

REVIEW

Systematic reviews in epidemiology: why are we so far behind?

Kay Dickersin

Accepted 24 October 2001

The public health and biomedicine communities have long recognized the need for regular synthesis of the available literature. Given the explosion of scientific information, there is simply too much literature available for any one person to be up-to-date. Review articles, which summarize existing knowledge in an area, serve to fill this need and are published in journals (e.g. *Epidemiologic Reviews*, *Annual Reviews of Public Health*) or special volumes, often as invited papers.

There are at least two general purposes to health-related reviews—one to summarize analytical research (e.g. analyses related to possible risk factors for a disease, or the efficacy of an intervention), and the other to summarize descriptive information (e.g. disease mechanisms, incidence and prevalence of a condition) and generate hypotheses and debate. While the value of analytical reviews is mainly to guide future research and clinical practice, the value of reviews that summarize descriptive information is mainly to educate. Some reviews serve both purposes, and others serve only one.

Reviews of analytical research are more likely to affect directly decision making (e.g. research funding, treatment), yet their validity is more clearly threatened by bias. While the review author's opinion can be quite valuable in selecting, presenting, and interpreting information which tells a story, it is potentially harmful in deciding which work should be included and excluded in summarizing existing data. Because of their particular relevance to epidemiological research, reviews of analytical studies will be the focus of this paper.

Systematic Reviews in Biomedicine

Probably most investigators have written at least one review article. But where are the instructions for how to go about doing it? Until recently, little or no attention has been paid to teaching students a 'correct' way to perform a review. It seems as if a double standard has existed: whereas the biomedical community has generally agreed on the elements of a good quality primary research study, we have not agreed to similar standards for reviews. The traditional review article typically has had no standard format and often has had no quantitative synthesis. Relevant to our discussion here, instructions for

authors for specific epidemiology journals rarely specify the components of review articles they aim to publish (Table 1).

This state of affairs is indeed strange in epidemiology, a field that defines itself in terms of research methodology. A review article is essentially an observational study, where the population is not humans but individual research studies, even when those studies are clinical trials. Data collection for a review thus involves either taking a sample of a well-defined population (of studies) or attempting to identify all members of the population if an unbiased sample is impossible. Most epidemiologists would not consider undertaking an observational study without a hypothesis, a predefined protocol for data collection and analysis, predetermined inclusion criteria, and careful consideration of ways to minimize bias, yet reviews are undertaken every day in epidemiology without attention to such methodology.^{1,2}

Systematic Reviews of Results in Biomedicine

'Systematic reviews,' reviews which report and adhere to a scientific research methodology that attempts to minimize bias and errors, have been proposed and used in response to problems with the traditional narrative review. Systematic reviews can include a meta-analysis, where the results of similar but separate studies are quantitatively combined.

Early reports of methodological research related to systematic reviews of clinical trials appeared in the 1980s³ and are now fairly common. In this research arena, The Cochrane Collaboration^{4,5} provides a supportive environment for identifying research needs related to systematic reviews through the Cochrane Methods Groups, for example, the Reporting Bias Group and the NonRandomized Studies Group. The Cochrane Methodology Register, published quarterly in *The Cochrane Library*, includes over 1300 citations, the vast majority of which relate to clinical trials.

To assist with the identification of all relevant studies for systematic reviews of clinical trials, the Cochrane Collaboration has developed CENTRAL, the most comprehensive database of controlled trials and other studies in existence today,⁶ for use by reviewers and others; CENTRAL includes over 300 000 citations as of December 2000. Several related efforts are underway to register ongoing trials, many of which may never be published. The most notable efforts are those of the US government,⁷ the UK government, and 'meta-registers' such as Current Controlled Trials⁸ and TrialsCentral.org.⁹

Center for Clinical Trials and Evidence-based Healthcare, Department of Community Health Box G-S2, 169 Angell Street, Providence, Rhode Island 02912, USA. E-mail: Kay_Dickersin@brown.edu

Table 1 Examples of epidemiology journals and instructions for authors on preparation of reviews and/or systematic reviews

