatment. Random allocation with adequate concealment of allocation protects against selection bias. Other means of selecting who receives the intervention of interest, particularly leaving it up to the providers and recipients of care, are more prone to bias because decisions about care can be related to prognosis and responsiveness to treatment.
2. Selection bias is sometimes used to describe a systematic error in reviews due to how studies are selected for inclusion. Publication bias is an example of this type of selection bias.
3. Selection bias, confusingly, is also sometimes used to describe a systematic difference in characteristics between those who are selected for study and those who are not. This affects the generalisability (external validity) of a study but not its (internal) validity.

BIAS , PERFORMANCE

Systematic differences between intervention groups in care provided apart from the intervention being evaluated. For example, if participants know they are in the control group, they may be more likely to use other forms of care. If care providers are aware of the group a particular participant is in, they might act differently. Blinding of study participants (both the recipients and providers of care) is used to protect against performance bias

BIAS , ATTRITION

Systematic differences between comparison groups in withdrawals or exclusions of participants from the results of a study. For example, participants may drop out of a study because of side effects of an intervention, and excluding these participants from the analysis could result in an overestimate of the effectiveness of the intervention, especially when the proportion dropping out varies by treatment group.

BIAS , DETECTION

Systematic difference between comparison groups in how outcomes are ascertained, diagnosed or verified (also called ascertainment bias).

BLINDING (SYNONYM: MASKING)

Keeping group assignment (eg. to treatment or control) secret from the study participants or investigators. Blinding is used to protect against the possibility that knowledge of assignment may affect patient response to treatment, provider behaviours (performance bias) or outcome assessment (detection bias). Blinding is not always practical (eg. when comparing surgery to drug treatment). The importance of blinding depends on how objective the outcome measure is; blinding is more important for less objective outcome measures such as pain or quality of life. See also single blind, double blind and triple blind.

SINGLE BLIND (SYNONYM: SINGLE MASKED)

The investigator is aware of the treatment/intervention the participant is getting, but the participant is unaware. See also blinding, double blind, triple blind.

DOUBLE BLIND (SYNONYM: DOUBLE MASKED)

Neither the participants in a trial nor the investigators (outcome assessors) are aware of which intervention the participants are given. The purpose of blinding the participants (recipients and providers of care) is to prevent performance bias. The purpose of blinding the investigators (outcome assessors, who might also be the care providers) is to protect against detection bias. See also blinding, single blind, triple blind, concealment of allocation.

TRIPLE BLIND (SYNONYM: TRIPLE MASKED)

An expression that is sometimes used to indicate that knowledge of which study participants are in which comparison group is kept secret from the statistician doing the analysis as well as from the study participants and investigators (outcome assessors). See also blinding, single blind, double blind.

CONCEALMENT OF ALLOCATION

The process used to prevent foreknowledge of group assignment in a randomised controlled trial, which should be seen as distinct from blinding. The allocation process should be impervious to any influence by the individual making the allocation by having the randomisation process administered by someone who is not responsible for recruiting participants; for example, a hospital pharmacy, or a central office. Using methods of assignment such as date of birth and case record numbers (see quasi random allocation) are open to manipulation. Adequate methods of allocation concealment include: centralised randomisation schemes; randomisation schemes controlled by a pharmacy; numbered or coded containers in which capsules from identical-looking, numbered bottles are administered sequentially; on-site computer systems, where allocations are in a locked unreadable file; and sequentially numbered opaque, sealed envelopes. Also refer to single, double and triple blind above.

CONFOUNDING

A situation in which a measure of the effect of an intervention or exposure is distorted because of the association of exposure with other factor(s) that influence the outcome under investigation.

CONTAMINATION

In clinical trials, the inadvertent application of the intervention being evaluated to people in the control group or inadvertent failure to apply the intervention to people assigned to the intervention group.

CONTROL

1. In clinical trials comparing two or more interventions, a control is a person in the comparison group that receives a placebo, no intervention, usual care or another form of care.
2. In case-control studies a control is a person in the comparison group without the disease or outcome of interest.
3. In statistics, control means to adjust for or take into account extraneous influences or observations.
4. Control can also mean programs aimed at reducing or eliminating the disease when applied to communicable (infectious) diseases.

