

Cochrane Reviewers' Handbook 4.2.2

Updated March 2004

ABOUT THE HANDBOOK

Editors

Phil Alderson, Sally Green, and Julian Higgins.

How to cite this version of the Handbook

Alderson P, Green S, Higgins JPT, editors. Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]. In: The Cochrane Library, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd.

or

Alderson P, Green S, Higgins JPT, editors. Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]. <http://www.cochrane.org/resources/handbook/hbook.htm> (accessed 31st January 2004).

When referring to a specific section or subsection refer to it by the title and section number, NOT page numbers. For example:

Alderson P, Green S, Higgins JPT, editors. Formulating the problem. Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]; Section 4. In: The Cochrane Library, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd.

or

Alderson P, Green S, Higgins JPT, editors. Formulating the problem. Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]; Section 4. <http://www.cochrane.org/resources/handbook/hbook.htm> (accessed 31st January 2004).

When referring to a section that has section editors listed (currently only section 8), use:

Deeks JJ, Higgins, JPT, Altman DG, editors. Analysing and presenting results. In: Alderson P, Green S, Higgins J, editors. Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]; Section 8. In: The Cochrane Library, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd.

or

Deeks JJ, Higgins, JPT, Altman DG, editors. Analysing and presenting results. In: Alderson P, Green S, Higgins J, editors. Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]; Section 8. <http://www.cochrane.org/resources/handbook/hbook.htm> (accessed 31st January 2004).

Contact addresses

Phil Alderson
UK Cochrane Centre
Summertown Pavilion
Middle Way
Oxford OX2 7LG
United Kingdom

6.4 Performance bias

Performance bias refers to systematic differences in the care provided to the participants in the comparison groups other than the intervention under investigation. To protect against unintended differences in care and placebo effects, those providing and receiving care can be 'blinded' so that they do not know the group to which the recipients of care have been allocated. Some research suggests that such blinding is important in protecting against bias (Karlowski 1975, Colditz 1989, Schulz 1995). Studies have shown that contamination (provision of the intervention to the control group) and cointervention (provision of unintended additional care to either comparison group) can affect study results (CCSG 1978, Sackett 1979b). Furthermore, there is evidence that participants who are aware of their assignment status report more symptoms, leading to biased results (Karlowski 1975). For these reasons, reviewers may want to consider the use of 'blinding' as a criterion for validity. This can be done with the following questions: Were the recipients of care unaware of their assigned intervention? Were those providing care unaware of the assigned intervention?

A third question addressing blinding and detection bias is often added: Were persons responsible for assessing outcomes unaware of the assigned intervention? This addresses detection bias, as noted below.

Reviewers working on topics where blinding is likely to be important may want to develop specific criteria for judging the appropriateness of the method that was used for blinding. In some areas it may be desirable to use the same criterion across reviews, in which case a Collaborative Review Group (CRG) might want to agree to a standard approach for assessing blinding (Chalmers 1989, Schulz 1995, Jadad 1996, Moher 1996b).

6.5 Attrition bias

Attrition bias refers to systematic differences between the comparison groups in the loss of participants from the study. It has been called exclusion bias. It is called attrition bias here to prevent confusion with pre-allocation exclusion and inclusion criteria for enrolling participants. Because of inadequacies in reporting how losses of participants (e.g. withdrawals, dropouts, protocol deviations) are handled, reviewers should be cautious about implicit accounts of follow-up. The approach to handling losses has great potential for biasing the results and reporting inadequacies cloud this problem. What is reported, or more frequently implied, in study reports on attrition after allocation has not been found to be consistently related to bias (Schulz 1995). Thus reviewers should be cautious about using reported follow-up as a validity criterion, particularly when it is implied rather than explicitly reported. This is a general recommendation, however, and may not apply to certain topic areas that have higher quality reporting or where it is possible to obtain missing information from investigators.

6.6 Detection bias

Detection bias refers to systematic differences between the comparison groups in outcome assessment. Trials that blind the people who will assess outcomes to the intervention allocation should logically be less likely to be biased than trials that do not. Blinding is likely to be particularly important in research with subjective outcome measures such as

pain (Karlowski 1975, Colditz 1989, Schulz 1995). However, at least two empirical studies have failed to demonstrate a relationship between blinding of outcome assessment and study results. This may be due to inadequacies in the reporting of studies (Reitman 1988).