Journal	Instructions related to reviews	Publisher	Structured abstracts for 'original research'?
<i>American Journal of Epidemiology</i>	'The Review and Commentary sections contain reviews and comments on various aspects of epidemiologic research.'	Oxford	No
<i>Annals of Epidemiology</i>	'[We publish] review articles, reports from US Federal and International sources, Editorials, Commentaries, Brief Communications, Letters to the Editor, Book Reviews, and selected papers from major symposia.'	Elsevier	Yes (Purpose, Methods, Results, Conclusions)
<i>Epidemiology</i>	'Original manuscripts reviewing specific substantive or methodologic topics in epidemiology will be considered for publication. They should be written for a general epidemiologic audience, and limited to 5000 words.'	Lippincott, Williams and Wilkins	No
<i>Epidemiologic Reviews</i>	'Epidemiologic Reviews does not publish reports on original research.'	Oxford	No
<i>European Journal of Epidemiology</i>	'Commentaries and reviews [are published].'	Kluwer Academic Publishers	No
<i>International Journal of Epidemiology</i>	'[We] publish original work, reviews, articles of interest and letters in the fields of research and teaching epidemiology.'	Oxford	Yes (Background, Methods, Results, Discussion)
<i>Journal of Clinical Epidemiology</i>	'Contributions are welcomed from any pertinent source, but should be written clearly enough to be understood by scholarly clinical readers ... pertinent symposia and reviews will be considered for publication.'	Elsevier	No
<i>Journal of Epidemiology and Community Health</i>	Reviews not explicitly mentioned. Provides only format for original articles	BMJ	Yes (Objective, Design, Setting, Patients, Results, Conclusions)

Systematic Reviews of Observational Studies

Where are we now?

Systematic reviews of observational studies of aetiology are especially important. Studies of this type tend to be limited in size and thus by examining multiple similar studies simultaneously we can achieve insight into real and spurious associations. If the associations detected by these observational studies are causal, even small increased risks are likely to have a large public health impact when exposures are common.

Systematic reviews of observational studies can also be important in detecting the consequences (such as harm) associated with an intervention. Clinical trials are rarely of sufficient size or duration to follow a meaningful number of patients for adverse events, especially when the events are unexpected, and thus systematically keeping track of up-to-date information from observational studies is critical. Beral and her colleagues, for example, have spearheaded an international effort to examine the association between breast cancer and a variety of exposures, including oral contraceptives, post-menopausal hormones, and induced abortion.¹⁰ Systematic reviews of analyses of large databases are another way of evaluating the consequences of health care interventions using observational data.

Systematic reviews and meta-analyses of observational studies are now fairly common, increasing dramatically in number over the last two decades. In 1985, Louis and colleagues¹¹ were only

able to identify four meta-analyses of epidemiological studies, and yet 11 years later over 400 meta-analyses of observational studies were published in a single year.¹² Syntheses involving observational studies appear to comprise almost half of all published meta-analyses.¹³ Still, epidemiology journals have not committed to the important move towards systematizing the review process: Breslow and her colleagues¹ found that, in 1995, less than 40% of reviews in seven epidemiology journals were systematic.

Obstacles to systematic reviews in epidemiology

Despite the burgeoning use of systematic reviews and meta-analyses in epidemiology, the community as a whole has not yet embraced them as important tools in its armamentarium and the approach remains controversial.¹⁴⁻¹⁶ Why is this? Certainly, formal, published systematic reviews are superior to those we do 'in our heads' when formal quantitative syntheses are not available. There are many possible reasons for the controversy. First, the systematic review is inexorably bound in people's minds with meta-analysis. While the benefits to doing systematic reviews seem almost irrefutable, the potential problems of meta-analysis related to observational studies (e.g. selection bias and confounding) seem insurmountable to some.¹⁴ Because of the high likelihood of achieving spurious results in a meta-analysis of observational studies, systematic reviews should probably not emphasize this approach.