HISTORICAL CONTROL

Person or group for whom data were collected earlier than for the group being studied. Because of changes over time in risks, prognosis, healthcare, etc. there is a large risk of bias (in studies that use historical controls) due to systematic differences between the comparison groups.

METHODOLOGICAL QUALITY (SYNONYMS: VALIDITY, INTERNAL VALIDITY)

The extent to which the design and conduct of a study are likely to have prevented systematic errors (bias). Variation in quality can explain variation in the results of studies included in a systematic review. More rigorously designed (better 'quality') trials are more likely to yield results that are closer to the 'truth'. See also external validity, validity.

OPEN CLINICAL TRIAL

There are at least three possible meanings for this term:

1. A clinical trial in which the investigator and participant are aware which intervention is being used for which participant (ie. not double blind). Random allocation may or may not be used in such trials.
2. A clinical trial in which the investigator decides which intervention is to be used (non-random allocation). This is sometimes called an open label design (but some trials which are said to be "open label", are randomised).
3. A clinical trial which uses an open sequential design.

OPEN LABEL DESIGN

A study design in which the investigator is aware which intervention is being given to which participant (ie. not double blind). Some studies with an open label design are randomised trials, but some do not include a comparison group and, therefore, cannot be randomised. See also open clinical trial.

PLACEBO

An inactive substance or procedure administered to a patient, usually to compare its effects with those of a real drug or other intervention, but sometimes for the psychological benefit to the patient through a belief that s/he is receiving treatment. Placebos are used in clinical trials to blind people to their treatment allocation. Placebos should be indistinguishable from the active intervention to ensure adequate blinding.

QUASI-RANDOM ALLOCATION

A method of allocating participants to different forms of care that is not truly random; for example, allocation by date of birth, day of the week, medical record number, month of the year, or the order in which participants are included in the study (eg. alternation).

RANDOM ALLOCATION

A method that uses the play of chance to assign participants to comparison groups in a trial, eg. by using a random numbers table or a computer-generated random sequence. Random allocation implies that each individual or unit being entered into a trial has the same chance of receiving each of the possible interventions. It also implies that the probability that an individual will receive a particular intervention is independent of the probability that any other individual will receive the same intervention. See also concealment of allocation, quasi-random allocation, and randomisation.

RANDOMISATION (RANDOMIZATION)

Method used to generate a random allocation sequence, such as using tables of random numbers or computer-generated random sequences. The method of randomisation should be distinguished from concealment of allocation because there is a risk of selection bias despite the use of randomisation, if the allocation concealment is inadequate. For instance, a list of random numbers may be used to randomise participants, but if the list is open to the individuals responsible for recruiting and allocating participants, those individuals can influence the allocation process, either knowingly or unknowingly. (See below for definitions of *stratified*, *block* and *weighted* randomisation)

RUN-IN PERIOD

A period before a trial is commenced when no treatment is given. The data from this stage of a trial are only occasionally of value but can serve a valuable role in screening out ineligible or non-compliant participants, in ensuring that participants are in a stable condition, and in providing baseline observations. A run-in period is sometimes called a washout period if treatments that participants were using before entering the trial are discontinued.

STRATIFIED RANDOMISATION

In any randomised trial it is desirable that the comparison groups should be as similar as possible in regard to participant characteristics that might influence their response to the intervention. Stratified randomisation is used to ensure that equal numbers of participants with a characteristic, thought to affect prognosis or response to the intervention will be allocated to each comparison group. For example, in a trial of women with breast cancer, it may be important to have similar numbers of pre-menopausal and post-menopausal women in each comparison group. Stratified randomisation could be used to allocate equal numbers of pre- and post-menopausal women to each treatment group. Stratified randomisation is performed either by performing separate randomisation (often using random permuted blocks) for each strata, or by using minimisation method of randomisation that ensures that, at any point in a trial, roughly equal numbers of participants have been allocated to all the comparison groups. Permuted blocks should be used in trials using stratified randomisation. (Also called **block randomisation**).

