

Chapter 5: Defining the review question and developing criteria for including studies

Editors: Denise O'Connor, Sally Green and Julian PT Higgins.

Copyright © 2008 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd under “The Cochrane Book Series” Imprint.

This extract is made available solely for use in the authoring, editing or refereeing of Cochrane reviews, or for training in these processes by representatives of formal entities of The Cochrane Collaboration. Other than for the purposes just stated, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the copyright holders.

Permission to translate part or all of this document must be obtained from the publishers.

This extract is from *Handbook* version 5.0.1. For guidance on how to cite it, see Section 5.8. The material is also published in Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions* (ISBN 978-0470057964) by John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, Telephone (+44) 1243 779777; Email (for orders and customer service enquiries): cs-books@wiley.co.uk. Visit their Home Page on www.wiley.com.

Key Points

- A clearly defined, focused review begins with a well framed question. In Cochrane reviews, questions are stated broadly as review ‘Objectives’, and specified in detail as ‘Criteria for considering studies for this review’.
- The review question should specify the types of population (participants), types of interventions (and comparisons), and the types of outcomes that are of interest. The acronym PICO (**P**articipants, **I**nterventions, **C**omparisons and **O**utcomes) helps to serve as a reminder of these. These components of the question, with the additional specification of types of study that will be included, form the basis of the pre-specified eligibility criteria for the review.
- Cochrane reviews should include all outcomes that are likely to be meaningful, and not include trivial outcomes. Primary outcomes should be limited to a very small number and include adverse as well as beneficial outcomes.
- Cochrane reviews can focus on broad questions, or be more narrowly defined. There are advantages and disadvantages of each.

5.1 Questions and eligibility criteria

5.1.1 Rationale for well-formulated questions

As with any research, the first and most important decision in preparing a systematic review is to determine its focus. This is best done by clearly framing the questions the review seeks to answer. Well-formulated questions will guide many aspects of the review process, including determining eligibility criteria, searching for studies, collecting data from included studies, and presenting findings (Jackson 1980, Cooper 1984, Hedges 1994). In Cochrane reviews, questions are stated broadly as review ‘Objectives’, and specified in detail as ‘Criteria for considering studies for this review’. As well as focussing review conduct, the contents of these sections are used by readers in their initial assessments of whether the review is likely to be directly relevant to the issues they face.

A statement of the review’s objectives should begin with a precise statement of the primary objective, ideally in a single sentence. Where possible the style should be of the form ‘To assess the effects of [*intervention or comparison*] for [*health problem*] in [*types of people, disease or problem and setting if specified*]’. This might be followed by one or more secondary objectives, for example relating to different participant groups, different comparisons of interventions or different outcome measures.

The detailed specification of the review question requires consideration of several key components (Richardson 1995, Counsell 1997). The ‘clinical question’ should specify the types of population (participants), types of interventions (and comparisons), and the types of outcomes that are of interest. The acronym PICO (**P**articipants, **I**nterventions, **C**omparisons and **O**utcomes) helps to serve as a reminder of these. Equal emphasis in addressing each PICO component is not necessary. For example, a review might concentrate on competing interventions for a particular stage of breast cancer, with stage and severity of the disease being defined very precisely; or alternately focus on a particular drug for any stage of breast cancer, with the treatment formulation being defined very precisely.

5.1.2 Eligibility criteria

One of the features that distinguish a systematic review from a narrative review is the pre-specification of criteria for including and excluding studies in the review (eligibility criteria). Eligibility criteria are a combination of aspects of the clinical question plus specification of the types of studies that have addressed these questions. The participants, interventions and comparisons in the clinical question usually translate directly into eligibility criteria for the review. Outcomes usually are not part of the criteria for including studies: a Cochrane review would typically seek all rigorous studies (e.g. randomized trials) of a particular comparison of interventions in a particular population of participants, irrespective of the outcomes measured or reported. However, some reviews do legitimately restrict eligibility to specific outcomes. For example, the same intervention may be studied in the same population for different purposes (e.g. hormone replacement therapy, or aspirin); or a review may address specifically the adverse effects of an intervention used for several conditions (see Chapter 14, Section 14.2.3).

In Sections 5.2 to 5.5 we provide an overview of the key components of questions and study types with examples of useful issues to consider for each component and the subsequent development of eligibility criteria to guide inclusion of studies.