Second, and perhaps more importantly, there is little methodological research relating to performance of systematic

reviews of observational studies, which leaves potential reviewers feeling vulnerable at the many decision points involved in doing a typical review. Although many of the systematic reviews in epidemiology include investigation of heterogeneity, possible biases and confounding, and perform sensitivity analyses, there have been few studies performed specifically to investigate the methodology of reviews in epidemiology.

Third, there is limited training available for those seriously interested in doing systematic reviews of observational studies, and until recently there has been little guidance on how to report them.¹²

Finally, there is no academic or other reward system in place for those who specialize in doing systematic reviews, perhaps because journals classify them separately from original research, they are difficult to get funded, and when they are funded they are wrongly assumed to be easy and quick to do.¹⁷

The need for methodological research

The time has come for a generalized international collaborative effort to identify all epidemiological studies, and to begin systematically reviewing available data for important epidemiological questions. The only endeavour comparable to the Cochrane Collaboration for epidemiological studies, as far as this author knows, is the HuGE (Human Genome Epidemiologic) Network—which produces systematic, structured, peer-reviewed synopses of epidemiological aspects of human genes in relation to specific diseases.¹⁸

If we in the epidemiology community were to start today a massive effort comparable to that undertaken by the Cochrane Collaboration, we would be unprepared for the challenges facing us, given the dearth of methodological research related to systematic reviews of observational studies. This is surprising given our field's methodological bent. But if systematic reviews and meta-analyses are themselves regarded with suspicion, interest in doing research on related methods may be unlikely. Furthermore, at least in the US, methodological research related to systematic reviews is typically not funded. Methodological research related to clinical trials may be more likely to be performed without formal funding sources because it can utilize funds available indirectly: clinical revenues, when available; funding allocated to trial conduct; and funding allocated for the conduct systematic reviews, such as those performed for the evidence-based practice reports supported by the Agency for Healthcare Research and Quality and many of the reviews performed by the Cochrane Collaboration.

What methodological research has been done so far relating to systematic reviews of observational studies and where are the gaps in our knowledge? Table 2 and the following paragraphs summarize major areas for suggested methodological research.

Dissemination of findings from epidemiological studies

Perhaps the major question facing those performing systematic reviews is whether there should be a minimum study quality for individual studies to be included in a review.^{19–21} (It should be noted that the quality of the study as it is reported, not as it was performed, is what is typically evaluated.) Proper reporting of observational studies has long been addressed in the literature,²² yet very little has been done to ensure implementation. For example, a standardized reporting structure for observational

studies, such as has been implemented for clinical trials,²³ has not yet been endorsed by journal editors. In another example, structured abstracts are still not ubiquitous in epidemiological journals (Table 1). Even when structured abstracts are required, they are not particularly informative regarding assessment of study design and other factors useful for those performing systematic reviews. Epidemiology journal editors should work together to develop uniform reporting structures²⁴ that will enhance report quality and inform the performance of systematic reviews.

Studies need to be done examining contributing factors to study quality and the relationship of those factors to study design, language in which the report is written, country of author or study origin, whether the study is published or not, the type of publication (e.g. MEDLINE-indexed, conference proceedings, thesis), and size of the association observed. Any association between quality and these items would indicate potential for bias when conducting reviews.

Identifying relevant studies

Selection bias is a major threat to a valid systematic review. Reporting bias, the tendency to publish findings depending on a bias at the investigator or editorial level, is one of the major ways in which selection bias is manifested. Publication or positive outcome bias, the selective reporting of studies based on the strength or direction of the results, has been identified for observational studies as well as clinical trials,^{25–27} though no in-depth analysis of observational studies has been done as it has for clinical trials. Other related areas of concern include bias against reporting or accepting 'negative' results in English language journals when the author's native language is not English;^{28–30} bias against accepting articles authored by one sex versus the other;³¹ bias against authors from countries different from the journal publisher or editor;³² and so forth. If systematic reviews limit their scope, for example, to English language articles, and if English language articles tend to report stronger associations,³³ then this would result in a biased review.

