Review

Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement

*David Moher, Deborah J Cook, Susan Eastwood, Ingram Olkin, Drummond Rennie, Donna F Stroup, for the QUOROM Group**

Summary

Background The Quality of Reporting of Meta-analyses (QUOROM) conference was convened to address standards for improving the quality of reporting of meta-analyses of clinical randomised controlled trials (RCTs).

Methods The QUOROM group consisted of 30 clinical epidemiologists, clinicians, statisticians, editors, and researchers. In conference, the group was asked to identify items they thought should be included in a checklist of standards. Whenever possible, checklist items were guided by research evidence suggesting that failure to adhere to the item proposed could lead to biased results. A modified Delphi technique was used in assessing candidate items.

Findings The conference resulted in the QUOROM statement, a checklist, and a flow diagram. The checklist describes our preferred way to present the abstract, introduction, methods, results, and discussion sections of a report of a meta-analysis. It is organised into 21 headings and subheadings regarding searches, selection, validity assessment, data abstraction, study characteristics, and quantitative data synthesis, and in the results with "trial flow", study characteristics, and quantitative data synthesis; research documentation was identified for eight of the 18 items. The flow diagram provides information about both the numbers of RCTs identified, included, and excluded and the reasons for exclusion of trials.

Interpretation We hope this report will generate further thought about ways to improve the quality of reports of meta-analyses of RCTs and that interested readers, reviewers, researchers, and editors will use the QUOROM statement and generate ideas for its improvement.

Lancet 1999; **354:** 1896–900
See Commentary page ????????

*Other members listed at end of paper

University of Ottawa, Thomas C Chalmers Centre for Systematic Reviews, Ottawa (D Moher MSc)**; McMaster University, Hamilton** (D J Cook MD)**, Ontario, Canada; University of California, San Francisco** (S Eastwood ELS(D))**; Stanford University, Stanford, CA** (I Olkin PhD)**; *JAMA*, Chicago, IL** (D Rennie PhD)**; and Centers for Disease Control and Prevention, Atlanta, GA, USA** (D F Stroup PhD)

Correspondence to: Dr David Moher, Thomas C Chalmers Centre for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, Ontario K1H 8L1, Canada (e-mail dmoher@uottawa.ca)

Introduction

Health-care providers and other decision-makers now have, among their information resources, a form of clinical report called the meta-analysis,[1-4] a review in which bias has been reduced by the systematic identification, appraisal, synthesis, and, if relevant, statistical aggregation of all relevant studies on a specific topic according to a predetermined and explicit method.[3] The number of published meta-analyses has increased substantially in the past decade.[5] These integrative articles can be helpful for clinical decisions, and they may also serve as the policy foundation for evidence-based practice guidelines, economic evaluations, and future research agendas. The value of meta-analysis is evident in the work of the international Cochrane Collaboration,[6,7] the primary purpose of which is to generate and disseminate high-quality systematic reviews of health-care interventions.

Like any research enterprise, particularly one that is observational, the meta-analysis of evidence can be flawed. Accordingly, the process by which meta-analyses are carried out has undergone scrutiny. A 1987 survey of 86 English-language meta-analyses[8] assessed each publication on 23 items from six content areas judged important in the conduct and reporting of a meta-analysis of randomised trials: study design, combinability, control of bias, statistical analysis, sensitivity analysis, and problems of applicability. The survey results showed that only 24 (28%) of the 86 meta-analyses reported that all six content areas had been addressed. The updated survey, which included more recently published meta-analyses, showed little improvement in the rigour with which they were reported.[9]

Several publications have described the science of reviewing research,[1] differences among narrative reviews, systematic reviews, and meta-analyses,[2] and how to carry out,[3,4,10] critically appraise,[11–15] and apply[16] meta-analyses in practice. The increase in the number of meta-analyses published has highlighted such issues as discordant meta-analyses on the same topic[17] and discordant meta-analyses and randomised-trial results on the same question.[18]

An important consideration in interpretation and use of meta-analyses is to ascertain that the investigators who did the meta-analysis not only report explicitly the methods they used to analyse the articles they reviewed, but also report the methods used in the research articles they analysed. The meta-analytical review methods used may not be provided when a paper is initially submitted: even when they are, other factors such as page limitations, peer review, and editorial decisions may change the content and format of the report before publication.