Stratified randomisation is used to keep the characteristics such as age and weight as consistent as possible between the study groups (Jadad 1998 p.6). To achieve this, investigators must first identify factors (also known as **strata**) that are known to be related to the outcome of the study. Once these factors are identified, the next step is to produce a separate block randomisation scheme for each factor to ensure that the groups are balanced within each strata (Last 1995 p.162).

BLOCK (RESTRICTED) RANDOMISATION

In any given study, despite investigators' attempts to generate random sequences of allocation, the number and characteristics of the participants allocated to each of the study groups are likely to be inconsistent. To minimise these differences, investigators can use some strategies known as restricted (or block) randomisation or stratified randomisation (Jadad 1998 p.5). **Restricted randomisation** ensures that the number of participants in all the study groups are as close as possible. By creating 'blocks' of sequences, the researcher can ensure that the same number of participants are allocated to the study groups within each block. For example, in a study with three groups (1, 2, and 3), the investigators can create six blocks: 123, 132, 213, 231, 312, 321. If they use a die to generate the sequences, then they can decide how each of the six numbers of the die will correspond to each of the blocks. Each of the sequences will determine the allocation of three participants at the same time, not only one at a time. For example, if it is decided that 1 corresponds to the sequence 123, then the three participants that enter the study after a die has shown 1 will be allocated in that order: the first participant to group 1, the second to group 2, and the third to group 3 (Jadad 1998 p.5-6).

WEIGHTED (UNEQUAL) RANDOMISATION

In certain studies (such as those with a higher probability of producing unexpected adverse events), investigators may not desire the same number of participants in each of the study groups and can

decide to allocate unequal numbers to each group, while preserving the homogeneity of the distribution of the characteristics of the participants across the study groups. This is called **weighted** or **unequal** randomisation. This type of randomisation tends to be used by investigators who wish to expose fewer participants to the experimental group because of lack of knowledge of adverse effects (Jadad 1998 p.6).

APPENDIX C – ADDITIONAL SOURCES OF INFORMATION

EVIDENCE HIERARCHIES

Petticrew M and Roberts H. Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health* 2003; 57(7): 527-9.

- This paper discusses a framework for the use of evidence in public health decision making, by accounting for the strengths and weaknesses of different research design approaches. While the framework includes a role for RCTs, it also emphasises the importance of matching specific questions with different research types. For example, while RCTs are considered best for answering questions of effectiveness, safety and cost-effectiveness; qualitative studies may be best for other questions such as those of salience (ie. does it matter?), appropriateness and satisfaction.

INCONSISTENCY: HETEROGENEITY AND HOMOGENEITY OF EVIDENCE

Higgins JPT, Thompson SG, Deeks JJ and Altman DG. Measuring inconsistency in meta-analysis. *BMJ* 2003; 327: 557-60.

- This article discusses the concept of consistency of results in meta-analysis, the desirability of consistent results across studies; and the need for tests of heterogeneity in order to assess how generalisable the results of a meta-analysis are.
- Specifically, this paper discusses the merits of the I^2 test for heterogeneity among the results of studies (which is included in Cochrane systematic reviews with meta-analysis); and compares this test of consistency with other measures.

INTENTION-TO-TREAT (ITT) ANALYSIS

Heritier SR, Gebski VJ and Keeck AC. Inclusion of patients in clinical trials analysis: the intention-to-treat principle. *MJA*, 2003; 179: 438-40.

- This article provides a definition and examples of ITT analysis. It examines some of the advantages and the disadvantages of ITT analysis. This also includes examples where ITT analysis is not a highly desirable option for analysis.
- Includes worked examples of ITT as well as other forms of analysis to deal with participant loss. These examples also involve comparison of the estimates of effects produced by adopting different approaches to analysis when dealing with participant loss.

Also refer to the studies listed under *Attrition bias*.

META-ANALYSIS

Egger M, Smith GD, Phillips AN. Meta-analysis: Principles and procedures. *BMJ* 1997; 315: 1533-7.

- This study examines the steps of conducting a systematic review, emphasising that the review must be properly planned prior to beginning the review. It describes the characteristics of well-designed

meta-analyses, as well as common pitfalls in poorly designed meta-analyses. In conjunction with this discussion, practical steps are also described for the process of conducting a meta-analysis.

Lewis S and Clarke M. Forest plots: trying to see the wood for the trees. *BMJ* 2001; 322: 1479-80.