Another type of selection bias is the failure to identify all relevant reports for a systematic review. This would be acceptable if one could be assured that those reports identified were a random sample of all those eligible. Since this is not possible, it is critical that we understand the validity and reliability of the searching methods reviewers have available to them. Electronic searching of bibliographical databases is probably the most common method used to identify studies. In the field of clinical trials, research on the sensitivity and precision of MEDLINE has been ongoing since the early 1980s,^{34,35} and has expanded to include searching using other electronic databases and hand searching methods. In a nutshell, these studies have found that the most sensitive searches (e.g. hand searching) are not precise and compel the reader to review thousands of irrelevant articles or abstracts. Electronic searching, especially of issues before 1991, when important clinical trials indexing changes were made, is not terribly sensitive. Electronic searching can only retrieve articles that are present in the database being searched (e.g. conference proceedings and non-indexed journals would not typically be included), and only if they are indexed under the 'key words' selected by the searcher. Little has been done to examine searching sensitivity and precision for observational studies, and what has been done is limited to MEDLINE searching

Table 2 Methodological issues in systematic reviews and examples of research performed

Methodological issue	Examples of related methodological research	References to related methodological research
Study quality		
What dimensions of methodological quality of epidemiological studies are essential to avoid bias?	Establish and test elements of study quality, reproducibility and validity of quality scoring systems	19,29
Is there a relationship between study quality and study design, language, country of origin, publication, type of publication, and size of association observed?	Examine relationship between study quality and opportunities for bias	56,57
Selection bias		
What is the impact of publication bias?	Examine publication bias: Language bias, positive outcome bias (on the parts of authors, conference organizers, editors)	30,58,59
What is the impact of incomplete ascertainment of published studies on a systematic review?	Examine usefulness of contacting experts to ask about existence of other studies; compare results in grey literature and conference abstracts with indexed publications; examine sensitivity, precision, and cost of searching methods	35,39,40
What is the impact of incomplete ascertainment of all outcomes examined and analyses performed?	Examine effort required and success rate in contacting authors for additional information; compare published findings with findings obtained when authors contacted	39,60,61–63
Are the results of systematic reviews different when inclusion criteria are determined <i>post hoc</i> or non-reproducibly?	Compare findings of high-quality systematic reviews with traditional and lesser quality reviews; examine utility of blinding of abstractors to source, authors, and results when selecting studies for review	41,64,65
Information bias		
What is the impact on results of a systematic review when data must be deduced (e.g. from a graph) or assumed (e.g. from an abstract)?	Follow conference abstracts to determine proportion reaching full publication; compare reliability of information in abstracts with subsequent full reports; consider more uniform adoption of structured abstracts	42,44,66,67
What methods should be used to ensure that data collection is objective and reliable?	Test utility of having duplicate abstraction	–
Are the data provided in publications reliable and valid?	Compare results from systematic reviews of published data to reviews using individual participant data; consider development of CONSORT-like statement	61,68–70
Analysis		
Can the apparent heterogeneity of studies/results in the review be explained?	Examine association between results and study quality; study outcomes and selection of cases and controls, methods used for assessment of exposure and assessment of outcomes; possible confounding; role of test of heterogeneity; comparison of meta-analysis with single large study	47,52,70
Do the statistical methods used to combine individual studies influence overall estimates?	Compare results obtained using different assumptions and models, methods to adjust such as meta-regression; omission of studies because outcomes not available or in unusable format (e.g. no standard deviation available); effect of small numbers in cells	53,71–74
What is the effect of sensitivity analyses on the findings of the review?	Opportunities for bias; study features that should be routinely explored because of relationship to outcomes and methods to do this; when can/should data from studies of different designs be combined; differences in findings using individual level data versus grouped data?	46,56,61,75–77

of selected high impact clinical journals in general and internal medicine.³⁶ Finally, at least one study has shown that having more than one person to hand search a journal issue increased the proportion of trials identified.³⁷

The 'grey literature', which includes conference proceedings, books, and theses is not typically accessed using electronic searching and, along with unpublished studies and data, is often underascertained.³⁸ Studies need to be done to assess the size of the problem and to examine the utility of contacting experts to find out about other studies.³⁹ Furthermore, it would be useful to compare published findings with the additional findings obtained from making these contacts. Since one cannot hope to

find all reports published in the grey literature until a comprehensive register of observational studies is in place, it would be useful to know whether results published in the grey literature differ in any systematic way from results published in journals and bibliographically indexed publications. McCauley⁴⁰ found significantly larger estimates of an intervention's effect in 135 meta-analyses that depended on journal-published trials, compared to those that included 'publications' from the grey literature.

