COCHRANE CONSUMERS & COMMUNICATION REVIEW GROUP

How to cite this document: Ryan R, Hill S, Prictor M, McKenzie J. Cochrane Consumers and Communication group Study Quality Guide. CCCG http://cccrg.cochrane.org/author-resources. La Trobe University, Melbourne. Published May 2013. Approved (S. Hill) May 2013. Accessed [date]

STUDY QUALITY GUIDE

GUIDE FOR REVIEW AUTHORS ON ASSESSING STUDY QUALITY

CONTENTS

This guide was prepared by Rebecca Ryan and edited by Dominique Broclain, Dell Horey, Sandy Oliver and Megan Prictor. This paper was approved by Sophie Hill, Coordinating Editor of the Consumers and

Communication Review Group on 5 March 2007. It was substantially revised in 2011 by Rebecca Ryan, Sophie Hill, Megan Prictor and Joanne McKenzie. **Last updated: 30 May 2013**

People using this guide should check the website http://cccrg.cochrane.org/author-resources regularly to ensure they have the latest version.

QUICK LINKS TO KEY SECTIONS

2.0 INTRODUCTION

The primary purposes of this guide are:

- ☐ To provide authors and referees with general guidance on assessing the risk of bias of included studies;
- ☐ To provide authors and referees with information about the reporting of the risk of bias of included studies in order to improve the consistency of reporting;
- ☐ To assist authors to make decisions about appropriate assessment criteria for studies included in their reviews.

This guide is essential reading at both the title development and the protocol stages for people conducting Cochrane systematic reviews. It is also highly recommended at the review stage.

Structure

- ***Section 1*:** What is risk of bias?
- ***Section 2*:** Why assess risk of bias?
- ***Section 3*:** Bias: what you should consider
- ***Section 4*:** How to include risk of bias assessment in systematic reviews.
- ***Section 5:*** Risk of Bias assessment: randomised controlled trials and non-randomised controlled trials
- ***Section 6:*** Quality assessment: qualitative studies.
- ***Section 7:*** How to report risk of bias: tips for authors.

This document must be read in conjunction with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), which gives extensive guidance on the conduct of systematic reviews. Referencing to appropriate sections of the *Handbook* is given throughout. Readers are also referred to the *Cochrane Collaboration Open Learning Material for Reviewers*, Version 1.1, November 2002 (Updated in 2009), which provides additional information on methodological issues in conducting a systematic review. More recent online training materials are available for Cochrane contributors are http://training.cochrane.org/

Additional sources of information for review authors, with brief descriptions of their key points, are provided in Appendix A. Appendices to the accompanying ***Study Design Guide*** contain glossaries of useful terms for study design and the characterisation of study design. These are adapted primarily from the *Cochrane Glossary*, with additional material from other sources.

2.1 WHAT IS 'RISK OF BIAS'?

- *See Cochrane Handbook for Systematic Reviews of Interventions, Chapter 8.*
- *See Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.4 Quality assessment (available at http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf)*

Quality is itself a difficult concept to qualify or to quantify. It's a complex idea that means different things to different people[1], and there are numerous tools in existence with which to assess study quality. However, there is little evidence available to help guide the selection of quality assessment tools to be used in systematic reviews[2]. Further, the reporting of quality in published studies is often poor, which can increase the difficulty of assessing relevant information.

Despite these difficulties, most definitions of quality or validity[3] used in systematic reviews involve some measure of the methodological strength of the relevant study, or how able it is, through its design and its conduct, to prevent systematic errors, or **bias**.

The following is taken directly from the *Cochrane Handbook for Systematic Reviews of Interventions, Section 8.1*.

8.1 Introduction

The extent to which a Cochrane review can draw conclusions about the effects of an intervention depends on whether the data and results from the included studies are valid. In particular, a meta-analysis of invalid studies may produce a misleading result, yielding a narrow confidence interval around the wrong intervention effect estimate. The evaluation of the validity of the included studies is therefore an essential component of a Cochrane review, and should influence the analysis, interpretation and conclusions of the review.

The validity of a study may be considered to have two dimensions. The first dimension is whether the study is asking an appropriate research question. This is often described as 'external validity', and its assessment depends on the purpose for which the study is to be used. External validity is closely connected with the generalizability or applicability of a study's findings, and is addressed in Chapter 12.

The second dimension of a study's validity relates to whether it answers its research question 'correctly', that is, in a manner free from bias. This is often described as 'internal validity', and it is this aspect of validity that we address in this chapter. As

[1] For example, it can mean the internal and/or external validity of a study, the clinical relevance of the research question that is studied, the appropriateness of the data analysis and the presentation of the findings, or the ethical implications of the interventions under evaluation.
[2] There is limited empirical evidence on the relationships between specific aspects of study quality and the actual findings of studies. There is some evidence that studies which lack allocation concealment and proper blinding may overestimate the effects of interventions, but the effects of other elements of quality on outcomes has not yet been established empirically.
[3] **Validity**, in general terms, is the extent to which the result (of a measurement or of an intervention/ study) is likely to be true and free from systematic errors.

> *most Cochrane reviews focus on randomized trials, we concentrate on how to appraise the validity of this type of study. Chapter 13 addresses further issues in the assessment of non-randomized studies, and Chapter 14 includes further considerations for adverse effects. Assessments of internal validity are frequently referred to as 'assessments of methodological quality' or 'quality assessment'. However, we will avoid the term quality, for reasons explained below. In the next section we define 'bias' and distinguish it from the related concepts of random error and quality.*

Variations in study quality can explain differences in the findings of studies that are included in a systematic review. As a result, the quality of a study will reflect the strength of the evidence that can be drawn from it. In other words, it reflects the confidence that we can have that the results of a study reflect the 'truth'. By extrapolation, the quality of studies included in a systematic review will affect the confidence we can have in the results of the systematic review. See resources listed in ***Appendix A*** of this guide for further information on quality and validity.

Generally, assessment of study quality includes assessment of at least some elements of the **internal validity** of the study[4]. This is the degree to which the design and the conduct of the study avoid bias (Jadad 1998). Simply put, it is the degree to which we can have confidence that the results of the study reflect what is 'true.' Higher quality studies are more likely to produce results that are closer to the true result, as they are less prone to bias or distortions from the true value. Evaluating the internal validity of a study therefore involves assessing the measures that were used to try to prevent or minimise bias. These might include evaluation of elements such as randomisation, allocation concealment and blinding.

The guidance in this document focuses mainly on the assessment of internal validity and related issues for studies included in systematic reviews. Internal validity is the least context-dependent aspect of study quality, and some elements of internal validity are empirically related to study results. For example, trials with inadequate allocation concealment have been shown to consistently overestimate the effects of interventions. **Assessment of study quality should therefore *always* include an assessment of internal validity**.

Other aspects of study quality may also be considered and evaluated according to how relevant they are to the scope and to the particular question(s) addressed by the review. For instance, studies included in systematic reviews should also be evaluated in terms of **external validity**. This is the extent to which the results of the study can be applied, or **generalised**, to the population outside the study; in other words how the study's results apply to the real world.

[4] **Internal and external validity**: *Internal* validity is the extent to which the study's design, conduct, analysis and presentation have minimised or avoided biased comparisons of the intervention under investigation. *External* validity is the precision and extent to which it is possible to generalise the results of the study to other settings (ie. to real life clinical situations) (Jadad 1998).

That said, evaluating internal validity is necessary, but *not* sufficient, to comprehensively evaluate a study included in a systematic review. The guide that will follow on the analysis and interpretation of the results of studies included in a systematic review [still in development, May 2011] deals in more detail with issues relating to the interpretation and external validity of studies. The current document will focus primarily on aspects of quality relating to internal validity. It will address the quality assessment of randomised controlled trials (RCTs) and of studies with non-randomised design, as both are relevant to the scope of the Cochrane Consumers and Communication Review Group and may be included in systematic reviews coordinated by the Group.

Much of the research on quality assessment to date has focussed on the evaluation of RCTs. As a result, this is where the most comprehensive advice is currently available. However, different aspects of quality apply to studies of different designs, and recently there has been increasing interest in methods of assessing and reporting quality for a range of different study types. This document attempts to provide guidance to authors on quality assessment of both randomised and non-randomised studies.

2.1.1 GENERAL POINTS TO REVIEW AUTHORS ON REPORTING OF QUALITY

An important point to be aware of when conducting a systematic review is that the methods you have employed throughout the review need to be **transparently reported**. You should aim to produce a review that is **reproducible** by anyone who reads it. That is, anyone should be able to conduct the review on the basis of what you have written and come to similar conclusions. This is only possible if you **explicitly** report the details of each decision you make throughout the review process.

All decisions that you make throughout the course of conducting and writing your review need to be logical and justified. They must also be clearly and consistently reported. While these are important issues throughout the entire review process, they may be especially important when you start to consider how you will evaluate the quality of included studies (which criteria you will evaluate studies against); how you will report study quality as a component of the systematic review; and how you will incorporate quality assessment into interpreting the results of the review.