- This paper provides a brief outline of meta-analysis and forest plots. It explains what each component of a forest plot represents, and what can be determined from them as part of a meta-analysis.

QUALITY ASSESSMENT

Jüni P, Altman DG and Egger M. Assessing the quality of controlled clinical trials. *BMJ* 2001; 323: 42-6.

- This paper discusses the various different elements of quality that are necessary to assess for studies included in systematic reviews. It provides a detailed description of different sources of bias in studies of healthcare interventions; as well as a description of the ways in which study quality can be incorporated into meta-analysis and systematic reviews.

Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, *et al.* (2003) Evaluating non-randomised intervention studies. *Health Technology Assessment* 7(27).

<http://www.nchta.org/execsumm/summ727.htm>

- This paper presents an empirical comparison of randomised and non-randomised studies. It concludes that results from these different study types are sometimes, but not always, different; and discusses cases where non-randomised studies may produce seriously misleading results.
- This study also evaluates quality assessment tools for non-randomised studies and identifies key components that should be included in any quality assessment.

Schultz KF, Chalmers I, Hayes RJ and Altman DG. Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273(5): 408-12.

- This review assessed the relationships between different aspects of trial quality and trial results. It reports empirical evidence that several elements of study quality (allocation concealment and blinding), if not adequately performed, tend to overestimate the estimates of effect in trials.

***Systematic reviews: CRD's guidance for undertaking reviews in health care.* 3rd edition. Centre for Reviews and Dissemination, University of York, January 2009**

(http://www.york.ac.uk/inst/crd/systematic_reviews_book.htm); see in particular section 1.3.4 on study quality assessment.

- This guide gives an overview of study quality issues, bias and why it is necessary to assess the quality of studies included in systematic review. It also provides ways of approaching quality assessment of studies, with discussion in terms of answering different types of research questions.

QUALITATIVE RESEARCH

Dixon-Woods M, Agarwal S, Jones D, Young B and Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy* 2005; 10(1): 45-53b.

- This paper reviews and discusses a range of approaches to qualitative methods and the possibilities and barriers to incorporating qualitative and quantitative evidence within systematic reviews.

Dixon-Woods M, Bonas S, Booth A, Jones DR, Sutton A, Shaw RL, Smith JA, and Young B. How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research* 2006; 6(1): 27-44.

- This paper describes the process of attempting to include qualitative research in a systematic review together with quantitative research, including the difficulties faced in trying to do so.
- It also discusses a range of problems in current systematic review approaches which may render current methods incompatible with the approaches of qualitative research.

Oliver S, Harden A, Rees R, Shepherd J, Brunton G, Garcia J and Oakley A. An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation* 2005; 11(4): 428-46.

- This paper reviews a method used to include and evaluate both qualitative and quantitative data within a systematic review. It also discusses some of the rationale for attempting to do so and some of the possible benefits of including a broader range of evidence in systematic reviews.

RANDOMISED CONTROLLED TRIALS

Jadad A. (1998) 'Randomised controlled trials: A user's guide.' *BMJ Books*, London.

- This book provides a detailed overview of RCTs, from basic definitions, methods of randomisation and different types of RCTs to the assessment of RCT quality, sources of bias and reporting of individual trials. It may be especially useful for people without formal epidemiological training.

SYSTEMATIC REVIEWS

Mulrow CD. Systematic reviews: rationale for systematic reviews. *BMJ* 1994; 309: 597-9.

- This paper discusses the advantages of systematic reviews, including meta-analysis. It describes how systematic reviews can effectively represent and integrate large quantities of information; that they increase power; that both the generalisability and the consistency of findings can be assessed effectively; and that a systematic review can also allow explanation of inconsistencies in the data.

See also Jadad A. (1998) 'Randomised controlled trials: A user's guide.' *BMJ Books*, London, particularly Chapter 6.

APPENDIX D – ADDITIONAL INFORMATION ABOUT STUDY DESIGN

D.1 DESCRIPTIVE AND COMPARATIVE STUDY DESIGNS

There are many different types of study designs, and each type is appropriate for addressing different types of research question. There are also many different ways of classifying study designs. The following aims to give an outline of the different types of studies that a review author might encounter when conducting a systematic review on interventions to communicate with or involve consumers in health care.