Selection bias can also occur in making decisions about which studies to include and exclude from systematic reviews. Ideally, all inclusion criteria are set before reviewing the studies

identified. But this is not always practical, since investigators may already know the literature fairly well, and eligibility criteria almost certainly need refinement based on the available information. One might imagine that 'high quality' systematic reviews, such as those that set eligibility criteria before beginning data collection, or those that mask reviewers to authors and results, might have different findings than those that did not. This has been tested by grouping reviews using (or not using) various methodologies, and comparing average effect sizes observed across reviews of each type.⁴¹

Information bias and selective reporting of outcomes

Information bias is a threat to any observational study and systematic reviews are no exception. One of the biggest challenges in doing a review is that one typically obtains information from a published article, not directly from individual study participants. This is not unlike obtaining data through a proxy respondent, with all its attendant problems. This is further complicated when one considers that sometimes only abstracts or other short summary reports are available, and these provide incomplete, sometimes preliminary, and even incorrect information. It would be useful to estimate the proportion of the literature (and information) one would miss if one chose to ignore abstracts.⁴² It would also be useful to be able to estimate the reliability of the information in abstracts, compared to full text articles⁴³ and the key information missed if one were to rely on short reports.⁴⁴

If one wanted to examine the possibility for information bias when data are obtained from systematic reviews of published data, one could compare results from reviews of this type to reviews using individual participant data on the same topic. As noted earlier, development of CONSORT-like guidelines for publishing observational studies, as well as adoption of structured abstracts, would address concerns about information bias generally, and increase clarity and comparability of published reports.⁴⁵

The quality and heterogeneity of epidemiological research

Performing an analysis of results as part of a systematic review, including a meta-analysis, is fraught with difficulties and opportunities for bias. To begin with, included studies are to at least some degree heterogeneous. Many of the meta-analyses of observational studies that have been conducted to date have explored the various sources of heterogeneity and their impact on a review's overall findings. For example, they have examined possible associations between results and study design,⁴⁶ selection of cases and controls,⁴⁷ methods used for assessment of exposure⁴⁸ and assessment of outcomes; differential effects in high- and low-risk groups;⁴⁹ and possible confounding.⁵⁰ Other areas for exploration would be the appropriate role of tests of heterogeneity and the frequency with which meta-analyses agree with single large observational studies.

Heterogeneity is typically explored using 'sensitivity analysis', where associations are estimated under various assumptions, for example including only studies conducted after a certain point in time, or studies performed in certain populations. Sensitivity analysis is useful for exploring bias: it might elucidate study features that should be routinely explored because of their relationship to outcomes of interests or the appropriateness of

combining data from studies of different designs or using different types of control groups. This type of exploration is quite common in epidemiological systematic reviews.^{51,52}

The statistical methods used to combine individual studies may influence overall estimates and results obtained using different assumptions and models should be compared.⁵³⁻⁵⁵ Other areas for study include the effect of omitting studies from a review or meta-analysis because outcomes are not available or, even if they are, they might not be in usable format (e.g. analysis of continuous data using mean change may not be possible if standard errors are more available).⁵⁴

Conclusions

Systematic reviews make sense and are here to stay. The epidemiological community should:

- Undertake methodological research related to systematic reviews;
- Develop methods to enhance the quality of reviews, including registers of observational studies conducted and registers of variables examined in these studies;
- Consider forming an international collaboration to undertake systematic reviews across all subject areas;
- Ensure that the science of systematically reviewing and interpreting available knowledge is taught as part of every epidemiology training programme; and
- Work with journal editors to standardize and improve the quality of reporting of observational studies, as well as the quality and profile of systematic reviews.

