

Methodology and Reports of Systematic Reviews and Meta-analyses

A Comparison of Cochrane Reviews With Articles Published in Paper-Based Journals

Alejandro R. Jadad, MD, DPhil; Deborah J. Cook, MD, MSc; Alison Jones, BA; Terry P. Klassen, MD, MSc; Peter Tugwell, MD, MSc; Michael Moher, MB; David Moher, MSc

Context.—Review articles are important sources of information to help guide decisions by clinicians, patients, and other decision makers. Ideally, reviews should include strategies to minimize bias and to maximize precision and be reported so explicitly that any interested reader would be able to replicate them.

Objective.—To compare the methodological and reporting aspects of systematic reviews and meta-analyses published by the Cochrane Collaboration with those published in paper-based journals indexed in MEDLINE.

Data Sources.—The Cochrane Library, issue 2 of 1995, and a search of MEDLINE restricted to 1995.

Study Selection.—All 36 completed reviews published in the Cochrane Database of Systematic Reviews and a randomly selected sample of 39 meta-analyses or systematic reviews published in journals indexed by MEDLINE in 1995.

Data Extraction.—Number of authors, trials, and patients; trial sources; inclusion and exclusion criteria; language restrictions; primary outcome; trial quality assessment; heterogeneity testing; and effect estimates. Updating by 1997 was evaluated.

Results.—Reviews found in MEDLINE included more authors (median, 3 vs 2; $P < .001$), more trials (median, 13.5 vs 5; $P < .001$), and more patients (median, 1280 vs 528; $P < .001$) than Cochrane reviews. More Cochrane reviews, however, included a description of the inclusion and exclusion criteria (35/36 vs 18/39; $P < .001$) and assessed trial quality (36/36 vs 12/39; $P < .001$). No Cochrane reviews had language restrictions (0/36 vs 7/39; $P < .01$). There were no differences in sources of trials, heterogeneity testing, or description of effect estimates. By June 1997, 18 of 36 Cochrane reviews had been updated vs 1 of 39 reviews listed in MEDLINE.

Conclusions.—Cochrane reviews appear to have greater methodological rigor and are more frequently updated than systematic reviews or meta-analyses published in paper-based journals.

JAMA. 1998;280:278-280

THE COCHRANE Collaboration (CC) is an international organization that aims to help people make well-informed decisions about health care by preparing, maintain-

ing, and promoting the accessibility of systematic reviews on the effects of health care interventions.^{1,2} The reviews produced within the CC are prepared following standardized instructions published in the *Cochrane Handbook*, which is widely available to reviewers in paper, CD-ROM, or Internet format. These reviews are included in the *Cochrane Database of Systematic Reviews*, a quarterly electronic publication.^{1,2} This electronic format allows reviewers to update or modify the Cochrane reviews in response to new evidence or comments and criticisms from readers. To safeguard the rigor and relevance of the Cochrane reviews, the CC has several levels of peer review, including assessment of protocols by editors and reviewers and evaluation of the reviews by methodology and content experts and by potential users.^{1,2}

There are major differences between

the peer review process followed by most paper-based journals and the system used by the CC. Few paper-based journals provide authors with explicit instructions to report systematic reviews and meta-analyses; their editors and peer reviewers only provide input after the reviews have been completed. Once the reviews are published there is limited opportunity for corrections of omissions or mistakes and few incentives to produce and publish updated versions.

It has been suggested recently that Cochrane reviews are less prone to bias than systematic reviews or meta-analyses published in paper-based journals.³ To date, however, we have been unable to identify other studies assessing the impact of the CC on the methodology and reports of systematic reviews and meta-analyses.

Against this background, we designed a study to (1) compare the methodological and reporting aspects of Cochrane reviews with those published in paper-based journals and (2) evaluate the frequency with which systematic reviews and meta-analyses, published by the CC or in paper-based journals, are updated.