Study designs can be classified, in broad terms, as **descriptive** or **comparative**.

D.1.1 DESCRIPTIVE STUDY DESIGNS

Descriptive studies record observations (such as opinions, behaviours, clinical measurements, and so on) in a defined population. These study designs can therefore look for associations (for example, the relationship of a disease with a particular characteristic in a certain group of people). They do **not** aim to examine causal relationships (whether a change in one variable, the intervention, caused changes in another variable, the outcome).

Descriptive studies can show correlations (association) between different variables and can also suggest hypotheses to be investigated further by appropriate comparative studies. An opinion survey is a common example of descriptive study: all the participants are questioned (sampled) at a single point in time, and the outcomes are self-reported by participants. Comparisons can be made within this dataset (for example, a comparison of men's opinions versus women's opinions). However in such a design nobody is exposed to any intervention.

These are several types of descriptive study designs, including the following:

- **Case series:** is an observational study of a series of cases, for example, describing the clinical course of a particular disease or condition.
- **Cross sectional survey:** is a study that examines the relationship between diseases (or other health-related factors) and other variables²⁰ of interest in a particular population at a single point in time. These studies therefore examine the associations between variables. A cross-sectional survey might, for example, examine the level of education and the risk of dementia in people aged over 70 to see whether an association between education and dementia exists. As these studies focus on describing a particular illness, condition or factor at *one* particular point in time, (usually) for a defined population group, they can **not** determine whether changes in one variable caused the other - only whether the two variables are associated. Cross-sectional surveys present what is essentially a snap-shot at a particular point in time, rather than a progression over time, but this can be useful for generating hypotheses to be tested using other types of research design.

²⁰ A **variable** is any characteristic, effect or factor that can take different values.

- **Correlational study:** is a study examining the association between two different variables between groups. For example, by comparing cigarette smoking and lung cancer rates in different countries, it might be possible to suggest an association between the two. However, the main weakness of this type of study is that it does not demonstrate that those people who developed lung cancer actually smoked.

Descriptive studies provide information about **associations** between different variables. However, it is important to remember that descriptive study designs do not aim to examine causal relationships.

D.1.2 COMPARATIVE STUDY DESIGNS

Study designs aiming to examine causal relationships, change or efficacy²¹ are based on **comparisons**. This may involve comparing the same participants before and after an intervention; or comparing two or more groups that have received different treatments. Although the distinction between descriptive studies and comparative studies appears very clear, in reports of other people's research it is not always clear what has been done or which design category the study falls into. For example, it can be difficult to decide whether a particular study is observational with different groups, or a controlled trial without randomisation.

Comparative study designs can be either observational or experimental.

D.1.2.1 OBSERVATIONAL STUDIES

In observational studies, the groups to be compared are made up of participants whose allocation and exposure are not controlled or manipulated by investigators. Sometimes it can be useful to think of observational designs as studies where nature is allowed to take its course; that is, where the investigator does *not* actively manipulate or control what happens to the groups that are being studied. In these studies, variations that naturally exist in terms of interventions or exposure to a particular factor are studied for their effects on health outcomes. It is also worth noting that observational studies are not confined only to those situations where nature takes its course. They can, for example, be about 'usual practice;' for example, comparing outcomes among pupils in two different schools.

There are different types of observational study, but a common example of an observational design is the **case-control study**. The case-control study starts by identifying people with the outcome of interest (the cases) and those people from the same population without the outcome (the controls). Investigators can then compare the frequency of the exposure (or intervention or risk factor) among those who have the outcome of interest and those who do not. This can help to determine whether a relationship exists between these two variables²².

²¹ **Efficacy** refers to the extent to which a particular intervention produces the desired effect under ideal conditions.

²² Case-control studies can be particularly useful when the outcome of interest is rare and it is possible to reliably measure the past exposure to the factor of interest.

D.1.2.2 EXPERIMENTAL STUDIES

A study is considered to be **experimental** when the investigator controls some of the conditions of the study. In experimental designs, commonly known as **trials**, investigators test the effect of different 'interventions'²³. In such studies, the intervention is introduced to one group of participants, and compared with another group who does not receive the intervention (the control group)²⁴. The investigator therefore actively allocates participants to the groups to be compared.