Several investigators have suggested guidelines for reporting of meta-analyses.[3,19] However, a consensus across disciplines has not developed. After the initiative to

improve the quality of reporting of randomised controlled trials (RCTs),[20–22] we organised the Quality of Reporting of Meta-analyses (QUOROM) conference to address these issues as they relate to meta-analyses of RCTs. This report summarises the proceedings of that conference. The issues discussed might also be useful for reporting of systematic reviews (ie, meta-analysis, as defined above, without statistical aggregation), particularly of RCTs.

Methods

The QUOROM steering committee began with a comprehensive review of publications on the conduct and reporting of meta-analyses. The databases searched included MEDLINE and the Cochrane Library,[23] which consists of the Cochrane Database of Systematic Reviews, the Cochrane Controlled Trials Register, the York Database of Abstracts of Reviews of Effectiveness, and the Cochrane Review Methodology Database. We examined reference lists of the retrieved articles and individual personal files. Articles of potential relevance were retrieved and critically appraised by the QUOROM steering committee. The committee generated a draft agenda for the conference, which included six domains requiring discussion and debate. The content areas were slightly modified during preliminary discussions at the conference and are reported as: the search for the evidence; decision-making on which evidence to include; description of the characteristics of primary studies; quantitative data synthesis; reliability and issues related to internal validity (or quality); and clinical implications related to external validity (or generalisability).

In planning the QUOROM conference, the steering committee identified clinical epidemiologists, clinicians, statisticians, and researchers who conduct meta-analysis as well as editors from the UK and North America who are interested in meta-analysis. These 30 individuals were invited to a conference in Chicago on Oct 2–3, 1996. Participants were surveyed before the meeting to elicit their views on current reporting standards of meta-analyses and whether these needed improvement. In addition, they were sent relevant citations for review and were asked to indicate in which of the six groups they wished to participate.

The conference included small-group and plenary sessions. Each small group had a facilitator who was a member of the steering committee and was responsible for ensuring the discussions of as many as possible of the issues relevant to their specific remit. Each small group also had a recorder, who was responsible for documenting the main points and the consensus on each issue discussed during that session; the recorder presented the group's consensus during the plenary sessions. During the plenary sessions, an elected scribe from each small group was responsible for recording the principal points relevant to that group's charge that arose during the plenary discussion.

The participants in each small group were asked to identify items that they thought should be included in a checklist of standards that would be useful for investigators, editors, and peer reviewers. We asked that, whenever possible, items included in the checklist be guided by research evidence that suggested that a failure to adhere to the particular checklist item proposed could lead to biased results. For example, a substantial lack of sensitivity and specificity of MEDLINE searches is evident.[24] Therefore, the checklist suggests that investigators explicitly describe all search strategies used to locate articles for inclusion in a meta-analysis. In considering whether candidate items were essential, each subgroup used a modified Delphi technique[25] that was replicated in the plenary sessions.