For example, if your review includes only RCTs, how will you decide if a study is really an RCT? For studies that report that they ‘randomly’ assigned participants but do not provide any details about their randomisation method, will you decide to include or exclude them from your review? Will you await response to author contact to confirm that they used a truly randomised technique? What will you do if the authors are not contactable? Decisions like these that you

make during the review process need to be explicitly reported, so that the logical processes you have followed throughout the review are clear to readers.

There are many complex decisions to make when conducting a systematic review. These decisions may be further complicated by variable quality of included studies, poor quality of reporting, complex interventions, complex outcomes or means of outcome assessment, or by other factors. Such factors may make the task of systematically reviewing the evidence on a particular intervention an involved one. However, these factors are not in themselves problematic, as long as you are **transparent** and **systematic** in the way that you report these issues and the decisions that you reach as you progress through the review.

2.1.2 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

- Assessment of study quality should include explicit and systematic evaluation of **internal** validity. This should be related to the review's findings, its limitations, and to the overall question being addressed by the review. Specific elements of study quality that should be considered are covered in later sections of this guide, ***see Section 2.4*** onwards.
- Aspects of the study's **external** validity should also be systematically considered and reported when examining the broader questions of the relevance and applicability of the study.
- The aim is to produce a systematic review that could be replicated on the basis of what you have written; therefore all decisions made throughout the review process must be explicitly reported in a **clear**, **logical** and **transparent** way. This should include reporting of the various decisions you made when conducting the review. For example, how did you decide whether to include a study or not? What decisions did you have to make about how studies were reported?

2.2 WHY ASSESS RISK OF BIAS?

- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8*
- *See Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.4 Quality assessment (available at http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf)*
- *See Spiegelhalter et al (2000)*
- *See Moher et al (1999)*

Studies, even well-designed ones, are never perfect in their design or execution, and there are many different ways in which they can be compromised or limited. These limitations (or various biases, see the following ***Section 2.3***) can affect the results of a study. Biased studies are more likely to produce misleading results than those that are rigorously designed and conducted.

A systematic review is only as good as the studies upon which it is built. Including biased studies in a systematic review can therefore produce misleading results. Even if high quality methods are followed for the conduct of the review itself, if studies with serious biases are included and these are not adequately accounted for or acknowledged, poor quality evidence will arise from the review.

Assessing the risk of bias of a study that might be included in a Cochrane review can be useful in a number of ways.

- Authors can set a minimum threshold in order for a study to be included in the review. For example, review authors may decide to exclude all RCTs that do not have adequate sequence generation (randomisation). Note that such decisions should be made and reported in advance, or *a priori.*
- A second use for risk of bias assessment can be to initially include studies at higher risk of bias, but to successively remove them from the analysis. This can allow review authors to determine how much, if at all, the result is affected. This approach is called a **sensitivity analysis** as it tests how 'sensitive' the results are to the risk of bias of the included studies.
- Finally, risk of bias assessment of the included studies can be used to help guide and structure the discussion and interpretation of the review's results, to assist in determining the strength of evidence that can be drawn from the review, and to guide recommendations or implications and directions for future research.

Variations in the validity of studies can explain differences in the results of the studies that are included in a systematic review. More rigorous studies – or those at lower risk of bias - are more likely to yield results that are closer to the 'truth'. Analysis of results from studies of variable validity can result in false conclusions being drawn about the effects of an intervention[5]. Even if there is no variability in the results of the studies included in a systematic review, it is still important to **systematically evaluate**, or **critically appraise**, all of the included studies. In such an instance, while the results may all be consistent, (that is, they report similar effects of the intervention on specific outcomes), the studies from which they come may still have serious methodological flaws. This means that the conclusions drawn from the review would be much weaker: you could have less confidence that they accurately reflect the true effect of the intervention than if the studies had all been rigorously designed and conducted and had yielded consistent results.

Therefore it is vital to be able to recognise and systematically evaluate the likely contribution of different sources of bias in included studies. It is essential to assess **whether** and **how** these may have affected the results in order to determine how confident you can be in the study's results, and that they reflect reality rather than the influence of other factors.

2.2.1 STUDY DESIGN VERSUS STUDY QUALITY

- *See Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.4 Quality assessment, available at http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf*

While design and quality are related characteristics of any study, they actually represent different components of a study. **Design** refers to how the study has been set up, or its framework. **Quality**, in comparison, is a complex concept that reflects how well the study was designed *and* how well it was executed. Hence, while the design of a study affects its quality (in part, how able it is to prevent bias), both quality and design determine the strength of the evidence that can be drawn from the results of any particular study.

An important point to realise is that not all study designs are equal when it comes to answering different types of research question. The **ranking** or **hierarchy** of different study designs depends on the question that is being asked (see the Group's ***Study Design Guide*** for more information; see also the *Cochrane Handbook for Systematic Reviews of Interventions,* Section 5.5; and the *CRD's guidance for undertaking reviews in health care*, Section *1.2.2.3*). When considering studies of effectiveness (that is, the effectiveness of

[5] For example, if the study *over*estimates the effects of an intervention, the study may make a **'false positive'** conclusion (ie. wrongly conclude that the intervention works). If, in comparison, the study *under*estimates the effects of an intervention, it may make a **'false negative'** conclusion (ie. wrongly conclude that there is no effect of the intervention).

therapy or some other type of intervention), the questions tend to focus on comparisons, or how an intervention compares with no intervention or with alternative intervention(s). To answer questions about the effects of interventions– which is what most Cochrane reviews address - comparative studies that minimise bias will be the highest in the hierarchy, or the most suitable types of studies to investigate these questions.

It is for these reasons that RCTs are regarded so highly. RCTs are considered to be the 'gold standard' for addressing questions of effectiveness, as they are designed to minimise bias. For example, randomly assigning participants to treatment groups reduces the likelihood that the groups will differ on important baseline characteristics. This includes those characteristics of which the investigator may not be aware. By using chance it is likely that the two groups will be equivalent on important characteristics. Although it is possible to control for known confounders using other study designs,[6] randomisation is the only was to control for confounders that are not known. This means that selection bias is likely to be minimised in RCTs; whereas the means of allocating participants in other study designs may not be as 'fair,' and groups of participants may differ on important characteristics at baseline.

Many different ways of classifying or ranking studies have been developed[7]. Although these schemes differ, broad similarities also exist. Below (Figure 1) is a general classification scheme or hierarchy of studies in terms of the suitability of their design to answer questions of effectiveness. Such hierarchies are based on the susceptibility of different study designs to bias.

Figure 1: General hierarchy of study designs to answer questions of effectiveness

RCTs

Non-randomised studies, that is:

- Quasi-randomised controlled trials
- Controlled before-and-after studies (CBAs)
- Interrupted time series studies (ITSs)

Controlled observational studies

- Cohort studies
- Case-control studies
- Case series

[6] For example, by 'matching' participants in the different study groups on important characteristics.

[7] See Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.4 Quality assessment (available at *http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf*). The exact schema adopted is not essential, but it is important to have some idea of the general level of evidence that different types of study may represent.

Many Cochrane systematic reviews specifically include RCTs only, as they are considered the strongest type of evidence to answer effectiveness questions. Reviews that fall within the scope of the Consumers and Communication Review Group, however, include many 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. Further, many kinds of research question may benefit from the inclusion of non-randomised studies, such as public health and health promotion interventions.

We have therefore decided to follow the advice developed by the Cochrane Effective Practice and Organisation of Care (EPOC) Review Group (see also http://epoc.cochrane.org/epoc-resources). This advice allows review authors to decide whether to include a limited range of experimental study designs other than RCTs in their analysis and 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 studies (CBAs)
- Interrupted time series (ITS)

It should be noted that any hierarchy of study designs should be treated with flexibility and used as a **guide** only. For example, RCTs should rank at the top of such a hierarchy only when they are well-conducted.[8] This is why it is necessary to assess the risk of bias of the studies included in a review – even if it only includes RCTs – because not all studies of the same basic design will be equally well conducted, and this may affect the results and conclusions of the review.

The nature and type of the review question, as well as pragmatic considerations such as author resources, should determine the choice of studies to include in a review. In terms of providing answers to a review question, the suitability of different studies should be obvious if the question is formulated clearly and explicitly. A review should be based on the best quality evidence that is available. If authors decide to include studies other than RCTs in their review, that their reasons for doing so must be **clearly** and **explicitly** stated.

2.2.1.1 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

- Think about the type of question you want to answer with your systematic review**.** What types of study design would be most appropriate to answer this question?

[8] If there are serious flaws or limitations in the design or conduct of an RCT, it should be considered as being of lower quality than one that is performed well.

- On the basis of the above, what studies will you include in your review? Will you include studies other than RCTs? If you decide that it is necessary to do so in order to answer your review question, you will need to explicitly **justify your decision** in the review. **This must be stated clearly at the protocol stage of your review, together with a rationale for your decision.**
- As study designs can be fundamentally different in their structure and conduct, they also need to be differentially assessed for quality. That is, different elements of quality, or quality criteria, need to be considered for different study designs. **Specific quality items to be assessed for each different type of study design to be included in the review must be clearly and systematically described at the protocol stage.** These items must be described clearly and separately for each of the different study designs to be included in the review.