In trials comparing at least one intervention, people in the **control** comparison group can receive usual care, no intervention, placebo or another form of care, depending on the intervention. For example, in an experimental trial investigating the effects of a drug, the control group may receive a placebo pill to control for the effects of the active drug (that is, so that those involved in the study do not know who has received the active pills and who has received the placebo pills). Alternatively, a trial investigating the effects of psychological counselling (intervention), might allocate the control group to 'usual care.' When considering comparative studies, it is therefore important to consider what sort of control group has been used and whether it is appropriate for the intervention of interest

RANDOMISED CONTROLLED TRIALS (RCTs)

In essence, different types of experimental study design are determined by the method used to **assign** or **allocate** participants to the different groups to be studied. In **randomised controlled trials (RCTs)** the investigator randomly assigns people to groups that will receive (intervention group) or not receive (control group) one or more interventions. The outcomes measured are then compared between the groups. (*Section 1.3.2* of this document deals with randomisation methods and RCTs in more detail).

NON RANDOMISED STUDIES

In other experimental study designs investigating the effects of interventions, investigators also assign participants to different groups. However, in these cases the method used to assign participants to groups is *not* a random method²⁵.

There are several different types of non-randomised studies. These include the following:

²³ An **intervention** in healthcare studies is any treatment, procedure or program that can potentially produce a change to health and health outcomes.

²⁴ In trials comparing at least one intervention, people in the **control** comparison group can receive usual care, no intervention, placebo or another form of care, depending on what the intervention is. For example, in an experimental trial investigating the effects of a particular drug, the control group may receive a placebo pill to take into account (control for) the intervention (the active pill). In comparison, a trial investigating the effects of psychological counselling (intervention), might use a control group that receives 'usual care.' When considering comparative studies, it is important to consider what sort of control group has been used and whether it is appropriate for the particular intervention of interest.

²⁵ Note that there is no clear line between randomised and non-randomised studies, and that many different classification schemes have been developed to attempt to describe study designs ore clearly. The guidance provided in these documents is consistent with the advice provided by The Cochrane Collaboration. More detail on specific methods of randomisation are given in *Section 1.3.2*.

- **Quasi-randomised studies** that attempt to randomly allocate participants, but the method of randomisation is not *truly* random (these studies are also known as **quasi-randomised studies**);
- **Cohort** studies;
- **Controlled before-and-after (CBA)** studies;
- **Interrupted-time-series (ITS)** studies; and
- **Case-control** studies.

The major difference between the different types of experimental study designs is therefore the **way in which participants are allocated to groups** in the study. While this may seem like a small difference, the method of allocation has important implications for the study, its strengths and weaknesses overall, and ultimately the potential for bias in the findings.

The main issue with allocating people to the different groups in an experimental study is the **potential for bias**. The role of randomisation is to minimise the differences between people in the intervention and control groups. If the allocation is truly random, and the sample size is large enough, the characteristics of both groups should be similar – by chance alone. This is true of the factors that we know about (such as age, gender etc), as well as those that we haven't considered or cannot measure (such as attitude to medication).

In The Cochrane Collaboration, and more widely in the area of evidence-based medicine, RCTs are considered to represent 'true' experimental designs; that is, the most rigorous study design to evaluate the effects of interventions. Increasingly, the value and role of non-randomised studies and qualitative research designs, particularly as adjuncts to the inclusion of RCTs in systematic reviews, are being recognised as providing important contributions to reviews.

D.1.3 SUMMARY TABLE OF STUDY DESIGNS

DESCRIPTIVE		<ul style="list-style-type: none"> • Case series • Cross sectional survey • Correlational
COMPARATIVE	Observational	<ul style="list-style-type: none"> • Case-control • Cohort
	Experimental	<ul style="list-style-type: none"> • Randomised controlled trials (including cluster RCTs) • Non-randomised studies: <ul style="list-style-type: none"> ○ Quasi randomised trials ○ Controlled before and after studies ○ Interrupted Time series

D.1.4 A NOTE ON TIMING AND STUDY DESIGN

It is important to determine whether a study was conducted forward in time (**prospectively**), or backwards (**retrospectively**) in time.