METHODS

Reviews were included in the study if they included only randomized controlled trials, and if they were published in the Cochrane Database of Systematic Reviews or in paper-based journals indexed by MEDLINE. For the latter, they had to be coded by indexers as meta-analyses, or described by the authors as systematic reviews, meta-analyses, integrative research reviews, overviews, quantitative syntheses, and pooling or combining studies.

For the purposes of this study, all 36 Cochrane reviews published in the second issue of 1995 of the *Cochrane Library* (Update Software, Oxford, England) were included. They were compared with 39 randomly selected paper-based systematic reviews and meta-analyses, published in 32 different journals, identified using a refined search strategy of MEDLINE restricted to 1995 (Table). To

From the Health Information Research Unit (Dr Jadad), Department of Clinical Epidemiology and Biostatistics (Drs Jadad and Cook), and the Department of Medicine (Dr Cook), McMaster University, Hamilton, Ontario; Thomas C. Chalmers Center for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario (Ms Jones, Drs Klassen and Tugwell, and Mr D. Moher); the Departments of Pediatrics (Mr D. Moher) and Medicine (Dr Tugwell and Mr D. Moher), University of Ottawa, Ottawa, Ontario; and Department of Primary Health Care, University of Oxford, Oxford, England (Dr M. Moher).

Presented at the Third International Congress on Peer Review in Biomedical Publication, Prague, Czech Republic, September 20, 1997.

The views expressed in the article do not necessarily represent the views of the UK National Health Service.

Reprints: Alejandro R. Jadad, MD, DPhil, Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, McMaster University, 1200 Main St W, Hamilton, Ontario L8N 3Z5, Canada (e-mail: jadada@hsc.mcmaster.ca).

Search Strategy to Identify Systematic Reviews and Meta-analyses

Set	Search Terms
001	Meta-analysis: pt:sh.
002	Meta-anal: or metaanal: .tw.
003	Quantitativ review: or quantitativ overview: .tw.
004	Systematic review: or systematic overview: .tw.
005	Methodologic review: or methodologic overview: .tw.
006	Integrative research review: or research integration: .tw.
007	Review: pt, sh. or review: .tw. or overview: .tw.
008	Quantitativ synthes: .tw.
009	1 or 2 or 3 or 4 or 5 or 6 or 8
010	MEDLINE: or medlars: .tw., sh. or embase: .tw.
011	Scisearch: or psychinfo: or psycinfo: .tw.
012	Psychlit: or psyclit: .tw.
013	Hand search: or manual search: .tw.
014	Electronic database: or bibliographic database: .tw.
015	Pooling: or pooled analys: or mantel haenszel: .tw.
016	Peto: or der simonian: or dersimonian: or fixed effect: .tw.
017	10 or 11 or 12 or 13 or 14 or 15 or 16
018	7 and 17
019	9 or 18

ascertain the rate with which reviews are updated, the reviews published in 1995 and included in this study were followed up until June 1997 by searching the second issue of 1997 of the *Cochrane Library* and using the same MEDLINE strategy restricted to the period from January 1996 to June 1997.

Hard copies of all the reviews were obtained, but the identity of the authors and their affiliations were masked. The journal name was masked in the paper-based reviews. We could not mask the source of the documents as the reviews had different formats and we did not scan them. Using criteria defined a priori, 2 of us (A.R.J. and D.M.), independently, extracted the following information: number of authors, trials, and patients included; sources of trials; description of primary outcomes, inclusion and exclusion criteria, and pooled-effect estimates; and use of language restrictions, formal assessment of trial quality, and heterogeneity testing. Once data extraction was completed, the extractors met, compared their findings, and reached consensus. The number of those reviews that had been updated by 1997 was also evaluated.

The Mann-Whitney *U* test was used to compare the number of sources used and the number of trials and patients included in the reviews; χ^2 tests were used to compare the proportion of reviews that described the inclusion and exclusion criteria, heterogeneity testing, and primary outcomes, and the Fisher exact test was used for the number of reviews with language restrictions and with descriptions of the quantitative effect estimates. Values of $P < .05$ were regarded as statistically significant.