5.2 Defining types of participants: which people and populations?

The criteria for considering types of people included in studies in a review should be sufficiently broad to encompass the likely diversity of studies, but sufficiently narrow to ensure that a meaningful answer

can be obtained when studies are considered in aggregate. It is often helpful to consider the types of people that are of interest in two steps. First, the diseases or conditions of interest should be defined using explicit criteria for establishing their presence or not. Criteria that will force the unnecessary exclusion of studies should be avoided. For example, diagnostic criteria that were developed more recently – which may be viewed as the current gold standard for diagnosing the condition of interest – will not have been used in earlier studies. Expensive or recent diagnostic tests may not be available in many countries or settings.

Second, the broad population and setting of interest should be defined. This involves deciding whether a special population group is of interest, determined by factors such as age, sex, race, educational status or the presence of a particular condition such as angina or shortness of breath. Interest may focus on a particular setting such as a community, hospital, nursing home, chronic care institution, or outpatient setting. Box 5.2.a outlines some factors to consider when developing criteria for the ‘Types of participants’.

The types of participants of interest usually determine directly the participant-related eligibility criteria for including studies. However, pre-specification of rules for dealing with studies that only partially address the population of interest can be challenging. For example, if interest focuses on children, a cut-point such as 16 years old might be desirable, but does not determine a strategy for dealing with studies with participants aged from 12 to 18. Use of arbitrary rules (such as “more than 80% of the participants are under 16”) will not be practical if detailed information is not available from the study. A phrase such as “the majority of participants are under 16” may be sufficient. Although there is a risk of review authors’ biases affecting *post hoc* inclusion decisions, this may be outweighed by a common sense strategy in which eligibility decisions keep faith with the objectives of the review rather than with arbitrary rules. Difficult decisions should be documented in the review, and sensitivity analyses can assess the impact of these decisions on the review’s findings (see Chapter 9, Section 9.7).

Any restrictions with respect to specific population characteristics or settings should be based on a sound rationale. It is important that Cochrane reviews are globally relevant, so justification for the exclusion of studies based on population characteristics should be explained in the review. For example, focusing a review of the effectiveness of mammographic screening on women between 40 and 50 years old may be justified on the basis of biological plausibility, previously published systematic reviews and existing controversy. On the other hand, focusing a review on a particular subgroup of people on the basis of their age, sex or ethnicity simply because of personal interests when there is no underlying biologic or sociological justification for doing so should be avoided. When it is uncertain whether there are important differences in effects among various subgroups of people, it may be best to include all of the relevant subgroups and then test for important and plausible differences in effect in the analysis (see Chapter 9, Section 9.6). This should be planned *a priori*, stated as a secondary objective and not driven by the availability of data.

Box 5.2.a: Factors to consider when developing criteria for ‘Types of participants’

- How is the disease/condition defined?
- What are the most important characteristics that describe these people (participants)?
- Are there any relevant demographic factors (e.g. age, sex, ethnicity)?
- What is the setting (e.g. hospital, community etc)?
- Who should make the diagnosis?
- Are there other types of people who should be excluded from the review (because they are likely to react to the intervention in a different way)?

- How will studies involving only a subset of relevant participants be handled?

5.3 Defining types of interventions: which comparisons to make?

The second key component of a well-formulated question is to specify the interventions of interest and the interventions against which these will be compared (comparisons). In particular, are the interventions to be compared with an inactive control intervention (e.g. placebo, no treatment, standard care, or a waiting list control), or with an active control intervention (e.g. a different variant of the same intervention, a different drug, a different kind of therapy)?

When specifying drug interventions, factors such as the drug preparation, route of administration, dose, duration, and frequency should be considered. For more complex interventions (such as educational or behavioural interventions), the common or core features of the interventions will need to be defined. In general, it is useful to consider exactly what is delivered, at what intensity, how often it is delivered, who delivers it, and whether people involved in delivery of the intervention need to be trained. Review authors should also consider whether variation in the intervention (i.e. based on dosage/intensity, mode of delivery, frequency, duration etc) is so great that it would have substantially different effects on the participants and outcomes of interest, and hence may be important to restrict.

Box 5.3.a outlines some factors to consider when developing criteria for the ‘Types of interventions’ (and comparisons).