- **Prospective studies** are those where people are allocated to groups that are exposed or receive an intervention *before* the outcomes are measured. Prospective studies can be either descriptive or comparative. For example, a prospective cohort study might assemble a group of participants (cohort) and follow them forward in time (into the future). This can happen for a defined time period, or until a specific event occurs (such as death, illness onset, clinical outcomes).
- **Retrospective studies**, in comparison, are studies in which the outcomes have *already* happened to the participants, before the study began. A retrospective study can be either descriptive or comparative. For example, a retrospective cohort study might identify a group of participants from past records (eg patient records or clinical data) and follow them from the time of the records until the present to measure or compare outcomes between subsets of the cohort.
- **Cross-sectional studies (surveys)** are studies where data are recorded at one particular point in time (during a single day, week or month) for all participants; they can also be either descriptive or comparative. Outcomes that are measured at this one time point can then be compared between different subgroups of participants, for example, between men and women, young and old, sick and well.

APPENDIX E – QUALITATIVE RESEARCH AND STUDY DESIGN

- See chapters 20 and 21 of the [Cochrane Handbook](#).

Qualitative research designs, like quantitative designs, can be used to explore, describe or explain particular events or phenomena. While qualitative methods are commonly used for exploration or description, they are also used extensively to explain phenomena. Numerous analytic techniques are now available for qualitative researchers (see, for example, Miles and Huberman, 1994 and Dixon-Woods *et al.*, 2005; 2006).

Qualitative and quantitative studies often differ markedly in their aims. Qualitative studies may, for example, produce hypotheses in order to explain phenomena. The role of experimental studies, in comparison, is usually to test such hypotheses²⁶. As the aims of qualitative and quantitative studies differ, the methods that they employ are also very different. However, it is a common misconception that qualitative and quantitative approaches are mutually exclusive. Some researchers would be surprised to realise that they routinely use approaches from the paradigm of the 'other' camp: for example, when incorporating open-ended questions into a fixed-choice self-completion questionnaire (for the quantitative researcher); or systematically collecting quantitative information (such as age or length of an experience) during interviews (for the qualitative researcher). Feinstein's *Clinimetrics* is dedicated to the use of qualitative methods for the design of clinical indexes, rating scales and other expressions that are used to describe or measure clinical phenomena in medicine (Feinstein, 1987).

QUALITATIVE RESEARCH DESIGNS

Qualitative research itself can be defined in many ways. It has, for example, been defined as research that does not use numbers (Seale, 1999). While this can be true, many qualitative studies do rely on numbers (such as simple counts, cross-tabulations, qualitative comparisons analysis) to report their findings. Including numbers in qualitative research is one way of avoiding the presentation of data as mere anecdotes. Qualitative methods, therefore, are not simply studies that do not use numbers. Nor are they necessarily restricted to the use of descriptive designs alone.

What, then, is so different or specific about qualitative research designs?

- The reasoning implicit in qualitative work moves from observation to hypothesis (induction) (that is, the hypotheses are usually developed during or after the research is completed). This contrasts with the deductive approach of quantitative research which aims to verify or falsify *a priori* hypotheses.

²⁶ In fact, the purpose of experimental studies is to answer specific, discrete research questions – such as whether a particular intervention will have the hypothesised effect on a particular population of people.

- In qualitative research, three data collection methods are commonly used:
 - Observation (to gain direct experience of the phenomena studied);
 - Interviews (to elicit others' perspectives); and
 - Document analysis/ examination (looking for regular patterns in texts, conversations, etc).
- Qualitative research develops explanations that fit the facts, rather than using predetermined operational definitions such as those often used in quantitative research. Whereas quantitative research starts with a research question or hypothesis that precedes data collection, qualitative researchers are encouraged **not** to separate the design, data collection and analysis stages of research. Instead, qualitative researchers are encouraged to move backwards and forwards between the raw data and the process of conceptualisation.
- Quantitative methods aim for reliability; that is, for consistency on retesting. In comparison, qualitative methods aim to uncover how people behave and what people mean when describing their own experiences, attitudes and behaviours.