RESULTS

Reviews published in paper-based journals included more authors (median, 3 vs 2; $P < .001$), more trials (median, 13.5 vs 5; $P < .001$), and more patients (median, 1280 vs 528; $P < .001$) than Cochrane reviews.

More Cochrane reviews, however, included a description of the inclusion and exclusion criteria (35/36 vs 18/39; $P < .001$) and assessed trial quality (36/36 vs 12/39; $P < .001$). No Cochrane reviews had language restrictions (0/36 vs 7/39; $P < .01$). There were no statistically significant differences between the Cochrane reviews and those published in paper-based journals in the number of sources of trials (median, 3 vs 2; $P = .12$), the frequency of heterogeneity testing (47% vs 54%; $P = .56$), or the description of quantitative effect estimates (92% vs 90%; $P = .55$). None of the Cochrane reviews and only 3 of the reviews published in paper-based journals described the primary outcome(s) of interest.

By June 1997, more Cochrane reviews had been updated (18/36 vs 1/39; $P < .001$).

COMMENT

Compared with reviews published in paper-based journals, Cochrane reviews include elements that make them less prone to bias, such as the description of inclusion and exclusion criteria and formal assessment of trial quality.⁴ The methodological deficiencies that we found in reviews published in paper-based journals are similar to those reported by others.⁴⁻⁶ Perhaps paper-based journals could improve the methodology and reporting of the systematic reviews they publish by influencing the review process at a much earlier stage and by encouraging more frequent updates or correction of published material using other media such as the Internet.

Almost all reviews included in this study used meta-analyses (quantitative data synthesis), but only half of them incorporated formal assessments of heterogeneity testing. This may indicate that most reviewers use meta-analysis as the "default action" to synthesize the information provided by the individual studies. The proportion of reviews that should include meta-analysis is unknown, but is likely to vary depending on unique features of the topics and studies being summarized.

The number of recorded sources of trials used by Cochrane reviews is likely to be an underestimate, as most Cochrane reviews use specialized registers that typically include extensive searches of bibliographic databases (ie, MEDLINE) and hand searches of journals.

Most reviews did not specify the primary outcome(s). This could have been

due to oversight (unlikely), to the lack of instructions for authors, to the lack of information on primary outcomes in the individual studies being reviewed, or to the perception that reporting the primary outcome is not as important in a review as it may be in an individual trial. The latter could be acceptable from the users' perspectives, as they would be given the opportunity to select freely the outcomes of greatest interest to them in different circumstances. However, from the researchers' perspective, the lack of clearly identified primary outcomes hinders empirical methodological studies designed to assess the effect of different characteristics of the reviews (ie, language restrictions, quality assessment, etc) on the direction of the results.

Cochrane reviews were updated more frequently than their paper-based counterparts. It is unclear, however, whether 50% is below or above the optimal update rate for systematic reviews. The comparatively low update rate among paper-based reviews suggests that editors of such journals are not sufficiently interested in publishing updated versions of previously published systematic reviews and meta-analyses or, if they are interested, that authors are not aware of such interest. Alternatively, authors of reviews published in paper-based journals may lack the interest or the resources to update them. We did not contact and ask the authors of these nonupdated reviews because of our intention to follow up the reviews in time. Regardless of the causes for not updating reviews, these reviews could be updated within the framework and with the support of the CC. Alternatively, it would be worthwhile for authors and editors to consider the publication of any updated review as research letters.

Our study also had limitations. The inclusion of reviews published in journals indexed only by MEDLINE limits the generalization of the results to reviews indexed in other databases. The lack of masking as to the origin of the reviews (Cochrane or MEDLINE) and the close links of most of us with the CC could have also introduced bias. Data extraction under masked conditions, however, has been shown to have little effect on systematic reviews.⁷

We will follow up these 2 groups of reviews to judge the impact of a more mature set of tools produced by the CC, particularly the Cochrane Criticisms Editor, as well as any set of standardized instructions for authors of systematic reviews and meta-analyses produced by paper-based journals. The Cochrane Criticisms Editor is an electronic facility included in the *Cochrane Library* for the delivery of comments on Cochrane reviews to editors and authors.² This sys-

tem is expected to further improve the validity and relevance of the reviews.

