

Module 8: Selecting studies for your review

This module will discuss how you decide which of the studies your searching found should be included in the review. We will look at ways you can do this as reliably as possible and with the minimum of bias.

Learning objectives

- Understand that selecting studies is a multi-stage process
- Be able to design an eligibility form
- Be aware of the need to pilot inclusion criteria
- Be aware of ways to improve the reliability of study selection
- Be aware of the existence of duplicate publications
- Be aware of ways to obtain sufficient information to make a reliable decision on study eligibility

Relevant section of the *Cochrane Handbook for Systematic Reviews of Intervention*

- Chapter 7: Selecting studies and collecting data

Other relevant material

- BMJ paper by Martin Tramèr and colleagues ([Tramèr MR et al BMJ 1997; 315:635-640](#)).

Where does this go in a Cochrane review?

- Early on in the Methods section of protocol and review



Read [Section 7.2](#) of the *Cochrane Handbook for Systematic Reviews of Interventions*

To select the relevant studies from our searches, we need to sift through them

This module relates to Section 7.2 of the *Cochrane Handbook for Systematic Reviews of Intervention* and you should read this now.

Sift and sift again

We saw in the previous module, *Searching for studies*, that in a systematic review we attempt to find every study that has ever been done addressing our question. We try to do this by running *sensitive* searches. Inevitably, when we do this, we find lots of reports of studies that *could* be relevant, and we then have to decide which ones *are* relevant to our question.

It's important to remember that our decisions about which studies to include should be based on the *design* of those studies, and *not* the results. If we allow ourselves to be swayed by the results of the studies, we might exclude a perfectly eligible study because we don't like or believe the results.

If you have already run some searches, you will know that these sensitive searches usually turn up hundreds or thousands of records. Most of these will have come from an electronic database, and you'll only have limited information about the study, like the example below.

ID:CN-00240563
TI:Antibiotic prophylaxis of wound infections in skin surgery [see comments].
AU:Bencini, P. L., Galimberti, M., Signorini, M., and Crosti, C.
SO:Archives of Dermatology
YR:1991
VL:127
NO:9
PG:1357-60
AB:A controlled prospective study of 2165 outpatients undergoing skin surgery was performed to evaluate the utility and the effects of several antibiotic schedules for prophylaxis of wound infections. The patients were divided into four groups. Twenty-three of the 541 group A patients, given no antibiotics, had wound infections. Eight of the 542 group B patients, given systemic antibiotics from immediately after surgery until the third day, had wound infections. Four of the 540 group C patients, treated only with local sterile antibiotic powder sprinkled into the wound during surgery, had wound infections develop, and only one infection occurred in the 542 group D patients given systemic antibiotics from 2 days before surgery until the second day after surgery. This last schedule was the best for prophylaxis of wound infections in contamination-prone regions. Local antibiotic administration is a simple method for prevention of infections in routine skin surgery.
KY:Administration, Cutaneous; Adolescence; Adult; Bandages; Cefazolin; Ad [Administration & Dosage]; *Cefazolin; Tu [Therapeutic Use]; Female; Human; Incidence; Injections, Intramuscular; Italy; Ep [Epidemiology]; Male; Middle Age; Powders; Premedication; Prospective Studies; Skin; Pa [Pathology]; *Skin; Su [Surgery]; Skin Neoplasms; Pa [Pathology]; *Skin Neoplasms; Su [Surgery]; Staphylococcal Infections; Ep [Epidemiology]; Surgical Wound Infection; Ep [Epidemiology]; *Surgical Wound Infection; Pc [Prevention & Control]; Time Factors
DE:RCT.
CC:SR-HANDSRCH, SR-SKIN

Comment [v1]: http://www.mrc-bsu.cam.ac.uk/cochrane/handbook/index.htm#chapter_7/7_2_selecting_studies.htm



Activity: write a list of the further information you think you might need to tell whether this study is relevant

It isn't usually possible to be certain from this sort of record whether it will turn out to be included in the review or not. There usually simply isn't enough information to be quite sure. Try to list the extra information you might want to help you decide if this was a randomised controlled trial relevant to a review of antibiotic use for preventing wound infection in skin surgery.

Here are a few suggestions

- Was the study randomised and how (ie how were the patients divided into different groups)?
- What type of surgery was performed?
- How was resultant infection determined?
- Were all patients followed up
- Did they have any other type of intervention or co-morbidity that may have influenced the results?