For further information, see *Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.4 Quality assessment*
(available at http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf)

2.2.2 QUALITY OF STUDIES VERSUS THE QUALITY OF REPORTING OF STUDIES

See Cochrane Handbook for Systematic Reviews of Interventions, Section 8.3.2

A major problem when assessing study quality is that the reporting of studies is often poor. Authors often neglect to provide adequate details of the methods they used to reduce bias. Even if the methods are reported, many authors simply report such measures as being 'done,' rather than providing detail that might allow an independent evaluation of the adequacy of the approach they used. This often makes the systematic assessment of study quality quite difficult. In cases where there is inadequate reporting, a review author may decide to assume that adequate measures were taken, or alternatively, that they were not taken[9]. Either decision may, however, be incorrect and may lead to an inaccurate representation of the study when included in the systematic review.

Given that many journals and other publications impose strict word limits on study reports, it is perhaps not surprising that authors often fail to elaborate on their methods. Standards for the reporting of trials have now been developed, and will help to standardise and improve the

[9] That is, as a review author, you could decide to consistently rate those instances where there is insufficient detail available in a consistent manner as either done or not done. In either case you should report this decision in the text of your review.

reporting of future studies[10]. However, those studies published previously remain likely to pose significant problems for review authors, when insufficient detail is reported to fully assess study quality[11]. This is often true in the case of Cochrane reviews, where the inclusion of studies is typically not restricted by the year of publication and it is common for a review to include studies dating back several decades.

Rather than assuming that something was not done if it was not reported, a more thorough approach should be taken. Review authors should contact the study authors in the first instance to request further information. If this is not possible or not successful[12], the reporting of quality should reflect this underlying uncertainty. The Cochrane Risk of Bias tool (covered in more detail in following sections) encourages this type of transparent reporting and provides a format for providing this type of information within a review. For example, if details about a particular issue are not provided in a trial report, this can be reported as 'not described'.

Building in this capacity for reporting uncertainty may allow more accurate reporting on the status of the study than simply assuming that critical elements of the study were done or not done. This may, in turn, be more meaningful when it comes time to interpret the study's results and to incorporate them into the review's overall findings, and is also informative to readers of the review.

2.2.2.1 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

- Variations in study quality (validity) can explain differences in the results of studies included in a systematic review. More rigorous or higher quality studies are more likely to closely estimate the truth. Study quality should therefore be assessed **systematically**, **transparently** and in an **explicit** manner.
- Poor reporting of elements of a study that relate to quality is not the same as poor quality of the study itself.
- The Consumers and Communication Review Group recommends that if not enough detail is provided to assess quality criteria for included studies:
  - Review authors should contact the study authors to see if more information about the design and/or conduct of the study can be obtained (for sample letters contact the Managing Editor).

[10] For example, the Consolidated Standards of Reporting Trials (CONSORT) statement was developed to improve the reporting of RCTs (see www.consort-statement.org); while the Transparent Reporting of Evaluations with Non-randomised Designs (TREND) statement provides guidelines for the standardisation of reporting for non-randomised studies (see www.cdc.gov/trendstatement/). See Appendix A of this guide for further information.

[11] While the reporting of study quality is always improving, older studies may be more limited in their reporting. As a review author, you need to be consistent and clear about how you intend to deal with studies where a lack of information or detail is problematic in their evaluation.

[12] For example, for studies conducted a long time ago (10 or 20 years), even if study authors can be contacted they may no longer have access to the data or the specific details of the conduct of the study.

- Where further information is not available, it is possible to report this directly using the Risk of Bias tool within reviews.

2.3 BIAS: WHAT YOU SHOULD CONSIDER ABOUT INCLUDED STUDIES

- *See Cochrane Handbook for Systematic Reviews of Interventions, Chapter 8*
- *See also Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.4 Quality assessment, available at http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf*
- *See also Jadad (1998): Chapters 3 and 4*

Bias is any systematic error or factor, whether it is recognised or not, that distorts the results or inferences of a study. It can occur because a study design is poor, or because a study is poorly conducted. As mentioned above, although some study designs are better than others at avoiding or minimising bias, you cannot assume that all studies of a particular design will have been conducted equally well, or as well as they might have been. This is why it is necessary to assess the risk of bias of all studies included in a review: to be able to tell how well likely it is that there are systematic errors that might affect the results of the studies and the review. Bias can arise from many different sources. There are five major sources of bias that can arise in studies examining the effects of healthcare interventions:

- Selection bias
- Performance bias
- Attrition bias
- Detection bias
- Reporting bias

The table below is taken directly from The Cochrane Handbook for Systematic Reviews of Interventions, Section 8.4.

Table 8.4.a: A common classification scheme for bias

Type of bias	Description	Relevant domains in the Collaboration's 'Risk of bias' tool
Selection bias.	Systematic differences between baseline characteristics of the groups that are compared.	• Sequence generation. • Allocation concealment.
Performance bias.	Systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.	• Blinding of participants and personnel. • Other potential threats to validity.
Detection bias.	Systematic differences between groups in how outcomes are determined.	• Blinding of outcome assessment. • Other potential threats to validity.

Attrition bias.	Systematic differences between groups in withdrawals from a study.	• Incomplete outcome data
Reporting bias.	Systematic differences between reported and unreported findings.	• Selective outcome reporting (see also Chapter 10).

2.3.1 SELECTION BIAS

- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8.4, 8.9 (sequence generation), and 8.10 (allocation sequence concealment)*

Selection bias occurs at the point of allocating participants to groups at the beginning of a trial, and results from systematic differences in the characteristics of the participants in each study group. These differences can relate to prognostic factors, or can be due to the differential responsiveness of participants to the intervention.

Minimising selection bias:

☐ **Random assignment** of large numbers of participants to study groups can minimise this type of bias (see ***Study Design Guide, section 1.3.2.3*** for more information on randomisation). However, allocation must also be **adequately concealed**, so that neither the investigator nor the participant can influence which group a participant is entered into. Randomisation without adequate allocation concealment does *not* adequately protect against selection bias.

- **Allocation sequence concealment** means that the process of allocating participants or actually placing them to the different groups to which they have been randomly assigned must be concealed from the person recruiting participants into the trial. If the allocation is not concealed, there is potential for systematic differences in characteristics between participants allocated to the different arms or treatment groups of the study. Evaluating allocation concealment involves a judgement about how the randomisation sequence was applied when participants were actually placed in the study groups, and whether it was possible for the researcher to subvert the randomisation process. For example, if random numbers are picked from an open envelope, there is potential to replace slips and to re-choose until the 'right' allocation is selected for a particular participant. Such practices subvert the randomisation process, so that it is no longer truly random.

☐ **Allocation concealment is different to blinding**[13]. 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

[13] **Blinding** prevents participants, providers and/or outcome assessors from knowing which participants are receiving the intervention(s) in the study.

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

☐ Allocation concealment has been shown to impact importantly on the results of trials. Several studies have demonstrated that when allocation concealment is inadequate or unclear, the treatment effect is overestimated.

☐ Allocation concealment can be achieved in different ways, but different approaches can have varying levels of success; that is, they are more or less likely to be interfered with or subverted.

2.3.1.1 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

RANDOMISATION

'**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.
For more information refer to the *Cochrane Handbook* section 8.9.

ALLOCATION SEQUENCE CONCEALMENT

Adequate allocation concealment approaches (sequence for allocating participants to groups is truly hidden from investigators):

- Central computer randomisation and allocation (for example, allocation by a central office that is unaware of the participants' characteristics);
- On-site computer from which assignment can only be determined after entering the patient's data;

☐ Pre-numbered or coded identical containers administered serially to participants;

☐ Serially (sequentially) numbered, opaque sealed envelopes.

There are also other approaches, similar to those above, that if administered by a different person to the one who generated the allocation scheme can also be considered adequate. Certain other approaches might also be new or innovative and not fit with those described above but can still provide adequate allocation concealment.

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

- Any 'open methods' (ie. transparent before allocation) eg lists on noticeboards, open lists, open envelopes;

☐ Non-opaque envelopes;

☐ Odd or even date or medical record number;

☐ Dates of birth or days of week;

☐ Alternation.

In cases where studies do not report any approach to concealing the allocation sequence, it should be assumed that the adequacy of approaches is **unclear**. Examples of this might include cases where studies state that a list or table or random numbers was used but do not provide further details; or stating that sealed envelopes were used .

Consider:

☐ Was there any way that those conducting the study could have known what the random allocation sequence was? Could allocation to groups be manipulated?[14]

☐ Were the study groups comparable at the start of the study, just after randomisation?

2.3.2 PERFORMANCE BIAS

- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8.11*

This bias arises from systematic differences in the way that care is provided, or from exposure to factors other than the intervention that is being studied.

☐ Performance bias occurs during the treatment and/or delivery of the intervention(s).

[14] If the answer to this questions is 'yes,' the method of allocation concealment can be considered inadequate.