Heading	Subheading	Descriptor	Reported? (Y/N)	Page number
Title		Identify the report as a meta-analysis [or systematic review] of RCTs[26]		
Abstract		Use a structured format[27]		
		Describe		
	Objectives	The clinical question explicitly		
	Data sources	The databases (ie, list) and other information sources		
	Review methods	The selection criteria (ie, population, intervention, outcome, and study design); methods for validity assessment, data abstraction, and study characteristics, and quantitative data synthesis in sufficient detail to permit replication		
	Results	Characteristics of the RCTs included and excluded; qualitative and quantitative findings (ie, point estimates and confidence intervals); and subgroup analyses		
	Conclusion	The main results		
		Describe		
Introduction		The explicit clinical problem, biological rationale for the intervention, and rationale for review		
Methods	Searching	The information sources, in detail[28] (eg, databases, registers, personal files, expert informants, agencies, hand-searching), and any restrictions (years considered, publication status,[29] language of publication[30,31])		
	Selection	The inclusion and exclusion criteria (defining population, intervention, principal outcomes, and study design[32]		
	Validity assessment	The criteria and process used (eg, masked conditions, quality assessment, and their findings[33–36])		
	Data abstraction	The process or processes used (eg, completed independently, in duplicate)[35,36]		
	Study characteristics	The type of study design, participants' characteristics, details of intervention, outcome definitions, &c,[37] and how clinical heterogeneity was assessed		
	Quantitative data synthesis	The principal measures of effect (eg, relative risk), method of combining results (statistical testing and confidence intervals), handling of missing data; how statistical heterogeneity was assessed;[38] a rationale for any a-priori sensitivity and subgroup analyses; and any assessment of publication bias[39]		
Results	Trial flow	Provide a meta-analysis profile summarising trial flow (see figure)		
	Study characteristics	Present descriptive data for each trial (eg, age, sample size, intervention, dose, duration, follow-up period)		
	Quantitative data synthesis	Report agreement on the selection and validity assessment; present simple summary results (for each treatment group in each trial, for each primary outcome); present data needed to calculate effect sizes and confidence intervals in intention-to-treat analyses (eg 2×2 tables of counts, means and SDs, proportions)		
Discussion		Summarise key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process (eg, publication bias); and suggest a future research agenda		

Quality of reporting of meta-analyses

Results

The conference resulted in the QUOROM statement: a checklist (table) and a flow diagram (figure). The checklist of standards for reporting of meta-analyses describes our preferred way to present the abstract, introduction, methods, results, and discussion sections of a report of a meta-analysis. The checklist is organised into 21 headings and subheadings to encourage authors to provide readers with information on searches, selection, validity assessment, data abstraction, study characteristics, quantitative data synthesis, and trial flow. Authors are asked to provide a flow diagram (figure) providing information about the number of RCTs identified, included, and excluded and the reasons for excluding them.[10]

Pretesting

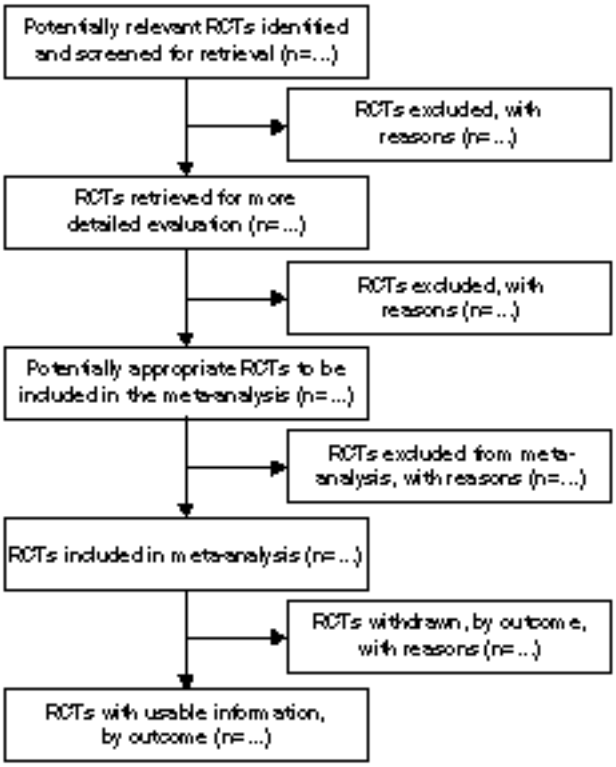
After development of the checklist and flow diagram, two members of the steering committee (DM, DJC) undertook pretesting with epidemiology graduate students studying meta-analysis, residents in general internal medicine, participants at a Canadian Cochrane Center workshop, and faculty members of departments of medicine and of epidemiology and biostatistics. One group of candidates for a master's degree in epidemiology used the checklist and flow diagram to report their meta-analyses as if their work were being submitted for publication. Feedback from these four groups was positive, most users stating that the checklist and flow diagram would be likely to improve reporting standards. Modifications of the checklist (eg, inclusion of a statement about major findings) and changes to the flow diagram (eg, more detail) were incorporated.