Box 5.3.a: Factors to consider when developing criteria for ‘Types of interventions’

- What are the experimental and control (comparator) interventions of interest?
- Does the intervention have variations (e.g. dosage/intensity, mode of delivery, personnel who deliver it, frequency of delivery, duration of delivery, timing of delivery)?
- Are all variations to be included (for example is there a critical dose below which the intervention may not be clinically appropriate)?
- How will trials including only part of the intervention be handled?
- How will trials including the intervention of interest combined with another intervention (co-intervention) be handled?

5.4 Defining types of outcomes: which outcome measures are most important?

5.4.1 Listing relevant outcomes

Although reporting of outcomes should rarely determine eligibility of studies for a review, the third key component of a well-formulated question is the delineation of particular outcomes that are of interest. In general, Cochrane reviews should include all outcomes that are likely to be meaningful to clinicians, patients (consumers), the general public, administrators and policy makers, but should not include outcomes reported in included studies if they are trivial or meaningless to decision makers. Outcomes considered to be meaningful, and therefore addressed in a review, will not necessarily have been reported in individual studies. For example, quality of life is an important outcome, perhaps the most important outcome, for people considering whether or not to use chemotherapy for advanced cancer, even if the available studies are found to report only survival (see Chapter 17). Including all

important outcomes in a review will highlight gaps in the primary research and encourage researchers to address these gaps in future studies.

Outcomes may include survival (mortality), clinical events (e.g. strokes or myocardial infarction), patient-reported outcomes (e.g. symptoms, quality of life), adverse events, burdens (e.g. demands on caregivers, frequency of tests, restrictions on lifestyle) and economic outcomes (e.g. cost and resource use). It is critical that outcomes used to assess adverse effects as well as outcomes used to assess beneficial effects are among those addressed by a review (see Chapter 14). If combinations of outcomes will be considered, these need to be specified. For example, if a study fails to make a distinction between non-fatal and fatal strokes, will these data be included in a meta-analysis if the question specifically relates to stroke death?

Review authors should consider how outcomes may be measured, both in terms of the type of scale likely to be used and the timing of measurement. Outcomes may be measured objectively (e.g. blood pressure, number of strokes) or subjectively as rated by a clinician, patient, or carer (e.g. disability scales). It may be important to specify whether measurement scales have been published or validated. When defining the timing of outcome measurement, authors may consider whether all time frames or only selected time-points will be included in the review. One strategy is to group time-points into pre-specified intervals to represent 'short-term', 'medium-term' and 'long-term' outcomes and to take no more than one of each from each study for any particular outcome. It is important to give the timing of outcome measure considerable thought as it can influence the results of the review (Gøtzsche 2007).

As Cochrane reviews are increasingly included in Overviews of reviews (see Chapter 22), harmonization of outcomes across reviews addressing related questions will facilitate this process. It may be helpful for review authors to consider those measures used in related reviews when defining the type and timing of measurement within their own review. In addition, several clinical areas are developing agreed core sets of outcome measures for use in randomized trials, and consideration of these in defining the detail of measurement of outcomes selected for the review is likely to be helpful.

Various sources can be used to develop a list of relevant outcomes, including the clinical experiences of the review authors, input from consumers and advisory groups (see Chapter 2), and evidence from the literature (including qualitative research about outcomes important to those affected). Further information about the use of qualitative research to inform the formulation of review questions, including types of outcome measures, can be found in Chapter 20.

While all important outcomes should be included in Cochrane reviews, trivial outcomes should not be included. Authors need to avoid overwhelming and potentially misleading readers with data that are of little or no importance. In addition, indirect or surrogate outcome measures, such as laboratory results or radiologic results (e.g. loss of bone mineral content as a surrogate for fractures in hormone replacement therapy), are potentially misleading and should be avoided or interpreted with caution because they may not predict clinically important outcomes accurately. Surrogate outcomes may provide information on how a treatment might work but not whether it actually does work. Many interventions reduce the risk for a surrogate outcome but have no effect or have harmful effects on clinically relevant outcomes, and some interventions have no effect on surrogate measures but improve clinical outcomes.