These efforts should be supported by the National Institutes of Health and other funding agencies, with realistic budgets. This may help academic promotions committees to recognize the importance of the research of reviews, which will ultimately contribute meaningfully to our knowledge base.

References

- ¹ Breslow RA, Ross SA, Weed DL. Quality of reviews in epidemiology. *Am J Public Health* 1998;**88**:475-77.
- ² Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987;**106**:485-88.
- ³ Chalmers I, Hetherington J, Elbourne D, Keirse MJNC, Enkin M. Materials and methods used in synthesizing evidence to evaluate the effects of care during pregnancy and childbirth. In: Chalmers I, Enkin M, Keirse MJNC (eds). *Effective Care in Pregnancy and Childbirth*. Oxford: Oxford University Press, 1989, pp.39-65.
- ⁴ Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. *Am J Med* 1982;**72**:233-40.
- ⁵ Dickersin K, Manheimer E. The Cochrane Collaboration: evaluation of health care and services using systematic reviews of the results of randomized controlled trials. *Clin Obstet Gynecol* 1998;**41**:315-31.
- ⁶ Dickersin K, Manheimer EW, Wieland LS, Robinson KA, Lefebvre C, McDonald S and the CENTRAL Development Group. Development of a centralized register of controlled clinical trials: The Cochrane Collaboration's CENTRAL Database. *Eval Health Prof* 2001 (in press).
- ⁷ McCray AT. Better access to information about clinical trials. *Ann Intern Med* 2000;**133**:609-14.
- ⁸ Tonks A. Registering clinical trials. *Br Med J* 1999;**319**:1565-68.
- ⁹ Schiff H, Haran C. Clinical trials watch. *MAMM* 2001;38.