Discussion

In developing the checklist, we identified supporting scientific evidence for only eight of 18 items to guide the reporting of meta-analyses of RCTs.[26-39] Some of this evidence is indirect. For example, we ask authors to use a structured abstract format. The supporting evidence for this item was collected by examining abstracts of original reports of individual studies[27] and may not pertain specifically to the reporting of meta-analyses. However, the QUOROM group judged this a reasonable approach by analogy with other types of research reports and pending further evidence about the merits of structured abstracts for meta-analyses.

We have asked authors to be explicit in reporting the criteria used when assessing the "quality" of trials included in meta-analyses and the outcome of the quality assessment. There is direct and compelling evidence to support recommendations about reporting on the quality of RCTs included in a meta-analysis. A meta-analytic database of 255 obstetric RCTs provided evidence that trials with inadequate reporting of allocation concealment (ie, keeping the intervention assignments hidden from all participants in the trial until the point of allocation) overestimated the intervention effect by 30% compared with trials in which this information was adequately reported.[33] Similar results for several disease categories and methods of quality assessment have been reported.[34] These findings suggest that inclusion of reports of low-quality RCTs in meta-analyses is likely to alter the summary measures of the intervention effect.

We also ask authors to be explicit in reporting assessment of publication bias, and we recommend that the discussion should include comments about whether the results obtained may have been influenced by such bias. Publication bias derives from the selective publishing of studies with statistically significant or directionally positive results,[40–42] and it can lead to inflated estimates of efficacy in meta-analyses. For example, trials of single alkylating agents versus multiple-agent cytotoxic chemotherapy in the treatment of ovarian cancer have been analysed.[39] Published trials yielded significant results in favour of the multiple-agent therapy, but that finding was not supported when the results of all trials—both those published and those registered but not published—were analysed.



Progress through the stages of a meta-analysis for RCTs

The statement asks authors to be explicit about the publication status of reports included in a meta-analysis. Only about a third of published meta-analyses report the inclusion of unpublished data.[29,43] Although one study found that there were no substantial differences in the dimensions of study quality between published and unpublished clinical research,[42] another suggested that intervention effects reported in journals were 33% greater than those reported in doctoral dissertations.[44] The role of the "grey literature" (difficult to locate or retrieve) was examined in 39 meta-analyses that included 467 RCTs, 102 of which were grey literature.[29] Meta-analyses limited to published trials, compared with those that included both published and grey literature, overestimated the treatment effect by an average of 12%. There is still debate between editors and investigators about the importance of including unpublished data in a meta-analysis.[43]

We have asked authors to be explicit in reporting whether they have used any restrictions on language of publication. Roughly a third of published meta-analyses have some language restrictions as part of the eligibility criteria for including individual trials.[30] The reason for such restrictions is not clear, since there is no evidence to support differences in study quality, and there is evidence that language restrictions may result in a biased summary.

The reports of 127 RCTs written in English, compared with those reported in four other languages, showed little or no difference in several important methodological features.[45] Similar results have been reported elsewhere.[31] The role of language restrictions has been studied in 211 RCTs included in 18 meta-analyses in which trials published in languages other than English were included in the quantitative summary.[30] Language-restricted meta-analyses overestimated the treatment effect by only 2% on average compared with language-inclusive meta-analyses. However, the language-inclusive meta-analyses were more precise.[30]

Reports of RCTs with statistically positive results are more likely than those with negative results to be published in English.[31] Likewise, there is emerging evidence to suggest that reports of RCTs from certain countries mostly have statistically positive results.[46]

We used several methods to generate the checklist and flow diagram: a systematic review of the reporting of meta-analyses; focus groups of the steering committee; and a modified Delphi approach during the conference. Although we did not involve certain users of meta-analyses (policy-makers or patients), we formally pretested this document with representatives of several constituencies who would use the recommendations and made modifications accordingly.