5.4.2 Prioritizing outcomes: main, primary and secondary outcomes

Main outcomes

Once a full list of relevant outcomes has been compiled for the review, authors should prioritize the outcomes and select the main outcomes of relevance to the review question. The main outcomes are the essential outcomes for decision-making, and are those that would form the basis of a ‘Summary of findings’ table. ‘Summary of findings’ tables provide key information about the amount of evidence for important comparisons and outcomes, the quality of the evidence and the magnitude of effect (see Chapter 11, Section 11.5). There should be no more than seven main outcomes, which should generally not include surrogate or interim outcomes. They should not be chosen on the basis of any anticipated or observed magnitude of effect, or because they are likely to have been addressed in the studies to be reviewed.

Primary outcomes

Primary outcomes for the review should be identified from among the main outcomes. Primary outcomes are the outcomes that would be expected to be analysed should the review identify relevant studies, and conclusions about the effects of the interventions under review will be based largely on these outcomes. There should in general be no more than three primary outcomes and they should include at least one desirable and at least one undesirable outcome (to assess beneficial and adverse effects respectively).

Secondary outcomes

Main outcomes not selected as primary outcomes would be expected to be listed as secondary outcomes. In addition, secondary outcomes may include a limited number of additional outcomes the review intends to address. These may be specific to only some comparisons in the review. For example, laboratory tests and other surrogate measures may not be considered as main outcomes as they are less important than clinical endpoints in informing decisions, but they may be helpful in explaining effect or determining intervention integrity (see Chapter 7, Section 7.3.4).

[Box 5.4.a](#) summarizes the principal factors to consider when developing criteria for the ‘Types of outcomes’.

Box 5.4.a: Factors to consider when developing criteria for ‘Types of outcomes’

- Main outcomes, for inclusion in the ‘Summary of findings’ table, are those that are essential for decision-making, and should usually have an emphasis on patient-important outcomes.
- Primary outcomes are the two or three outcomes from among the main outcomes that the review would be likely to be able to address if sufficient studies are identified, in order to reach a conclusion about the effects (beneficial and adverse) of the intervention(s).
- Secondary outcomes include the remaining main outcomes (other than primary outcomes) plus additional outcomes useful for explaining effects.
- Ensure that outcomes cover potential as well as actual adverse effects.
- Consider outcomes relevant to all potential decision makers, including economic data.
- Consider the type and timing of outcome measurements.

5.4.3 Adverse outcomes

It is important that Cochrane reviews include information about the undesirable as well as desirable outcomes of the interventions examined. Review authors should consider carefully how they will include data on undesirable outcomes in their review, and at least one undesirable outcome should be

defined as a primary outcome measure. Assessment of adverse effects is discussed in detail in Chapter 14.

5.4.4 Economic data

Decision makers need to consider the economic aspects of an intervention, such as whether its adoption will lead to a more efficient use of resources. Economic data such as resource use, costs or cost-effectiveness (or a combination of these) may therefore be included as outcomes in a review. It is useful to break down measures of resource use and costs to the level of specific items or categories. It is helpful to consider an international perspective in the discussion of costs. Economics issues are discussed in detail in Chapter 15.

5.5 Defining types of study

Certain study designs are more appropriate than others for answering particular questions. Authors should consider a priori what study designs are likely to provide reliable data with which to address the objectives of their review.

Because Cochrane reviews address questions about the effects of health care, they focus primarily on randomized trials. Randomization is the only way to prevent systematic differences between baseline characteristics of participants in different intervention groups in terms of both known and unknown (or unmeasured) confounders (see Chapter 8). For clinical interventions, deciding who receives an intervention and who does not is influenced by many factors, including prognostic factors. Empirical evidence suggests that, on average, non-randomized studies produce effect estimates that indicate more extreme benefits of the effects of health care than randomized trials. However, the extent, and even the direction, of the bias is difficult to predict. These issues are discussed at length in Chapter 13, which provides guidance on when it might be appropriate to include non-randomized studies in a Cochrane review.

A practical consideration also motivates the restriction of many Cochrane reviews to randomized trials. The efforts of The Cochrane Collaboration to identify randomized trials have not been matched for the identification of other types of studies. Consequently, including studies other than randomized trials in a review may require additional efforts to identify studies and to keep the review up to date, and might increase the risk that the result of the review will be influenced by publication bias. This issue and other bias-related issues important to consider when defining types of studies (e.g. whether to restrict study eligibility on the basis of language or publication status) are discussed in detail in Chapter 10.