- ¹⁰ Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52 705 women with breast cancer and 108 411 women without breast cancer. *Lancet* 1997;**350**: 1047–58.
- ¹¹ Louis TA, Fineberg HV, Mosteller F. Findings for public health from meta-analyses. *Annu Rev Public Health* 1985;**6**:1–20.
- ¹² Stroup DF, Berlin JA, Morton SC *et al.* and the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) Group. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA* 2000;**283**:2008–12.
- ¹³ Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *Br Med J* 1998;**316**:140–44.
- ¹⁴ Shapiro S. Meta-analysis/Shmeta-analysis. *Am J Epidemiol* 1994;**140**: 771–78.
- ¹⁵ Petitti DB. Of babies and bathwater. *Am J Epidemiol* 1994;**140**:779–82.
- ¹⁶ Goodman S. Commentary on meta-analysis. In: Hoffmeister H, Szklo M, Thamm M (eds). *Epidemiological Practices in Research on Small Effects*. Heidelberg, Germany: Springer, 1998, pp.113–18.
- ¹⁷ Petticrew M. Systematic reviews from astronomy to zoology: myths and misconceptions. *Br Med J* 2001;**322**:98–101.
- ¹⁸ Khoury MJLJ. Human Genome Epidemiologic reviews: the beginning of something HuGE. *Am J Epidemiol* 2000;**151**:2–3.
- ¹⁹ Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999;**282**:1054–60.
- ²⁰ Juni P, Altman D, Egger M. Assessing the quality of randomised controlled trials. In: Egger M, Smith GD, Altman DG (eds). *Systematic Reviews in Health Care: Meta-analysis in Context*. 2nd Edn. London: BMJ, 2001, pp.87–108.
- ²¹ Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care* 1996;**12**:195–208.
- ²² Bracken MB. Reporting observational studies. *Br J Obstet Gynaecol* 1989;**96**:383–88.
- ²³ 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.
- ²⁴ Ad hoc Working Group for Critical Appraisal of the Medical Literature. A proposal for more informative abstracts of clinical articles. *Ann Intern Med* 1987;**106**:598–604.
- ²⁵ Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* 1992;**267**:374–78.
- ²⁶ Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;**337**:867–72.
- ²⁷ Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *Br Med J* 1997;**315**: 640–45.
- ²⁸ 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.
- ²⁹ 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.
- ³⁰ Moher D, Pham B, Klassen TP *et al.* What contributions do languages other than English make on the results of meta-analyses? *J Clin Epidemiol* 2000;**53**:964–72.
- ³¹ Gilbert JR, Williams ES, Lundberg GD. Is there gender bias in JAMA's peer review process? *JAMA* 1994;**272**:39–42.
- ³² Nylenna M, Riis P, Karlsson Y. Multiple blinded reviews of the same two manuscripts. Effects of referee characteristics and publication language. *JAMA* 1994;**272**:149–51.
- ³³ Bassler D, Antes G, Egger M. Non-English reports of medical research. *JAMA* 2000;**284**:2996–97.
- ³⁴ Dickersin K, Hewitt P, Mutch L, Chalmers I, Chalmers TC. Perusing the literature: comparison of MEDLINE searching with a perinatal trials database. 6. *Contr Clin Trials* 1985;**6**:306–17.
- ³⁵ Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *Br Med J* 1994;**309**:1286–91.
- ³⁶ Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc* 1994;**1**:447–58.
- ³⁷ Clarke M, Westby M, McDonald S, Lefebvre C. How many hand-searchers does it take to change a light-bulb—or to find all of the randomised trials? (Poster No. 6). *2nd Symposium on Systematic Reviews: Beyond the Basics*. Oxford, UK, 5–7 January 1999.
- ³⁸ Brazier H. Poorly executed and inadequately documented? An analysis of the literature searches on which systematic reviews are based (Poster No. 1). *2nd Symposium on Systematic Reviews: Beyond the Basics*. Oxford, UK, 5–7 January 1999.
- ³⁹ McManus RJ, Wilson S, Delaney BC *et al.* Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. *Br Med J* 1998;**317**:1562–63.
- ⁴⁰ McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet* 2000;**56**:1228–31.
- ⁴¹ Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet* 1997;**350**:185–86.
- ⁴² Scherer R, Langenberg P. *Full Publication of Results Initially Presented in Abstracts*. (Cochrane Methodology Review.) Cochrane Library. Oxford: Update Software, 2001.
- ⁴³ Chokkalingham A, Scherer R, Dickersin K. Agreement of data in abstracts compared to full publications. *6th International Cochrane Colloquium*. 22–26 October, Baltimore, Maryland. 1998:A17.
- ⁴⁴ Deeks JJ, Altman DG. Inadequate reporting of controlled trials as short reports. *Lancet* 1998;**352**:1908.
- ⁴⁵ Scherer RW, Crawley B. Reporting of randomized clinical trial descriptors and use of structured abstracts. *JAMA* 1998;**280**:269–72.
- ⁴⁶ Ford ES, Smith SJ, Stroup DF, Steinberg KK, Mueller PW, Thacker SB. Homocyst(e)ine and cardiovascular disease: a systematic review of the evidence with special emphasis on case-control studies and nested case-control studies. *Int J Epidemiol* 2002;**31**:59–70.
- ⁴⁷ Pladevall-Vila M, Delclos GL, Varas C, Guyer H, Brugues-Tarradellas J, Anglada-Arisa A. Controversy of oral contraceptives and risk of rheumatoid arthritis: meta-analysis of conflicting studies and review of conflicting meta-analyses with special emphasis on analysis of heterogeneity. *Am J Epidemiol* 1996;**144**:1–14.
- ⁴⁸ Frumkin H, Berlin J. Asbestos exposure and gastrointestinal malignancy review and meta-analysis. *Am J Ind Med* 1988;**14**:79–95.
- ⁴⁹ Weiss HA, Quigley MA, Hayes RJ. Male circumcision and risk of HIV infection in sub-Saharan Africa: a systematic review and meta-analysis. *AIDS* 2000;**14**:2361–70.
- ⁵⁰ Tweedie RL, Mengersen KL. Lung cancer and passive smoking: reconciling the biochemical and epidemiological approaches. *Br J Cancer* 1992;**66**:700–05.
- ⁵¹ Realini JP, Goldzieher JW. Oral contraceptives and cardiovascular disease: a critique of the epidemiologic studies. *Am J Obstet Gynecol* 1985;**152**:729–98.
- ⁵² Longnecker MP, Berlin JA, Orza MJ, Chalmers TC. A meta-analysis of alcohol consumption in relation to risk of breast cancer. *JAMA* 1988;**260**:652–56.
- ⁵³ Berlin JA, Laird NM, Sacks HS, Chalmers TC. A comparison of statistical methods for combining event rates from clinical trials. *Stat Med* 1989;**8**:141–51.