The QUOROM group also discussed the format of a meta-analysis report, how best to assess the impact of the QUOROM statement, and how best to disseminate it. The format we recommend includes 15 subheadings that reflect the sequential stages in the conduct of the meta-analysis within the text of the report of a meta-analysis. The checklist included in the statement can also be used during the planning, performing, and reporting of a meta-analysis and during peer review of the report after its submission to a journal.

We delayed publication of the QUOROM statement until its impact on the editorial process had been assessed. We organised an RCT involving eight medical journals to assess the impact of use of QUOROM criteria on journal peer review. Accrual is now complete and we will report the trial results elsewhere.

After about 5 weeks of electronic posting we had received five comments from investigators, whom we thank for their thoughtful consideration of the statement. Several issues, in particular in relation to terminology, cannot be addressed in the statement at present. The QUOROM group is agreed on the importance of making changes to the checklist in the light of documented evidence and must resist changes based on opinion or anecdotal evidence unless there is a compelling rationale for doing otherwise. Nonetheless, the issues raised have been noted for consideration and discussion in future.

Several queries addressed the distinction between the meta-analysis and systematic review. As we indicate in the introduction, and throughout the statement, the QUOROM group agreed to observe the distinction as defined by the Potsdam consultation on meta-analysis.[3]

We were also asked to clarify the checklist item asking investigators to interpret their results in light of the totality of evidence. Increasingly, several meta-analyses on the same topic are reported.[47-49] If other similar reports are available, authors should discuss their results as they relate to such evidence.

For the QUOROM statement to continue to be useful, it must remain evidence based and up to date. Members of the QUOROM group need to survey the literature continually to help inform themselves about emerging evidence on reporting of meta-analyses. This information needs to be collated and presented annually for two purposes. The first is decisions on which checklist items to keep, delete, or add; these decisions can be made similarly to the selection of the original items. The second purpose is so that an up to date summary on the reporting of meta-analyses can be prepared. These efforts are being coordinated through a website. This approach is similar to the CONSORT initiative.

In summary, our choice of items to include in a meta-analysis report was based on evidence whenever possible, which implies the need to include items that can systematically influence estimates of treatment effects. Currently, we lack a detailed understanding of all the factors leading to bias in the result of a meta-analysis. Clearly, research is required to help improve the quality of reporting of meta-analyses. Such evidence may also act as a catalyst for improving the methods by which meta-analyses are conducted.

The QUOROM checklist and flow diagram are available on *The Lancet*'s website [www.thelancet.com]. We hope that this document will generate further interest in the field of meta-analysis and that, like the CONSORT initiative, the QUOROM statement will become available in different languages and locations as it is disseminated. We invite interested readers, reviewers, researchers, and editors to use the QUOROM statement and generate ideas for improvement.

Contributors

David Moher, Deborah Cook, Susan Eastwood, Ingram Olkin, Drummond Rennie, and Donna Stroup developed the QUOROM statement. They all planned the meeting, participated in regular conference calls, identified and secured funding, identified and invited participants, and planned the meeting agenda. All of them helped write the report, including revisions.