Bias due to the selective reporting of results is different from bias in outcome assessment. This source of bias may be important in areas where multiple outcome measures are used, such as evaluations of treatments for rheumatoid arthritis (Gotzsche 1989). Therefore, reviewers may want to consider specification of predefined primary outcomes and analyses by the investigators as indicators of validity. Alternatively, selective reporting of particular outcomes could be taken to suggest the need for better reporting and efforts by reviewers to obtain missing data.

6.7 Approaches to summarising the validity of studies

6.7.1 Simple approaches

There are several ways to rate validity. One is to rate individual criteria as 'met', 'unmet', or 'unclear' and to use individual criteria, such as adequacy of allocation concealment, in sensitivity analyses (see section 8.10). However, if several explicit criteria are used to assess validity, it is desirable to summarise these so as to derive an overall assessment of how valid the results of each study are. A simple approach to doing this is to use three categories such as the following:

Risk of bias	Interpretation	Relationship to individual criteria
A. Low risk of bias	Plausible bias unlikely to seriously alter the results	All of the criteria met
B. Moderate risk of bias	Plausible bias that raises some doubt about the results	One or more criteria partly met
C. High risk of bias	Plausible bias that seriously weakens confidence in the results	One or more criteria not met

The relationships suggested above will most likely be appropriate if only a few assessment criteria are used and if all the criteria address only substantive, important threats to the validity of study results. In general and when possible, reviewers should obtain further information from the authors of a report when it is unclear whether a criterion was met.

6.7.2 'Quality' scales and checklists

David Moher and his colleagues identified 25 scales and 9 checklists that have been used to assess the validity and 'quality' of randomised controlled trials (Moher 1995, Moher 1996b). These scales and checklists include anywhere from 3 to 57 items and take from 10 to 45 minutes to complete. Almost all of the items in the instruments are based on suggested or 'generally accepted' criteria that are mentioned in clinical trial textbooks. Many of the instruments are liable to confuse the quality of reporting with the validity of

Emerson 1990. Emerson JD, Burdick E, Hoaglin DC, Mosteller F, Chalmers TC. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Controlled Clin Trials* 1990; 11:339-52.

Feinstein 1982. Feinstein AR, Horwitz RI. Double standards, scientific methods, and epidemiological research. *N Engl J Med* 1982; 307:1611-7.

Feinstein 1985. Feinstein AR. *Clinical Epidemiology: The Architecture of Clinical Research*. Philadelphia: Saunders, 1985: 39-52.

Gotzsche 1989. Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Controlled Clin Trials* 1989;10:3-56.

Horwitz 1979. Horwitz RI, Feinstein AR. Methodological standards and contradictory results in case-control research. *Am J Med.* 1979; 66:556-64.

Jadad 1996. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clin Trials* 1996; 17:1-12.

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

Karlowski 1975. Karlowski TR, Chalmers TC, Frenkel LD, Kapikian AZ, Lewis TL, Lynch JM. Ascorbic acid for the common cold: a prophylactic and therapeutic trial. *JAMA* 1975; 231: 1038-42.

Kleijnen 1997. Kleijnen J, Gotzsche P, Kunz RA, Oxman AD, Chalmers I. So what's so special about randomization? In: Chalmers, I, Maynard, A, editors. *Non-Random Reflections on Health Services Research*. London: BMJ, 1997; 93-106.

Kunz 1995. Kunz RA, Oxman AD. Empirical evidence of selection bias in studies of the effects of health care: a systematic review. Presented at the Cochrane Colloquium, Oslo, 5-8 October, 1995.

Kunz 1998. Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998; 317:1185-90.

Levine 1994. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, IV: how to use an article about harm. *JAMA* 1994; 271:1615-9.

Moher 1995. Moher D, Jadad A, Nichol G, Penman M, Tugwell T, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Controlled Clin Trials* 1995; 16:62-73.

Moher 1996b. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials: current issues and future directions. *Int J Tech Assess in Health Care* 1996; 12:195-208.

Moher 1998a. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609-13.

Oxman 1993b. Oxman AD, Guyatt GH. The science of reviewing research. *Ann NY Acad Sci.* 1993;703:125-33.

Reitman 1988. Reitman D, Chalmers TC, Nagalingam R, Sacks H. Can efficacy of blinding be documented by meta-analysis? Presented to the Society for Clinical Trials, San Diego, 23-26 May, 1988.

Sackett 1979b. Sackett DL. Bias in analytic research. *J Chronic Dis* 1979; 32:51-63.

Schulz 1994. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994;272:125-8