☐ It arises as the result of differences in the way that the intervention is delivered to the different study groups; that is, not only does the intervention differ between groups, the method of delivering it also differs. The impact of the intervention alone therefore cannot be assessed.

Minimising performance bias:

☐ Performance bias can be avoided when both the **participants** and the intervention **provider** are **blind** to the comparison group to which the participant belongs.

☐ Caregivers/providers, outcomes assessors, analysts and the participants themselves can all be eligible for blinding, leading to trials being described as single, double, triple, or even quadruple blind[15].

2.3.2.1 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

Consider:

- Were there any other factors that might have affected the study's results?

☐ Of those involved in the study, were the following blinded:

  - o Participants?
  - o Provider/ caregivers?

- We suggest that you **avoid** the use of terms like 'double blind' etc. unless you clearly and explicitly define them in your review (eg. 'we will use the term double blind to mean…'). Instead we suggest that you specifically report on each component of blinding **separately** (eg. whether participants were blind and how adequate the approach to blinding was; whether providers were blinded adequately, and so on). This makes it very clear who was blinded throughout the study; allows comment on the adequacy of each possible component of blinding; and allows the potential bias introduced by inadequate blinding to be identified clearly both by you the review author and the readers of your review.

☐ Blinding of outcome assessors is addressed under Detection Bias (at 2.3.3) below.

- Blinding of data analysts and manuscript writers is not explicitly addressed in the Cochrane Collaboration's Risk of Bias tool.

2.3.3 DETECTION (OR ASCERTAINMENT) BIAS

- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8.12*

[15] There is very little consensus regarding what actually constitutes single/ double/ triple blind, and although these terms are commonly used there is little consistency in the way that they are used. See Appendix A of this guide fore more on blinding and interpretations of blinding.

This bias arises due to differences in the way that outcomes are assessed. Detection bias occurs at the point of follow-up (that is, at the time of outcome assessment). The *Cochrane Handbook* says:

> *All outcome assessments can be influenced by lack of blinding, although there are particular risks of bias with more subjective outcomes (e.g. pain or number of days with a common cold). It is therefore important to consider how subjective or objective an outcome is when considering blinding. The importance of blinding and whether blinding is possible may differ across outcomes within a study…*
>
> *Blinding of outcome assessment can be impossible (e.g. when patients have received major surgery). However, this does not mean that potential biases can be ignored, and review authors should still assess the risk of bias due to lack of blinding of outcome assessment for all studies in their review.*

Minimising detection bias:

☐ Detection bias can be minimised when the outcome assessor is **blind** to participant groups. A lack of blinding can exaggerate the estimated effect of treatment. It is especially important that outcome assessors be blind to treatment allocation in cases where the outcomes are subjective (eg pain, satisfaction)[16].

2.3.3.1 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

Consider:

☐ Were outcome assessors blind to the treatment allocation?

2.3.4 ATTRITION BIAS[17]

- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8.13*

This bias arises due to systematic differences between study groups in the withdrawals or exclusion of people entered into the study. That is, it is due to systematic differences between study groups in the numbers of participants who are 'lost' during the study (ie. it occurs *after* the allocation and inclusion of participants in the study). Attrition bias occurs over the duration of the trial.

[16] Note that allocation concealment helps to prevent selection bias and protects the randomisation sequence *before* and *until* the interventions are given to the study participants. In contrast, blinding helps to prevent detection bias, by protecting the randomisation sequence *after* allocation.

[17] Attrition bias is also known as **exclusion bias**, as it may result from the post-randomisation exclusion of participants from the study. See Appendix A of this guide for more information on attrition bias.

For example, in a study examining the effects of a particular drug, more people receiving the active drug may leave a study due to side effects than those assigned to the control group. This could lead to an *over*estimation of the effectiveness of the intervention, as only those people who tolerated the intervention well remained in the study and were assessed. Attrition bias could also lead to an *under*estimation of the adverse effects or harms associated with the intervention, if those with the most severe side effects left the study and were not measured and included in analysis.

In the case of RCTs, random assignment is not itself sufficient to make inferences infallible. The validity of inferences depends on the assumptions that random assignment produced comparable groups at baseline, and that systematic attrition over the study period did not occur. Studies often fail to report how many participants who were enrolled in the study were followed throughout it, which makes sample attrition difficult to assess. Even if not reported however, it is important to note the attrition of the samples as a possible source of bias; and to interpret the results of studies with large rates of attrition with some caution.[18]

Minimising attrition bias:

☐ Attrition bias can be minimised when the **proportion** and **characteristics** of the participants lost to follow-up (that is, those lost from the trial *after* their inclusion in the trial and assignment to groups) are described and are comparable for the different study groups. Note that this should include an account of **everyone** who was enrolled and randomised by the study. The reasons for participants dropping out of the study can provide valuable qualitative information when interpreting and discussing the results of the study. Studies should account for every participant initially enrolled in the study for **each group**. Ideally, analysis should always be based on the intention-to-treat (ITT).

- **Intention-to-treat (ITT) analysis** involves analysing participants in the groups to which they were allocated at the start of the study, even if they didn't receive the intervention, if they deviated from the protocol or they withdrew from the study. Note that ITT analysis is most often applied to RCTs and not to other types of study design, although a similar approach can be taken with non-randomised studies.

  ITT analysis is a way of preserving the equivalence of the groups established at the start of the study (baseline) by the process of random allocation. If the study participants are randomly assigned to groups and there are adequate methods of allocation concealment, the groups should (theoretically) be similar at the start of the study and should be analysed on this basis. In order to prevent attrition bias, the groups also need to be similar at the

[18] The way that losses of participants throughout the study are handled has much potential to introduce bias. However, reported attrition following allocation has not been found to be clearly related to bias. It is not yet clear exactly what the relationship between attrition and bias are, so review authors should be cautious when dealing with this aspect of study execution.

end of the trial (Heritier *et al* 2003). The alternative approach, to analyse participants according to whether the participants did or did not receive the intervention, can markedly change the characteristics of the different study groups, and so make the process of random allocation ineffective.

Often studies report outcomes only for those participants that were followed up or completed the study; or they do not provide enough information to determine whether ITT analysis was used. Even if ITT analysis is **not** used, it is important that the participants lost from the study (after they were enrolled in the study) are clearly accounted for. The way in which they were dealt with in the study's analysis must also be clearly and completely described. Many studies fail to report any details on these issues.

- See the *Cochrane Handbook* *section 16.2*
- See ***Appendix A*** of this guide, for additional sources of information (listed under **attrition bias** and **ITT analysis**).

2.3.4.1 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

- Often the numbers of participants involved throughout studies (participant flow) are poorly reported, although this is improving in more recent trial reports. Review authors should contact study authors to request further information about participant flow if it is not clear from the study report.
- Note that even where participant flow is reported, participant numbers are often implied rather than explicitly stated or reported at each stage of the study. Authors of such studies should be contacted to provide further information to clarify participant flow. If information is not available, the potential bias introduced by participant attrition from the study should be noted in the review.

Consider:

☐ Was follow-up adequate?

- How many participants completed the study for each group and for each outcome?
- How many participants did not complete the study for each group and for each outcome?

☐ Were the study groups comparable at the end of the study?

☐ Were those followed up comparable to those who were lost from the study?

☐ Was ITT analysis used?

- If ITT analysis was not used, how were those who dropped out (or were excluded post randomisation) treated in the analyses?

2.3.5 REPORTING BIAS

- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8.14 and Section 10*
- *See also Egger et al (2003) and Song et al (2004) for more information*

This bias arises because of the existence of systematic differences between findings that are reported and those that are not. This can happen in a number of ways.

Within studies

Selective outcome reporting: it is possible for only some outcomes to be included in the trial report. This means that some of the outcomes have been omitted from the report. For example, a study might fail to report on all primary outcomes that were pre-specified; may not report on a key outcome that would usually be reported for that type of study; may report data incompletely for one or more outcomes; or may otherwise fail to convince a reader that all outcomes that were pre-specified have been fully reported. A study can also incompletely report all data for a given outcome, for example, selectively reporting data only at some but not all pre-specified time points. Selective outcomes reporting arising in any of these ways may systematically distort results and introduce bias.

Across studies

Reporting biases: at the study level, it can also happen that the direction of the research findings affects the dissemination of the research. The best-recognised type of reporting bias is known as publication bias.

Publication bias arises due to the fact that studies with statistically significant 'positive' results are more likely to be published than those studies with results that are not. This means that in the collected published research literature on a given topic there are less likely to be studies reporting non-significant results. This introduces a systematic error into the collected research evidence, for example, systematic reviews that do not include unpublished studies may be more likely to overestimate the effects of the intervention.

There are also other sorts of reporting biases that can be introduced when conducting a systematic review if the search strategy is not comprehensive[19]. Examples are shown in the table below.

[19] **Publication and associated biases** can be formally evaluated using various statistical techniques, including **funnel plots**. See the Cochrane Handbook chapter 10.4 for details.

The table below is taken directly from *The Cochrane Handbook for Systematic Reviews of Interventions*, Section 10.1.