QUOROM participants

D G Altman (ICRF/NHS Centre for Statistics in Medicine, Oxford, UK); J A Berlin (University of Pennsylvania, Philadelphia, PA, USA); L Bero (University of California, San Francisco, CA, USA); W DuMouchel (AT&T Laboratories, New York, NY, USA); K Dickersin (Brown University, Providence, RI, USA); J J Deeks (ICRF/NHS Centre for Statistics in Medicine, Oxford, UK); P Fontanarosa (*JAMA*, Chicago, IL, USA); N Geller (National Heart, Lung, and Blood Institute, Bethesda, MD, USA); F Godlee (*BMJ*, London, UK); S Goodman (*Annals of Internal Medicine*, Philadelphia, PA, USA); R Horton (*The Lancet*, (London, UK); P Huston (University of Ottawa, Ottawa, Canada); A R Jadad (McMaster University, Hamilton, Canada); K Kafadar (University of Colorado, Denver, CO, USA); T Klassen (University of Alberta, Edmonton, Canada); S Morton (RAND, Santa Monica, CA, USA); C Mulrow (University of Texas, San Antonio, TX, USA); S Pyke (GlaxoWellcome, London, UK); H S Sacks (Mount Sinai School of Medicine, New York, NY, USA); K F Schulz, (Family Health International, Research Triangle Park, NC, USA); S G Thompson (Imperial College School of Medicine, London, UK); M Winker (*JAMA*, Chicago, IL, USA); S Yusuf (McMaster University, Hamilton, Canada).

Acknowledgments

We thank Iain Chalmers, Ted Colton, Sander Greenland, Brian Haynes, Edward J Huth, Alessandro Liberati, Tom Louis, Roy Pitkin, David Sackett, Trevor Sheldon, and Chris Silagy, for reviewing earlier drafts of this paper, and Jacqueline Page for helping with revisions.
Financial support was provided by Abbott Laboratories, Agency for Health Care Policy & Research, GlaxoWellcome, and Merck & Co.

References

1 Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987; **106:** 485–88.

2 Cook DJ, Mulrow C, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med* 1997; **126:** 376–80.

3 Cook DJ, Sackett DL, Spitzer W. Methodologic guidelines for systematic reviews of randomized controlled trials in health care from the Potsdam consultation on meta-analysis. *J Clin Epidemiol* 1995; **48:** 167–71.

4 Deeks J, Glanville J, Sheldon T. Undertaking systematic reviews of research on effectiveness CRD guidelines for those carrying out or commissioning reviews. CRD report no 4. York: NHS Centre for Reviews and Dissemination, University of York, 1996.

5 Chalmers I, Haynes RB. Reporting, updating, and correcting systematic reviews of the effects of health care. In: Chalmers I, Altman DG, eds. Systematic reviews. London: BMJ Publishing Group 1995: 86-95.

6 Bero L, Rennie D. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *JAMA* 1995; **274:** 1935–38.

7 Huston P. The Cochrane Collaboration helping unravel tangled web woven by international research. *Can Med Assoc J* 1996; **154:** 1389–92.

8 Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987; **316:** 450–55.

9 Sacks HS, Reitman D, Pagano D, Kupelnick B. Meta-analysis: an update. *Mt Sinai J Med* 1996; **63:** 216–24.

10 Mulrow CD, Oxman AD, eds. Cochrane Collaboration Handbook. In: The Cochrane Library [database on disk and CDROM]. Oxford: Cochrane Collaboration. Update Software: 1994, issue 4.

11 Oxman AD, Cook DJ, Guyatt GH, and the Evidence-Based Medicine Working Group. Users' guides to the medical literature: VI, how to use an overview. *JAMA* 1994; **272:** 1367–71.

12 Klassen TP, Jadad AR, Moher D. Guides for reading and interpreting systematic reviews: 1, getting started. *Arch Pediatr Adolesc Med* 1998; **152:** 700–04.

13 L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987; **107:** 224–33.

14 Olkin I. A critical look at some popular meta-analytic methods. *Am J Epidemiol* 1984; **140:** 287–88.

15 Olkin I. Statistical and theoretical considerations in meta-analysis. *J Clin Epidemiol* 1995; **48:** 133–46.

16 Guyatt GH, Sackett DL, Sinclair J, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature: IX, a method for grading health care recommendations. *JAMA* 1995; **274:** 1800–04.

17 Jadad AR, Cook DJ, Browman G. A guide to interpreting discordant systematic reviews. *Can Med Assoc J* 1997; **156:** 1411–16.