Specific aspects of study design and conduct should also be considered when defining eligibility criteria, even if the review is restricted to randomized trials. For example, decisions over whether cluster-randomized trials (Chapter 16, Section 16.3) and cross-over trials (Chapter 16, Section 16.4) are eligible should be made, as should thresholds for eligibility based on aspects such as use of a placebo comparison group, evaluation of outcomes blinded to allocation, or a minimum period of follow-up. There will always be a trade-off between restrictive study design criteria (which might result in the inclusion of studies with low risk of bias, but which are very small in number) and more liberal design criteria (which might result in the inclusion of more studies, but which are at a higher risk of bias). Furthermore, excessively broad criteria might result in the inclusion of misleading evidence. If, for example, interest focuses on whether a therapy improves survival in patients with a chronic condition, it might be inappropriate to look at studies of very short duration, except to make explicit the point that they cannot address the question of interest.

5.6 Defining the scope of a review question (broad versus narrow)

The questions addressed by a review may be broad or narrow in scope. For example, a review might address a broad question regarding whether antiplatelet agents in general are effective in preventing all thrombotic events in humans. Alternatively, a review might address whether a particular antiplatelet agent, such as aspirin, is effective in decreasing the risks of a particular thrombotic event, stroke, in elderly persons with a previous history of stroke.

Determining the scope of a review question is a decision dependent upon multiple factors including perspectives regarding a question's relevance and potential impact; supporting theoretical, biologic and epidemiological information; the potential generalizability and validity of answers to the questions; and available resources.

There are advantages and disadvantages to both broad and narrow questions, some of which are summarized in Table 5.6.a. The validity of very broadly defined reviews may be criticized for 'mixing apples and oranges', particularly when good biologic or sociological evidence suggests that various formulations of an intervention behave very differently or that various definitions of the condition of interest are associated with markedly different effects of the intervention.

In practice, a Cochrane review may start (or have started) with a broad scope, and be divided up into narrower reviews as evidence accumulates and the original review becomes unwieldy. This may be done for practical and logistical reasons, for example to make updating easier as well as to make it easier for readers to keep up to date with the findings. Individual authors in consultation with their CRGs must decide if there are instances where splitting a broader focused review into a series of more narrowly focused reviews is appropriate and the methods that are implemented to achieve this (see Chapter 3, Section 3.4.4). If a major change is to be undertaken, such as splitting a broad review into a series of more narrowly focused reviews, a new protocol will need to be published for each of the component reviews which clearly document the eligibility criteria for each one.

The advent of Cochrane Overviews of reviews (Chapter 22, Section 22.1.1), in which multiple Cochrane reviews are summarized, may affect scoping decisions for reviews. Overviews can summarize multiple Cochrane reviews of different interventions for the same condition, or multiple reviews of the same intervention for different types of participants. It may increasingly be considered desirable to plan a series of reviews with a relatively narrow scope, alongside an Overview to summarize their findings.

Table 5.6.a: Some advantages and disadvantages of broad versus narrow review questions

	Broad scope	Narrow scope
Choice of participants e.g. corticosteroid injection for shoulder tendonitis (narrow) or corticosteroid injection for any tendonitis (broad)	<p><i>Advantages:</i></p> <p>Comprehensive summary of the evidence.</p> <p>Ability to assess generalizability of findings across types of participants.</p> <p><i>Disadvantages:</i></p> <p>May be more appropriate to prepare an Overview of reviews (see Chapter</p>	<p><i>Advantages:</i></p> <p>Manageability for review team; ease of reading.</p> <p><i>Disadvantages:</i></p> <p>Evidence may be sparse.</p>