Table 10.1.a: Definitions of some types of reporting biases

Type of reporting bias	Definition
Publication bias	The *publication* or *non-publication* of research findings, depending on the nature and direction of the results
Time lag bias	The *rapid* or *delayed* publication of research findings, depending on the nature and direction of the results
Multiple (duplicate) publication bias	The *multiple* or *singular* publication of research findings, depending on the nature and direction of the results
Location bias	The publication of research findings in journals with different *ease of access* or *levels of indexing* in standard databases, depending on the nature and direction of results.
Citation bias	The *citation* or *non-citation* of research findings, depending on the nature and direction of the results
Language bias	The publication of research findings *in a particular language*, depending on the nature and direction of the results
Outcome reporting bias	The *selective reporting* of some outcomes but not others, depending on the nature and direction of the results

While there is empirical evidence that certain types of bias can substantially alter study results, the effects of other biases on results are not yet clear. This complicates the assessment of study quality, as ideally we only want to evaluate those sources of bias that have been shown empirically to affect results. While the effects of some potential sources of bias on study results are not yet clear, it is logical to suspect that certain biases, particularly selection bias, performance bias, attrition and detection biases, may influence study results. Hence, it is important to systematically and transparently assess studies against criteria that evaluate the likely impact of each of these biases.

- For more information, see the *Cochrane Handbook for Systematic Reviews of Interventions*, Sections 8 and 10

2.4 INCLUDING ASSESSMENTS OF STUDY QUALITY IN SYSTEMATIC REVIEWS

- *See the Cochrane Handbook for Systematic Reviews of Interventions, Section 8*
- See Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.4 Quality assessment (available at http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf)

2.4.1 GENERAL POINTS ON THE ASSESSMENT AND INCORPORATION OF STUDY QUALITY

Many checklists and scoring systems have been developed and used to assist people in assessing study quality, in particular for assessing the quality of randomised controlled trials (RCTs). The Cochrane Collaboration has developed a tool to assist review authors to assess the quality of studies included in reviews, known as the Risk of Bias tool. This tool specifies a number of different domains which may predispose studies to bias, and provides a detailed framework for assessing and reporting the risk of bias of included studies on each of these domains.

The Risk of Bias (RoB) tool can be used to assess the quality of RCTs. It can also be adapted to assess the quality of quasi-RCTs controlled before-and-after (CBA) studies, and interrupted time series studies (ITS), all of which are occasionally eligible for inclusion in systematic reviews of complex interventions such as those published through the CC&CRG.

2.4.1.1. SUMMARY: KEY POINTS FOR REVIEW AUTHORS

In the Consumers and Communication Review Group, we require that the Collaboration's Risk of Bias tool be used. This assessment should be reported explicitly in the review in Risk of Bias tables and review text.

Review authors should include the following text (tailored as necessary) in the Methods section ("Assessment of Risk of Bias in Included Studies") of their Protocol/Review:

> 'We will assess and report on the methodological risk of bias of included studies in accordance with the Cochrane Handbook (Higgins 2011) and the guidelines of the Cochrane Consumers and Communication Review Group (Ryan 2013), which recommends the explicit reporting of the following individual elements for RCTs: random sequence generation; allocation sequence concealment; blinding (participants,

personnel); blinding (outcome assessment); completeness of outcome data, selective outcome reporting; and other sources of bias [please specify]. We will consider blinding separately for different outcomes where appropriate (for example, blinding may have the potential to differently affect subjective versus objective outcome measures). We will judge each item as being at high, low or unclear risk of bias as set out in the criteria provided by Higgins 2011, and provide a quote from the study report and a justification for our judgement for each item in the risk of bias table.

Studies will be deemed to be at the highest risk of bias if they are scored as at high or unclear risk of bias for either the sequence generation or allocation concealment domains [or other domains of the tool; please adapt], based on growing empirical evidence that these factors are particularly important potential sources of bias (Higgins 2011).*

In all cases, two authors will independently assess the risk of bias of included studies, with any disagreements resolved by discussion to reach consensus. We will contact study authors for additional information about the included studies, or for clarification of the study methods as required. We will incorporate the results of the risk of bias assessment into the review through standard tables, and systematic narrative description and commentary about each of the elements, leading to an overall assessment the risk of bias of included studies and a judgment about the internal validity of the review's results.'

*If including quasi-RCTs add:

'We will assess and report quasi-RCTs as being at a high risk of bias on the random sequence generation item of the risk of bias tool.

*If including cluster RCTs add:

'For cluster-RCTs we will also assess and report the risk of bias associated with an additional domain: selective recruitment of cluster participants.'

*If including controlled before and after studies add:

'We will assess CBA studies against the same criteria as RCTs but report them as being at high risk of bias on both the random sequence generation and allocation sequence concealment items. We will exclude CBA studies that are not reasonably comparable at baseline.'

*If including interrupted times series add:

'We will assess and report on the following items for ITS studies: intervention independence of other changes; pre-specification of the shape of the intervention effect; likelihood of intervention affecting data collection; blinding (participants, personnel); blinding (outcome assessment); completeness of outcome data, selective outcome reporting; and other sources of bias [please add].'

If you are planning to assess specific items under 'other' sources of bias domain this should also be described. Note that assessing other sources of bias is not essential but should be guided by the specific study designs that you plan to include in the review. This might include design-specific issues (such as assessing selective recruitment of cluste participants for cluster-RCTs), baseline imbalances between groups or the likelihood of contamination.

2.5 ASSESSING THE RISK OF BIAS OF STUDIES IN REVIEWS

- *See Consumers and Communication Review Group data extraction template, available at: http://cccrg.cochrane.org/author-resources*
- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8*

RCTs are considered to represent the 'gold standard' study design for investigating the effects of interventions. This is because their design helps to avoid or to minimise many different kinds of bias[20] (see ***Study Design Guide*** for more details). However, it is important to remember that not all RCTs are well-designed or well-conducted. As a result, it is possible to have RCTs of variable quality. This in turn affects the strength of the evidence that can be drawn from their findings (see ***Section 2.2.2***: *Study design versus study quality*). Similar principles hold for the limited range of non-randomised studies that are eligible for inclusion in Cochrane reviews.

The *Cochrane Handbook* provides detailed advice, description of and practical instructions for review authors about the Risk of Bias tool.
This includes details about:

- What each domain of the tool involves;
- What issues are to be considered in making a judgement about the risk of bias in each domain; and
- How to report details from the study against the tool for each domain (practical decision rules for using the tool).

These details are captured in the RoB tables within RevMan 5, and appear within reviews so that readers have access to this information in a systematic way. For tips on how to enter data into RevMan 5, see "Risk of Bias" tables in the RevMan User Guide.

The RoB tool encourages transparent reporting of bias, by asking review authors not only to make a judgement about whether particular criteria are met (or not), but also providing information directly form the study to support the judgement made. This assists readers of the review in understanding what the study reported and how the review authors made their decisions about the study's risk of bias.

Authors are advised to consult the Handbook, especially Section 8 'Assessing risk of bias in included studies' for details on the tool and how to complete the risk of bias assessments in

[20] Note that bias is taken here to mean the various different sources of bias overed in earlier sections of this guide.

reviews. The following sections outline the RoB tool for use by review authors with the CC&CRG, for RCTs and quasi-RCTs, and for non-randomised studies (CBA and ITS studies).

2.5.1 ASSESSING RISK OF BIAS: RANDOMISED CONTROLLED TRIALS (RCTS)

- *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8, especially Table 8.5.a*

For assessing the risk of bias of RCTs included in reviews, authors should use the Cochrane Handbook and follow the guidance outlined.

Adapted from Cochrane Handbook Table 8.5.a: The Cochrane Collaboration's tool for assessing risk of bias

Domain	Review authors' judgement	Support for judgement
Random sequence generation*	***High risk*** ***Unclear*** ***Low risk***	**Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.** Quasi-RCTs must be rated as 'High Risk' for random sequence generation as the methods were not, by definition, truly random.
Allocation concealment	***High risk*** ***Unclear*** ***Low risk***	**Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.**
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*	***High risk*** ***Unclear*** ***Low risk***	**Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.**
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).*	***High risk*** ***Unclear*** ***Low risk***	**Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.**
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*	***High risk*** ***Unclear*** ***Low risk***	**Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and**

		any re-inclusions in analyses performed by the review authors.
Selective reporting	***High risk*** ***Unclear*** ***Low risk***	**State how the possibility of selective outcome reporting was examined by the review authors, and what was found.**
Other sources of bias See the *Cochrane Handbook* 8.15.1 for further examples of potential threats to validity, as well as 16.3.2 for issues relating to cluster trials (see also below), and 16.4.3 for cross-over trials	***Note: all answers should follow the format:*** ***High risk*** ***Unclear*** ***Low risk***	**State any important concerns about bias not addressed in the other domains in the tool.** **If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.**

2.5.1.1. ASSESSING RISK OF BIAS IN CLUSTER RCTS

Cluster RCTs are at particular risk of bias in terms of participant recruitment. We recommend that review authors assess whether selective recruitment of cluster members was adequently prevented during the study. This should be reported in the 'Other Sources of Bias' row in the standard Risk of Bias table for RCTs.