18 LeLorier J, Gregroire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* 1997; **337:** 536–42.

19 Shea B, Dubé C, Moher D. Assessing the quality of reports of meta-analyses: a systematic review of scales and checklists. In: Egger M, Davey Smith G, Altman DG, eds. Systematic reviews, 2nd edn. London: BMJ Publishing Group (in press).

20 The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. *JAMA* 1994; **272:** 1926–31.

21 The Asilomar Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Checklist of information for inclusion in reports of clinical trials. *Ann Intern Med* 1996; **124:** 741–43.

22 Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* 1996; **276:** 637–39.

23 The Cochrane Library [database on disk and CDROM]. Oxford: Cochrane Collaboration. Update Software, 1996, issue 3.

24 Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994; **309:** 1286–91.

25 Whitman N. The Delphi technique as an alternative for committee meetings. *J Nurs Educ* 1990; **29:** 377–79.

26 Dickersin K, Higgins K, Meinert CL. Identification of meta-analyses: the need for standard terminology. *Control Clin Trials* 1990; **11:** 52–66.

27 Taddio A, Pain T, Fassos FF, Boon H, Illersich AL, Einarson TR: Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association. *Can Med Assoc J* 1994; **150:** 1611–15.

28 Tramér M, Reynolds DJM, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* 1997; **315:** 635–40.

29 McAuley L, Moher D, Tugwell P. The influence of grey literature on meta-analysis. MSc Thesis: University of Ottawa, 1999.

30 Moher D, Pham B, Klassen TP, et al. Does the language of publication of reports of randomized trials influence the estimates of intervention effectiveness reported in meta-analyses? 6th Cochrane Colloquium; 1998.

31 Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997; **350:** 326–29.

32 Khan KS, Daya S, Collins JA, Walter S. Empirical evidence of bias in infertility research: overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. *Fertil Steril* 1996; **65:** 939–45.

33 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273:** 408–12.

34 Moher D, Pham B, Jones A, et al. Does the quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998; **352:** 609–13.

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

36 Berlin JA on behalf of the University of Pennsylvania meta-analysis blinding study group. Does blinding of readers affect the results of meta-analyses? *Lancet* 1997; **350:** 185–86.

37 Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. *JAMA* 1998; **279:** 1566–70.

38 Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 1994; **309:** 1351–55.

39 Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol* 1986; **4:** 1529–41.

40 Sterling TD, Rosenbaum WL, Weinkam JJ. Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa. *Am Statist* 1995; **49:**108–12.

41 Dickersin K, Min YI. NIH clinical trials and publication bias. *Online J Curr Clin Trials* 1993: April 28; doc no 50.

42 Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991; **337:** 867–72.

43 Cook DJ, Guyatt GH, Ryan G, et al. Should unpublished data be included in meta-analyses? Current convictions and controversies. *JAMA* 1993; **269:** 2749–53.

44 Smith ML. Publication bias and meta-analysis. *Eval Educ* 1980; **4:** 22–24.

45 Moher D, Fortin P, Jadad AR, et al. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet* 1996; **347:** 363–66.

46 Vickers A, Goyal N, Harland R, Rees R. Do certain countries produce only positive results? A systematic review of controlled trials. *Control Clin Trials* 1998; **19:** 159–66.

47 Kennedy E, Song F, Hunter R, Clark A, Gilbody S. Risperidone versus typical antipsychotic medication for schizophrenia (Cochrane Review). In: Cochrane Library, issue 3. Oxford: Update Software, 1999.

48 Davies A, Adena MA, Keks NA, Catts SV, Lambert T, Schweitzer I. Risperidone versus haloperidol: I, meta-analysis of efficacy and safety. *Clin Ther* 1998; **20:** 58–71.

49 Leucht S, Pitschel-Walz G, Abraham D, Kissling W. Efficacy and extrapyramidal side-effects of the new antipsychotics olanzapine, quetiapine, risperidone, and sertindole compared to conventional antipsychotics and placebo: a meta-analysis of randomized controlled trials. *Schizophrenia Res* 1999; **35:** 51–68.