What *is* normally possible by looking at these records is to tell whether the study *might* be relevant or is *very unlikely* to be relevant. You can then order the full paper copy of the ones that you think might be relevant. This saves you getting full paper copies of hundreds of articles that turn out to be irrelevant.

The first sift – pre-screening - is to decide which studies to retrieve in full.

Once you have these hard copies of the probably relevant studies, you can get on with comparing them with your review's inclusion criteria. This second stage is where you make definite decisions about whether studies are included or excluded from the review.

The second sift – selection - is to look again at these studies and decide which are to be included in your review

Practicalities of sifting

A quick note on language

The information you're using for pre-screening and selection might not be written in your first language. When looking at studies in a language you cannot read, you are likely to need help. How much help you need depends on whether you can identify certain key words to tell you whether a study might be relevant. If not, you will probably need to find someone who is familiar with that language to help. They may just be able to read an abstract and tell you whether it is worth ordering the full article. With a full paper, you may just be able to ask someone to read it and tell you whether it is eligible, or you may need to get certain parts, or all, of the paper translated. You may know people who speak a variety of languages, or your review group may be able to help you find a translator if necessary.

Human error

Both stages of sifting are going to be done by humans. Humans make mistakes and people doing reviews are no different. When looking at a lot of studies we may simply miss some information and mistakenly include or exclude the study. We all have certain prejudices or biases which might make us more or less likely to choose a particular study. Part of the review process is to try to minimise these mistakes and biases.

An important point is that if you exclude a study in the initial sift, it will rarely get another chance to be included. So reviewers usually give studies the benefit of the doubt at this early stage, and go on to obtain the full report.

Should you collect outcome data at the same time as eligibility information?

It might seem that since you are reading these papers anyway for making eligibility decisions, why not collect outcome data at the same time? Why read all the papers again later?

There are two main arguments against this. The first is simple – if you decide to exclude a study, you will have extracted its data unnecessarily, and will have wasted time and effort. We mentioned the second reason earlier on – you should try to avoid letting the results of studies sway your decision. If you look too hard at the results too soon, you might find it difficult not to be swayed.

Reducing mistakes

Having more than one person making decisions can reduce mistakes

How can we reduce mistakes? A common sense way to try and do this is to have more than one person make the decisions. Then, if one person makes a mistake, there will be one or more others who might get it right. In one review, pairs of reviewers independently sifted about 11,000 records. It took the reviewers between 11 and 28 hours to go through all these records. The reviewers concluded that having two people check each record was worthwhile. The reference for this study is listed at the start of the module under 'Other relevant material'. Therefore, assuming that your review has many fewer records than that large review, the time needed for this part of your review might be less than you had thought. So you should think about having more than one of you do this stage.

If you can't find one or more other people willing to commit the necessary amount of time, there are a couple of compromises

- Have a second person look at a sample of the records
- Have two reviewers each look at separate sets of records, and overlap the sets they look at so that they both do some of the records (for example one looks at years 1966-1988 and the other looks at 1986-2000)

If you find, through either of these ways of working, that you disagree on quite a few records, you may need to look again at the criteria you have set for the sifting of the records of studies. There are no rules about the level of disagreement above which you need to go back and redo the work. It should be low, as mistakes here might mean that you leave out relevant studies, which are unlikely ever to be looked at again.

If you do think you need to go back and recheck this sifting process, the most obvious option is to resift all the records, with two or more reviewers looking at all of them. Another option is for one of the reviewers to sift through all the records they did not look at first time.

In most circumstances, use at least two reviewers for final inclusion/exclusion decisions

For the second stage of the sifting, there will be fewer reports and it seems common sense that at least two reviewers look at each of them. Of course, there may be occasions when you think the decisions are likely to be so straightforward that this isn't necessary. Whatever you decide to do, report it in the Methods section of your review.

Reducing bias

All of us have prejudices that might affect our judgements about whether studies should be included or excluded. Experts may have pre-formed opinions which might affect their assessments of the relevance and validity of studies. On the other hand, it's difficult to make judgements if you know nothing at all about a topic. Other people might have opinions about the value of research published in particular journals, or research carried out in particular institutions.