Criteria for judging risk of bias in recruitment of participants in cluster designs[21]

Was selective recruitment of cluster members adequately prevented during the study?

Criteria for a judgement of 'Low risk' of bias.	Those involved in the identification and/or recruitment of the cluster participants did not have knowledge of the group allocation because one of the following, or an equivalent method, was employed: • Cluster participants were recruited prior to group assignment and the same participants were followed up over time. • Cluster participants were recruited post group assignment but o carried out by a person who was blinded to the group allocation; o carried out by a person who was unaware of characteristics of the participants (e.g. clinical characteristics of patients); o eligibility criteria were such that it was unlikely to be subverted by knowledge of group allocation (e.g. all patients attending a hospital within a specified period of time included); o invited by mail questionnaire with identical information to participants in the intervention and control arms.

[21] Modified from Table 1 of Brennan et al *Continuous quality improvement: effects on professional practice and healthcare outcomes.* Cochrane Database of Systematic Reviews 2009, Issue 4. Art. No.: CD003319. DOI: 10.1002/14651858.CD003319.pub2.

Criteria for the judgement of 'High risk' of bias.	Those involved in the identification and/or recruitment of the cluster participants may have had knowledge of the group allocation. • Cluster participant identification and/or recruitment undertaken post group allocation by a person who was unblinded and who may have had knowledge of characteristics of the cluster participants.
Criteria for the judgement of 'unclear risk' of bias.	Insufficient information to permit judgement of 'Low risk' or 'High risk'.

The purpose of this domain to assess the risk of selection bias from recruitment of cluster participants into the study. The risk of selection bias is also assessed through the domains 'Sequence generation' and 'Allocation concealment' (See Chapter 8, *Cochrane Handbook for Systematic Reviews of Interventions*). In cluster-randomised trials the hierarchical nature of the design allows for differential recruitment or consent of participants, through recruitment post-randomisation (Giraudeau B et al. PLoS Med 2009, 6:e1000065). This may potentially introduce bias from selective recruitment irrespective of adequate sequence generation and allocation concealment.

Selective recruitment can occur at multiple levels when there are multiple levels of clustering. For example, in a cluster-randomised trial where primary care practices are randomised to the intervention and control groups, selection bias could occur when recruiting practitioners to the trial, patients, or both. Practitioners' knowledge of their practice allocation could affect the type of practitioners who choose to be involved in the trial and if, for example, the practitioners are involved in the recruitment of patients, they may selectively recruit patients depending on their clinical characteristics.

In cluster-randomised trials with multiple levels of clustering (e.g. practices, practitioners, and patients), the risk of bias arising from selective recruitment may differ depending on the level at which the outcome is measured. Using the above example, if all practitioners within a practice were recruited, but practitioners recruited patients, then the risk of bias arising from selective recruitment will be low for outcomes measured at the practitioner level, but may be high for outcomes measured at the patient level.

The criteria in this table for judging the risk of selection bias from recruitment of cluster participants can be applied to other designs involving clustering such as controlled clinical trials and controlled before and after studies.

2.5.2 ASSESSING RISK OF BIAS: OTHER (NON-RCT) ELIGIBLE STUDY DESIGNS

Studies of different design have different issues associated with their design and conduct. Criteria used to evaluate the quality of RCTs cannot be applied directly (without adaptation) to all studies of other designs that might be included in a review. However, major elements of the

RoB criteria used to assess RCTs can be adapted to systematically assess non-randomised studies.

Because controlled studies that are not randomised are similar in their design in many respects to RCTs, they can be assessed using a similar but slightly adapted set of criteria. Assessment of the randomisation method and suitability is obviously not relevant for a non-randomised study. However, as allocation to intervention and control groups is not determined by randomisation, review authors should pay particular attention to how the groups are chosen, the possibility of bias arising from the non-random allocation of participants, and the potential influence of baseline differences in groups upon the outcomes of the study.

At the CC&CRG we follow the guidance developed by the EPOC group, which has tailored the RoB tool for use on study designs other than RCTs. Authors should follow the decision rules laid out in this additional guidance, and which complements the Handbook guidance, in those cases where studies other than RCTs are included. This guidance is presented below.

2.5.2.1 ASSESSING RISK OF BIAS FOR CBA STUDIES

See EPOC resources, available at http://epoc.cochrane.org/epoc-resources

In particular:

- *Risk of Bias Criteria*
- *Data collection template*
- *Data collection checklist*

Inclusion criteria:

The EPOC Review Group specifies that, to be included in a systematic review, the CBA study design must meet three key criteria. These are:

- 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).
- 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.

Risk of bias assessment:

If CBA studies meet these criteria, they are eligible (at least based on study design criteria) for inclusion in a systematic review and so need to be systematically assessed for their risk of bias.

The standard Risk of Bias tool for RCTs should be utilised for this purpose, and a form suitable for use on CBAs is shown in the table.

Adapted from Cochrane Handbook Table 8.5.a: The Cochrane Collaboration's tool for assessing risk of bias; adapted using EPOC's criteria for studies other than RCTs

Domain	Review authors' judgement	Support for judgement
Random sequence generation*	***High risk*** ***Unclear*** ***Low risk***	**Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.** **CBA studies must be rated as 'High risk'. Score 'unclear' if not specified in the paper.**
Allocation concealment	***High risk*** ***Unclear*** ***Low risk***	**Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.** **CBA should be scored 'High risk'. Score 'unclear' if not specified in the paper.**
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*	***High risk*** ***Unclear*** ***Low risk***	**Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.**
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).*	***High risk*** ***Unclear*** ***Low risk***	**Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.**
Incomplete outcome data* *Assessments should be made for each main outcome (or class of outcomes).*	***High risk*** ***Unclear*** ***Low risk***	**Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.**
Selective reporting*	***High risk*** ***Unclear*** ***Low risk***	**State how the possibility of selective outcome reporting was examined by the review authors, and what was found.**
Other sources of bias* For example: - Were the intervention and	***Note: all answers should follow the format:***	**State any important concerns about bias not addressed in the other domains in the tool.** **If particular questions/entries were pre-specified in**

control groups comparable at baseline (note, if groups were not reasonably equivalent and this was not adjusted through analysis, the study should be excluded). - Have measures been taken within the study to protect against contamination? See the *Cochrane Handbook* 8.15.1 for further examples of potential threats to validity.	***High risk*** ***Unclear*** ***Low risk***	**the review's protocol, responses should be provided for each question/entry.**

* If some primary outcomes were imbalanced at baseline, assessed blindly or affected by missing data and others were not, each primary outcome can be scored separately.

Action:

- *Rate included CBA studies on each of the risk of bias assessment criteria. Each criteria should be rated as low risk (done), high risk (not done), or unclear, with a description given.*
- *Authors should rate* ***Random sequence generation*** *as 'high risk' and* ***Allocation concealment*** *as 'high risk' (as indicated in the table above).*
- *Authors should pay particular attention to the element relating to baseline comparability of intervention and control groups under the* ***Other sources of bias*** *heading, and there is a high risk of groups being unbalanced in CBA designs.*

2.5.2.2 ASSESSING RISK OF BIAS FOR ITS STUDIES

Authors are referred to the guidance provided by EPOC on ITS risk of bias assessment. Authors may wish to seek advice from the Consumers and Communication Group (cochrane@latrobe.edu.au) on adapting this to their reviews.

See EPOC resources, available at http://epoc.cochrane.org/epoc-resources-review-authors, in particular:

- *Data collection template*
- *Data collection checklist*
- *EPOC Methods Paper: Including Interrupted Times Series (ITS) designs in an EPOC review*

The EPOC Review Group specifies that, to be included in a systematic review, studies of the ITS design must meet two key criteria. These are

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

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

If ITS studies meet these criteria, they are eligible (at least on study design criteria) for inclusion in a systematic review and so need to be systematically assessed for their risk of bias.

The EPOC guidelines for risk of bias assessment of ITS studies are available on their website.

Note that the scope of EPOC focuses more on healthcare structures and organisational aspects than does the scope of the Consumers and Communication Review Group. Review authors should bear this in mind, and will need to **adapt** the EPOC guidance to suit the needs of their specific review question.