- ⁵⁴ Engels EA, Schmid CH, Terrin N, Olkin I, Lau J. Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses. *Stat Med* 2000;**19**:1707–28.
- ⁵⁵ Goodman SN. Meta-analysis and evidence. *Control Clin Trials* 1989;**10**:188–204.
- ⁵⁶ 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.
- ⁵⁷ Moher D, Pham B, Jones A *et al*. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;**352**:609–13.
- ⁵⁸ Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;**9**(1Suppl.):15–21.
- ⁵⁹ Callahan ML, Wears RL, Weber EJ, Barton C, Young G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *JAMA* 1998;**280**:254–57.
- ⁶⁰ Herxheimer A. Data on file cited in pharmaceutical advertisements: What are they? International Congress on Biomedical Peer Review and Global Communication. Prague, Czech Republic, 19 September 1997.
- ⁶¹ Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: Is there a difference? *Lancet* 1993;**341**:418–22.
- ⁶² Steinberg KK, Smith SJ, Stroup DF *et al*. Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. *Am J Epidemiol* 1997;**145**:917–25.
- ⁶³ Hayward S, Brunton G, Thomas H, Ciliska D. Searching for the Evidence: Source, Time, and Yield. Fifth Annual Cochrane Colloquium. Amsterdam, The Netherlands, 8–12 October 1997: P282.
- ⁶⁴ Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *Can Med Assoc J* 1997;**156**:1411–16.
- ⁶⁵ Chalmers TC, Berrier J, Sacks HS, Levin H, Reitman D, Nagalingam R. Meta-analysis of clinical trials as a scientific discipline. II: Replicate variability and comparison of studies that agree and disagree. *Stat Med* 1987;**6**:733–44.
- ⁶⁶ Pitkin RM, Branagan MA. Can the accuracy of abstracts be improved by providing specific instructions? A randomized controlled trial. *JAMA* 1998;**280**:267–69.
- ⁶⁷ Froom P, Froom J. Deficiencies in structured medical abstracts. *J Clin Epidemiol* 1993;**46**:591–94.
- ⁶⁸ Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Stat Med* 1995;**14**:2057–79.
- ⁶⁹ Sanchez-Thorin JC, Cortes MC, Montenegro M, Villate N. The quality of reporting of randomized clinical trials published in Ophthalmology. *Ophthalmology* 2001;**108**:410–15.
- ⁷⁰ Friedenreich CM, Brant RF, Riboli E. Influence of methodologic factors in a pooled analysis of 13 case-control studies of colorectal cancer and dietary fiber. *Epidemiology* 1994;**5**:66–79.
- ⁷¹ Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology* 1993;**4**:218–28.
- ⁷² Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998;**17**:2815–34.
- ⁷³ Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *Br Med J* 1994;**309**:1351–55.
- ⁷⁴ Berlin JA, Antman EM. Advantages and limitations of metaanalytic regressions of clinical trials data. *Online J Curr Clin Trials* 1994; Doc. No. 134:[8425 words; 84 paragraphs].
- ⁷⁵ Ioannidis JP, Haidich AB, Pappa M *et al*. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001;**286**:821–30.
- ⁷⁶ Lijmer JG, Mol BW, Heisterkamp S *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–66.
- ⁷⁷ Marshall RJ. An empirical investigation of exposure measurement bias and its components in case-control studies. *J Clin Epidemiol* 1999;**52**:547–50.