Using two reviewers may also reduce the effect of personal biases

One way to minimise the effect of these personal biases is to have reviewers of different backgrounds making judgements about studies, for example an expert and a non-expert. For this to work, however, both reviewers need to be willing to accept that they may have biases and to listen to the other reviewer's views!

Blinding reviewers to details of the papers doesn't seem to make much difference

One suggestion for reducing bias is that we should remove as much identifying information from the papers as possible, such as the name of the journal, authors, institutions, etc. A randomised trial has been done to assess this approach (you can find the reference as Berlin 1997 in the *Cochrane Handbook for Systematic Reviews of Interventions*, section 7.2.4). The trialists repeated several meta-analyses, once with reviewers who saw the unblinded papers of the included studies, and once with reviewers blinded to many of the identifying details in the papers. It took many hours to remove as much identifying detail as possible, and it appeared to make little difference to the overall estimate of effect for the meta-analysis. So, unless someone refutes these results, it looks as though we don't normally need to 'blind' the papers, unless we think there are important reasons why not blinding the papers would be a problem for a particular review.

Another interpretation of the above study is that we are just not very good at concealing the identity of the papers. For instance, experts will often have read the paper before and may recognise the text.

Getting hold of extra information

You will often find that, even when you have the full report of a study, you don't have all the information you need. Maybe it doesn't tell you how the study was randomised, or maybe it doesn't tell you clearly who the participants were. This may happen later on as well, when you need outcome data.

You may be able to make assumptions about a study that everyone would agree are reasonable. For instance, if a report of a study tells you that it recruited women aged over 60, it is reasonable to assume that the women were all post-menopausal. There isn't really a need to ask the authors to clarify this. If you cannot make assumptions like this, you may need to create categories for features of trials called 'unclear' – it was 'unclear' whether the patients were masked to the treatment, etc.

You may need to contact the study authors if you need more information

The other way to deal with the situation is to try to get hold of more information. To do this you either need to find other reports of the same study, which contain more information, or try to contact the people who conducted the study. Make sure you've looked carefully for other reports – they will probably have come up in your search anyway, but you could do a quick extra search for other publications by the same author.

Resolving disagreements

Just as surely as we all make mistakes, people will disagree about the way they see things. So, when you have two people making decisions about including and excluding studies, there will be times when you come to a different decision about a study.

Some people like to measure and report how often this occurs as part of their review, to give readers an idea of how difficult the decisions were. The Cochrane Collaboration doesn't insist that this is done.

To resolve differences, you first need to work out why you came to different judgements. It may just be that one person missed a vital sentence. It may be that there wasn't enough information – in this case you might need to put that study aside and wait until you've got the extra information you need. Or maybe one reviewer has special knowledge about a study that isn't in the report – they may even have been one of the authors. If you really can't agree, ask another person to help you resolve this. You might choose another content expert or a methodologist, depending on the area of your disagreement.

It is important to plan how you are going to resolve disagreements early on in planning your review and to ensure your review team are happy with the planned process. This will help resolve any possible later disagreement.

Getting your work into RevMan

Once you've decided which studies are included, excluded, awaiting classification (the ones where you needed more information), or ongoing (not yet completed), you'll need to put them in the appropriate category in RevMan. This is explained in RevMan 'Help' under the heading 'Studies and References'.

*Look out for
duplicate
publications*

Pitfalls and problems - duplicate publications

Some studies result in more than one publication. Authors may publish the methods of the study, present preliminary data at a conference resulting in an abstract, then publish some results, and later publish longer term follow-up. There's nothing wrong with this, as long as you can tell it's all about the same study.

Sometimes, studies may be published more than once for other reasons – more publicity, more papers on the author's CV, or to allow different authors to be first authors. Sometimes, it's not easy to tell whether they are reports of the same study. This can cause problems for reviewers because we might count a study more than once and so give extra weight to it in our review.

This phenomenon was studied by Martin Tramèr and colleagues (Tramèr MR et al BMJ 1997;315:635-640.)

They found that 17% of reports of trials of a drug were duplicates and 28% of the data were duplicated. If this had not been spotted and a review had been done counting studies more than once, the drug would have appeared more effective than it actually was. So, we need to be alert to the possibility of duplicates. Look for

- Same authors in different orders
- Similar study inclusion and exclusion criteria
- Many reports of a study done in the same place at the same time
- Results tables that look familiar

Now you understand the process of selecting studies for your review, return to Module 7 and we will work on designing this Section of your data collection form