Adapted from Cochrane Handbook Table 8.5.a: The Cochrane Collaboration's tool for assessing risk of bias; adapted using EPOC's criteria for ITS studies and CCRG input

Domain	Review authors' judgement	Description
Was the intervention independent of other changes?	***High risk*** ***Unclear*** ***Low risk***	**Score "Low risk" if there are compelling arguments that the intervention occurred independently of other changes over time and the outcome was not influenced by other confounding variables/historic events during study period. *If Events/variables identified, note what they are.*** **Score "High Risk" if reported that intervention was not independent of other changes in time.**
Was the shape of the intervention effect pre-specified?	***High Risk*** ***Unclear*** ***Low Risk***	**Score "Low Risk" if point of analysis is the point of intervention OR a rational explanation for the shape of intervention effect was given by the author(s). Where appropriate, this should include an explanation if the point of analysis is NOT the point of intervention;** **Score "High Risk" if it is clear that the condition above is not met**
Was the intervention unlikely to affect data	***High Risk***	**Score "Low Risk" if reported that intervention itself was unlikely to affect**

collection?	***Unclear*** ***Low Risk***	**data collection (for example, sources and methods of data collection were the same before and after the intervention);** **Score "High Risk" if the intervention itself was likely to affect data collection (for example, any change in source or method of data collection reported).**
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*	***High risk*** ***Unclear*** ***Low risk***	**Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.**
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).******	***High risk*** ***Unclear*** ***Low risk***	**Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.**
Were incomplete outcome data adequately addressed? Assessments should be made for each main outcome (or class of outcomes). ***	***High Risk*** ***Unclear*** ***Low Risk***	**Score "Low Risk" if missing outcome measures were unlikely to bias the results (e.g. the proportion of missing data was similar in the pre- and post-intervention periods or the proportion of missing data was less than the effect size i.e. unlikely to overturn the study result). Score "No" if missing outcome data was likely to bias the results. Score "Unclear" if not specified in the paper (Do not assume 100% follow up unless stated explicitly).**
Are reports of the study free of suggestion of selective outcome reporting? *Assessments should be made for each main outcome (or class of outcomes).*	***High Risk*** ***Unclear*** ***Low Risk***	**State how the possibility of selective outcome reporting was examined by the review authors, and what was found.**
Was the study free from other risks of bias?	***High Risk*** ***Unclear*** ***Low Risk***	**State any important concerns about bias not addressed in the other domains in the tool.** **If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.** **Score "Low Risk" if there is no evidence of other risk of biases.** **e.g. should consider if seasonality is an issue (i.e. if January to June comprises the pre-intervention period and july to December the post, could the "seasons' have caused a spurious effect).**

*** If some primary outcomes were assessed blindly or affected by missing data and others were not, each primary outcome can be scored separately.

2.6 QUALITY ASSESSMENT: QUALITATIVE STUDIES

- *See Chapter 20 of the Cochrane Handbook for Systematic Reviews of Interventions*
- *NHS Centre for Reviews and Dissemination (2009) Systematic reviews: CRD's guidance for undertaking reviews in health care, available at http://www.york.ac.uk/inst/crd/index_guidance.htm, Chapter 6: Incorporating qualitative evidence in or alongside effectiveness reviews.*

Sometimes it is both appropriate and useful to include qualitative data as part of a systematic review. It is now becoming more common to include data from qualitative research alongside quantitative data in systematic reviews of effectiveness. Qualitative data can, for example, help to provide depth and explanation for observed outcomes; help to decide whether the intervention of interest is suitable for a particular target population; help to examine factors that might have influenced the results if the effect of the intervention is not in the expected direction; and so on. Such contributions can add meaning to the results of a systematic review and can help to address some of the questions which are likely to be important to the users of systematic reviews (for details, see ***Study Design Guide***, *Quantitative and qualitative designs*).

Note that inclusion of qualitative data in the analysis section of a review must be clearly justified by the authors and is subject to editorial approval.

Just as it is desirable to adopt a structured approach to the quality assessment of quantitative research, clear and transparent approaches to the assessment of qualitative research also need to be adopted. Several frameworks for assessing the quality of qualitative research now exist[22]

Like quantitative research, qualitative research needs to be transparently and systematically assessed. There is growing consensus on appropriate hierarchies of evidence arising from qualitative research and on criteria against which to assess qualitative data.

As the conduct and design of qualitative research is also fundamentally different to quantitative research, in that qualitative research does not attempt to 'answer' pre-formulated research questions, there is also little consensus on when quality assessment of qualitative research should occur (that is, whether it should occur prior to or during the synthesis of data

[22] For example, the Critical Appraisal Skills Program (CASP) tool for qualitative studies.

from included studies). Despite these differences, there is agreement that the aim of quality assessment of qualitative research should be to adopt a structured and transparent approach.

The Cochrane Qualitative Research Methods Group and other groups worldwide are in the process of developing standards for the assessment and reporting of qualitative data. Guidelines and recommendations for review authors will be developed and this document updated when they become available.

Authors must contact the editorial base of the Consumers and Communication Group directly if they wish to incorporate qualitative data into their review.

2.7 HOW TO REPORT THE RISK OF BIAS FOR INCLUDED STUDIES: TIPS FOR REVIEW AUTHORS

2.7.1 GENERAL POINTS ON REPORTING RISK OF BIAS

☐ *See Cochrane Handbook for Systematic Reviews of Interventions, Section 8*

- *See Systematic Reviews; CRD's guidance for undertaking reviews in health care; 1.3.5.1 Data synthesis (http://www.york.ac.uk/inst/crd/index_guidance.htm)*

A structured and explicit approach to assessing and reporting the risk of bias of included studies is an important component of a systematic review. While this information must be included in the review (for example, as a component of the 'Table of Included Studies' or as a separate table or figure generated by RevMan), this information is *not* all that is needed.

The text of the systematic review should attempt to integrate the risk of bias assessment by using these assessments to interpret the findings of the studies, and to draw conclusions.

We encourage review authors to **integrate** the methodological assessment of included studies into the wider aims of the review. This is important, as the point of performing risk of bias assessment is to come to an overall decision about how believable the evidence you have collected and analysed is, how it might affect the findings of the review, and how applicable it might be to the real world. This type of synthesis therefore involves considering both the internal and external validity of the studies included in the review.

In terms of internal validity, as well as discussing the individual risk of bias items considered for different studies, you might attempt to draw this together by considering some of the following, or similar, types of questions:

- How confident are you that the results of the included studies reflect what is true?
- Are there limitations in the design or execution of the studies that might render the results less believable? How severe are these limitations?
- Were there aspects of the studies that were particularly well designed and so increase your confidence in the results?

In terms of the external validity or generalisability of the studies, you might wish to consider and discuss some of the following types of questions when considering the body of evidence you have gathered and appraised as a whole:

- What works for whom? Which groups of the population do the results apply to? (e.g. Was the study only of children? Of adults? Of people with a particular disease but no comorbidities?)
- Has the intervention been evaluated in a particular setting only? (e.g. Was the intervention delivered in a hospital setting only? Are the results likely to be applicable to the community setting?)
- Do some versions of the intervention work better or best? (e.g. Is the intensity or frequency of delivery of the intervention likely to affect results? Are some versions ineffective? Are some versions effective only in some people?)

2.7.1.1 SUMMARY: KEY POINTS FOR REVIEW AUTHORS

We encourage authors to attempt to integrate their risk of bias assessment of individual studies into a narrative synthesis. This might include some or all of the following:

- **Narratively describe** and comment upon the risk of bias issues associated with the included studies for each of the rated items. Some issues to consider when trying to formulate an overall description of the body of evidence gathered in the review might include the following:
  - Did most of the included trials report adequate detail on their methods of sequence generation (or other allocation methods)?
  - Have you had to assume that they used adequate methods due to lack of information?
  - Did many studies adequately conceal allocation, and what methods did they use? How do the results of these studies compare to those that did not adequately conceal allocation?
  - Was blinding done adequately (for participants, providers, outcome assessors and analysts; for all outcome measures)?; or was it not possible to blind all of those involved in the study?
  - Was follow-up of participants throughout the study period described; and was it adequate?

☐ Draw some general comments on the **overall risk of bias** of studies included in the review. The purpose of doing so is to help to set the scene for the Results and Discussion sections of the review. As the results of the studies need to be interpreted in the light of the risk of bias of the studies, this is a critical part of your review. Some of the things you may wish to consider might include:

  - Are there particular limitations in the design, execution or reporting of the included studies that render the results less believable?
  - How severe are these limitations?

- Were there aspects of the included studies that were particularly well designed and so increase your confidence in the results?

It may be appropriate at this point, for example, to highlight any particular strengths and/or limitations of the included studies. The main focus should be to discuss and attempt to narratively synthesise the issues associated with the risk of bias of the included studies.

☐ For reviews where studies other than RCTs have also been included, we suggest that studies of different designs be discussed as **separate sections** in terms of their risk of bias and each of their respective strengths and weaknesses.

- For example, it might be appropriate to discuss the risk of bias of RCTs and non-randomised controlled studies separately: as their designs are different, the discussion will need to deal with different aspects relating to design. Once this is done, it may then be appropriate to contrast the different types of studies, and to draw comparisons between the different categories.
- It can be especially interesting, for example, to compare the findings of studies of different design and varying levels of quality to comment on the consistency of the findings of studies included in the review. For example, do non-randomised studies report the same direction and size of results as RCTs? Do studies with a low risk of bias report the same direction and size of results as studies with a high risk of bias?

APPENDIX A –ADDITIONAL SOURCES OF INFORMATION

(SEE ALSO APPENDIX C IN THE STUDY DESIGN GUIDE)

ALLOCATION CONCEALMENT AND SUBVERTING RANDOMISATION

Schulz KF. Subverting randomisation in controlled trials. *JAMA* 1995; 274(18): 1456-8.