	<p>22).</p> <p>Searching, data collection, analysis and writing may require more resources.</p> <p>Risk of ‘mixing apples and oranges’ (heterogeneity); interpretation may be difficult.</p>	<p>Findings may not be generalizable to other settings or populations.</p> <p>Scope could be chosen by review authors to produce a desired result.</p>
<p>Definition of an intervention</p> <p>e.g. supervised running for depression (narrow) or any exercise for depression (broad)</p>	<p><i>Advantages:</i></p> <p>Comprehensive summary of the evidence.</p> <p>Ability to assess generalizability of findings across different implementations of the intervention.</p> <p><i>Disadvantages:</i></p> <p>Searching, data collection, analysis and writing may require more resources.</p> <p>Risk of ‘mixing apples and oranges’ (heterogeneity); interpretation may be difficult.</p>	<p><i>Advantages:</i></p> <p>Manageability for review team; ease of reading.</p> <p><i>Disadvantages:</i></p> <p>Evidence may be sparse.</p> <p>Findings may not be generalizable to other formulations of the intervention.</p> <p>Scope could be chosen by review authors to produce a desired result.</p>
<p>Choice of interventions and comparisons</p> <p>e.g. alarms for preventing bed-wetting (narrow) or interventions for preventing bed-wetting (broad)</p>	<p><i>Advantages:</i></p> <p>Comprehensive summary of the evidence.</p> <p><i>Disadvantages:</i></p> <p>May be unwieldy, and more appropriate to present as an Overview of reviews (see Chapter 22).</p> <p>Searching, data collection, analysis and writing may require more resources.</p>	<p><i>Advantages:</i></p> <p>Manageability for review team.</p> <p>Clarity of objectives and ease of reading.</p> <p><i>Disadvantages:</i></p> <p>May have limited value when not included in an Overview.</p>

5.7 Changing review questions

While questions should be posed in the protocol before initiating the full review, these questions should not become a straitjacket that prevents exploration of unexpected issues (Khan 2001). Reviews are analyses of existing data that are constrained by previously chosen study populations, settings, intervention formulations, outcome measures and study designs. It is generally not possible to formulate an answerable question for a review without knowing some of the studies relevant to the question, and it may become clear that the questions a review addresses need to be modified in light of evidence accumulated in the process of conducting the review.

Although a certain fluidity and refinement of questions is to be expected in reviews as a fuller understanding of the evidence is gained, it is important to guard against bias in modifying questions. Data-driven questions can generate false conclusions based on spurious results. Any changes to the protocol that result from revising the question for the review should be documented in the section 'Differences between the protocol and the review'. Sensitivity analyses may be used to assess the impact of changes on the review findings (see Chapter 9, Section 9.7). When refining questions it is useful to ask the following questions:

- What is the motivation for the refinement?
- Could the refinement have been influenced by results from any of the included studies?
- Are search strategies appropriate for the refined question (especially any that have already been undertaken)?
- Are data collection methods appropriate to the refined question?

5.8 Chapter information

Editors: Denise O'Connor, Sally Green and Julian PT Higgins.

This chapter should be cited as: O'Connor D, Green S, Higgins JPT (editors). Chapter 5: Defining the review question and developing criteria for including studies. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Acknowledgements: This section builds on earlier versions of the *Handbook*. For details of previous authors and editors of the *Handbook*, see Chapter 1 (Section 1.4).

5.9 References

Cooper 1984

Cooper HM. The problem formulation stage. In: Cooper HM (editors). *Integrating Research: a Guide for Literature Reviews*. Newbury Park (CA): Sage Publications, 1984.

Counsell 1997

Counsell C. Formulating questions and locating primary studies for inclusion in systematic reviews. *Annals of Internal Medicine* 1997; 127: 380-387.

Gøtzsche 2007

Gøtzsche PC, Hróbjartsson A, Maric K, Tendal B. Data extraction errors in meta-analyses that use standardized mean differences. *JAMA* 2007; 298: 430-437.

Hedges 1994

Hedges LV. Statistical considerations. In: Cooper H, Hedges LV (editors). *The Handbook of Research Synthesis*. New York (NY): Russell Sage Foundation, 1994.

Jackson 1980

Jackson GB. Methods for integrative reviews. *Review of Educational Research* 1980; 50: 438-460.

Khan 2001

Khan KS, ter Riet G, Glanville J, Sowden AJ, Kleijnen J (editors). *Undertaking Systematic Reviews of Research on Effectiveness: CRD's Guidance for those Carrying Out or Commissioning Reviews (CRD Report Number 4)* (2nd edition). York (UK): NHS Centre for Reviews and Dissemination, University of York, 2001.

Richardson 1995

Richardson WS, Wilson MS, Nishikawa J, Hayward RSA. The well-built clinical question: a key to evidence based decisions. *ACP Journal Club* 1995: A12-A13.