QUALITATIVE RESEARCH METHODS

Examples of qualitative research include the following approaches (see also Dixon-Woods *et al*, 2005; 2006 for further details):

- **Observation:** Involves systematic observation and later interpretation of social behaviours within social settings. Note that these settings are not controlled by the researcher.
- **Participant observation:** Involves systematic observation where the researcher both observes and participates in the social setting of interest.
- **In-depth interviews:** Involves face-to-face discussion (structured or semi-structured) between the researchers and the participants. Generally they aim to explore in detail the participant's experiences and attitudes regarding a particular occurrence.
- **Focus groups:** Involves groups of participants (usually 12 to 15 participants) who are brought together to participate in guided discussion about a specific topic. The emphasis is on exploring, rather than explaining, particular topics.
- **Content analysis:** This analysis involves systematically analysing text via identified themes, and then coding, classifying and developing categories for analysis.
- **Discourse analysis:** Discourse analysis involves analysing the language itself and the context in which it is used. It is not, strictly speaking, analysis of grammar, but rather focuses on the way language is used to represent understandings of the world.

Qualitative research is uniquely focussed on meaning, context and depth. This raises different issues when including such studies in systematic reviews than those associated with the inclusion of quantitative studies. The two approaches have fundamentally different aims and strengths, but they are not mutually exclusive. Both approaches to research can be valuable and can be used effectively to complement one other (see Dixon-Woods *et al.*, 2006; and Oliver *et al*, 2005).

REFERENCES

- Cochrane Collaboration Open Learning Material for Reviewers, Version 1.1, November 2002. Eds, P Alderson and S Green, www.cochrane-net.org/openlearning/index.htm
- Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *The data collection checklist*. <http://www.epoc.cochrane.org/Files/Website%20files/Documents/Reviewer%20Resources/datacollectionchecklist.pdf>
- Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *The data abstraction form*. <http://www.epoc.cochrane.org/Files/Website/Reviewer%20Resources/Data%20Abstraction%20Form%20-%20EPOC%20-%202007-Feb-27.doc>
- Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *EPOC Methods Paper: Including Interrupted Times Series (ITS) designs in an EPOC review*. <http://www.epoc.cochrane.org/Files/Website/Reviewer%20Resources/inttime.pdf>
- Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *EPOC Methods Paper: Issues Related to Baseline Measures of Performance* <http://www.epoc.cochrane.org/Files/Website/Reviewer%20Resources/baseline.pdf>
- Cook TD and Campbell DT. (1979) 'Quasi-experimentation: design and analysis issues for field settings.' Rand McNally College Pub. Co., Chicago: xii, p. 405.
- Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, *et al.* (2003) Evaluating non-randomised intervention studies. *Health Technol Assess* 7(27). <http://www.ncchta.org/execsumm/summ727.htm>
- Dixon-Woods M, Agarwal S, Jones D, Young B and Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy* 2005; 10(1): 45 - 53b.
- Dixon-Woods M, Bonas S, Booth A, Jones DR, Sutton A, Shaw RL, Smith JA, and Young B. How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research* 2006; 6(1): 27- 44.
- Feinstein AR. (1987) 'Clinimetrics.' Yale University Press, New Haven: xi, p. 272.
- Harris IB. (2002) 'Qualitative methods.' In: Norman G, Van Der Vleuten C, Newble D, eds. *International handbook of research in medical education*. 1st ed. Kluwer Academic Dordrecht: p. 45-95.
- Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions 5.0.1* [updated September 2008]. www.cochrane-handbook.org
- Jadad A. (1998) 'Randomised controlled trials: A user's guide.' BMJ Books, London.
- Miles MB and Huberman AM. (1994) 'Qualitative data analysis: an expanded sourcebook.' 2nd ed, Sage Publications, Thousand Oaks, California: xiv, p. 338.
- Oliver S, Harden A, Rees R, Shepherd J, Brunton G, Garcia J and Oakley A. An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation* 2005; 11(4): 428-46.
- Seale C. (1999) 'The quality of qualitative research.' Sage Publications, London; Thousand Oaks, California: x, p. 214.
- Torgerson DJ and Sibbald B. What is a patient preference trial? *BMJ* 1998; 316: 360.