☐ This paper discusses the necessity of adequate allocation concealment in order to adequately randomise participants to groups in a RCT, giving examples, recommendations and further references on adequate approaches.

ATTRITION BIAS

Tierney JF. and Stewart LA. Theory and Methods: Investigating patient exclusion bias in meta-analysis. *International Journal of Epidemiology* 2005; 34 (1): 79-87.

Juni P. and Egger M. Commentary: Empirical evidence of attrition bias in clinical trials. *International Journal of Epidemiology* 2005; 34 (1): 87-8.

☐ This study (and commentary) analyses the effects of participant exclusion from trials (post-randomisation), and directly analyses the effects of attrition bias upon trial effect estimates.

☐ It includes a useful glossary of terms.

☐ This study demonstrates that pooled analysis of trials that excluded patients post-randomisation may be prone to bias. The authors conclude that analysis (in trials, meta-analyses and systematic reviews) should be based on all participants randomised at the start of the trial.

BASELINE PERFORMANCE MEASURES

Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *EPOC Methods Paper: Issues Related to Baseline Measures of Performance* (http://epoc.cochrane.org/epoc-resources)

- This paper gives further detail on issues associated with baseline performance measures in primary studies. It includes a rationale for measuring baseline values, the problem of imbalances at baseline, and a worked example of baseline imbalance.
- It also provides guidance on ways of adjusting analysis and data extraction to include baseline values; and how to incorporate information about baseline values into a review.

BLINDING

Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, Bhandari M and Guyatt GH. Physician interpretations and textbook definitions of blinding terminology in randomised controlled trials. *JAMA* 2001; 285(15): 2000-5.

☐ As physicians commonly use study blinding as a quality assessment criteria for studies they read, this study examined the definitions and interpretations of blinding used by physicians and by textbooks.

☐ The study reports little consensus about the use of blinding and associated terms (single, double, etc); and discusses the findings in terms of implications for the quality of trials.

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.

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

NON-RANDOMISED STUDIES AND QUALITY

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.hta.ac.uk/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.

Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. (2000) Bayesian methods in health technology assessment: a review. *Health Technol Assess* 4(34).

http://www.hta.ac.uk/execsumm/summ438.htm

This paper assesses the association between study (methodological) quality and estimates of effect sizes. In particular, it examines the differences between effect sizes generated by RCTs in comparison with quasi-randomised and observational studies. It discusses differences in effect estimates across various different non-randomised study designs.

☐ This study also recommends that further studies to directly compare the results of RCTs and quasi-randomised and observational studies be performed; and that further instruments be developed to assess study quality and the accuracy of effect estimates in relation to study quality.

PARTICIPANT FLOW AND FOLLOW-UP

Cakir B, Gebski VJ and Keech AC. Flow of participants in randomised studies. *MJA* 2003; 178: 347-9.

☐ Outlines the components of participant flow through a trial that should be included in a trial report and should be assessed when including such as study in a systematic review.

☐ This article draws clear distinctions between participant loss at different stages of a trial, and emphasises the implications of this for the quality of the study and the degree of confidence that should be placed in the results of the study.

PUBLICATION BIAS

Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. *Health Technol Assess* 2000;4(10) http://www.hta.ac.uk/execsumm/summ410.shtml

☐ This paper systematically examines and evaluates studies on methodological issues associated with publication bias. It attempts to identify empirical evidence about the existence and consequences of publication bias and other forms of dissemination-based biases.

PUBLIC HEALTH INTERVENTIONS

Weightman A, Ellis S, Cullum A, Sander L and Turley R. Grading evidence and recommendations for public health interventions: developing and piloting a framework, *Health Development Agency 2005*, http://www.nice.org.uk/niceMedia/docs/grading_evidence.pdf

☐ This paper is an extensive evaluation of different types of research design and their appropriateness for answering questions about the efficacy of public health interventions.

☐ This paper presents the results of the literature and development of the evidence grading framework, including a detailed description of the criteria used to assess studies. It also presents the results of the piloting of the provisional framework developed for grading evidence about public health interventions.

For further discussion of issues associated with public health interventions, also see the Cochrane Health Promotion and Public Health Review Group. Guidelines for *Systematic Reviews of Health Promotion and Public Health Interventions.* http://ph.cochrane.org/review-authors

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.

Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, et al. Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses. *Health Technol Assessment* 1999; 3(12) http://www.hta.ac.uk/execsumm/summ312.shtml

- ☐ This paper examines and discusses the various different elements of quality that are necessary to assess for studies included in systematic reviews.
- ☐ It discusses in detail the different aspects of study quality; which are likely to be most important (ie. to impact the most upon results), and makes recommendations for assessing study quality as part of systematic review conduct.

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.

Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. York: University of York; 2009 [Available from: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf] see Chapter 1.3.4 in particular.

- ☐ These resources give 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.

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 of RCTS, methods of randomisation and different types of RCTs to the assessment of RCT quality, sources of bias and reporting of individual trials.

SEARCHING

Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. (2003) *How important are comprehensive literature searchers and the assessment of trial quality in systematic reviews: empirical study*. HTA Methodology 7(26). http://www.hta.ac.uk/pdfexecs/summ701.pdf

- ☐ This paper examines the characteristics of trials that are difficult to locate by searching. It also examines the effects of including only easily accessible studies, compared with more difficult to locate studies (that require comprehensive search techniques), on reported treatment effects and on pooled treatment effects (from meta-analysis).
- ☐ It concludes that trials that are difficult to locate tend to be smaller and of poorer methodological quality than those that are easily accessible; and that the inclusion or exclusion of trials of low methodological quality has a substantial impact on the results and the conclusions reached by systematic reviews and meta-analyses.

See also Jadad 1998.

CONSORT STATEMENT

Moher D., Schulz K.F. and Altman D.G. The CONSORT statement: revised recommendations for improving the quality of report of parallel-group randomised trials. *The Lancet* 2001; 357 (9263): 1191-8.

- ☐ This paper outlines the CONSORT checklist, which has been developed to improve the reporting of RCTs, so that readers will be better able to determine why the study was undertaken, and how it was conducted and analysed.
- ☐ A copy of the most recent version of the CONSORT checklist (2010) is available at http://www.consort-statement.org/index.aspx?o=1031

TREND STATEMENT

Des Jarlais D.C., Lyles C., Crepaz N. and the TREND Group. Improving the reporting quality of nonrandomised evaluations of behavioural and public health interventions: The TREND statement. *American Journal of Public Health* 2004; 94(3): 361-6.

- This paper notes that non randomised trials and other studies should be included in systematic reviews, as excluding all data from studies other than RCTs will bias the evidence base towards those interventions which are 'simpler' to evaluate (ie. they can be evaluated using an RCT). This forms the basis for the development of the TREND statement, a consensus statement for the reporting of nonrandomised intervention studies.
- ☐ The authors describe the TREND checklist and its development (as companion to the CONSORT statement for RCTs), which provides guidelines for the transparent reporting of studies, including aspects of intervention and comparisons, the research design, as well as methods of adjusting for possible sources of bias.
- A copy of the TREND checklist is available at http://www.cdc.gov/trendstatement

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 Health Promotion and Public Health Field. *Guidelines for Systematic Reviews of Health Promotion and Public Health Interventions.* http://ph.cochrane.org/review-authors

Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *The data collection checklist*. http://epoc.cochrane.org/epoc-resources

Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *The data collection template.* http://epoc.cochrane.org/epoc-resources

Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *EPOC Methods Paper: Including Interrupted Times Series (ITS) designs in an EPOC review.* http://epoc.cochrane.org/epoc-resources

Cochrane Effective Practice and Organisation of Care (EPOC) Review Group. *EPOC Methods Paper: Issues Related to Baseline Measures of Performance*. http://epoc.cochrane.org/epoc-resources

Systematic reviews: CRD's guidance for undertaking reviews in health care. Centre for Reviews and Dissemination. York: University of York; 2009 [Available from: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf]

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.hta.ac.uk/execsumm/summ727.htm

Egger M, Jüni P, Bartlett C, Holenstein F, Sterne J. (1999) How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess* 7(1). http://www.hta.ac.uk/execsumm/summ701.htm

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

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

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

Juni P, Witschi A, Bloch R, Egger M. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA;* 282: 1054-60.

Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, et al. (1999) Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses. *Health Technol Assessment* 3(12).

Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. *Health Technol Assess* 2000;4(10) http://www.york.ac.uk/inst/crd/projects/publication_bias.htm

Spencer L, Ritchie J, Lewis J, Dillon L. *Quality in Qualitative Evaluation: A framework for assessing research evidence*. Government Chief Social researcher's Office. Crown

Copyright, 2003. http://www.civilservice.gov.uk/wp-content/uploads/2011/09/a_quality_framework_tcm6-38740.pdf

Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. (2000) Bayesian methods in health technology assessment: a review. *Health Technol Assess* 4(34). http://www.hta.ac.uk/execsumm/summ438.htm

Torgerson DJ. and Sibbald B. What is a patient preference trial? *BMJ* 1998; 316: 360.