This study was funded by the UK National Health Service Research and Development Programme grant 93/52/3. Dr Jadad was supported by a National Health Research Scholar Award from Health Canada, Ottawa, Ontario, and Dr Cook is a career scientist supported by the Ontario Ministry of Health, Toronto, Ontario.

A list of the articles included in this review is available upon request.

References

1. Bero L, Rennie D. The Cochrane Collaboration: preparing, maintaining, and disseminating reviews of the effects of health care. *JAMA*. 1995;274:1935-1938.
2. Jadad AR, Haynes RB. The Cochrane Collaboration: advances and challenges in improving evidence-based decision making. *Med Decis Making*. 1998;279:611-614.
3. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-634.
4. Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualita-

tive review of their methodology. *J Clin Epidemiol*. 1996;49:235-243.

5. Assendelft WJJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. *JAMA*. 1995;274:1942-1948.
6. Sacks HS, Reitman D, Pagano D, Kupelnick B. Meta-analysis: an update. *Mt Sinai J Med*. 1996;3-4:216-224.
7. Berlin JA, for the University of Pennsylvania Meta-analysis Blinding Study Group. Does blinding of readers affect the results of meta-analyses? *Lancet*. 1997;350:185-186.

Discussion Sections in Reports of Controlled Trials Published in General Medical Journals

Islands in Search of Continents?

Michael Clarke, DPhil; Iain Chalmers, MSc

Context.—Several journals have adopted the Consolidated Standards of Reporting Trials (CONSORT) recommendations to make assessment of the quality of randomized controlled trials (RCTs) easier. One of these recommendations is that the trial's results be discussed in light of the totality of the available evidence.

Objective.—To assess the extent to which reports of RCTs published in 5 general medical journals have discussed new results in light of all available evidence.

Design.—Assessment of the discussion sections in all 26 reports of RCTs published during May 1997 in *Annals of Internal Medicine*, *BMJ*, *JAMA*, *The Lancet*, and *The New England Journal of Medicine*.

Main Outcome Measure.—The inclusion or mention of a systematic review in the discussion section of each article.

Results.—In only 2 articles were the RCT's results discussed in the context of an updated systematic review of earlier trials. In a further 4 articles, references were made to relevant systematic reviews, but no attempts were made to integrate the results of the new trials in updated versions of these reviews. One article was probably the first published trial to address the question studied. The remaining 19 articles included no evidence that any systematic attempt had been made to set the reported trial's results in the context of previous trials.

Conclusion.—There is little evidence that journals have adequately implemented the CONSORT recommendation that results of an RCT be discussed in light of the totality of the available evidence.

JAMA. 1998;280:280-282

SEVERAL MAJOR health care journals have already adopted the Consolidated Standards of Reporting Trials (CONSORT) recommendations to make it easier for readers to assess the quality of controlled trials.¹ This is the first joint attempt by biomedical journals to improve the

quality of reports of controlled trials, a topic of research for 4 decades.² No other category of biomedical report has received such sustained attention, and this reflects the practical importance of controlled trials in guiding decisions in health care.

Previous research has highlighted deficiencies in descriptions of the materials and methods used and the analysis and presentation of results.² Like most similar articles before it, the CONSORT statement concentrates on these 2 elements of reports of controlled trials. By contrast, the quality of introduction and discussion sections in trial reports has received little

systematic scrutiny. The typical discussion section usually addresses a number of dimensions, but, crucially, it is in this section that readers will look for an answer to Bradford Hill's "bottom line" question for any research article: "What does it mean, anyway?"³ This was recognized in the CONSORT statement, which included the recommendation that trialists should "state general interpretation of the data in light of the totality (our emphasis) of the available evidence."¹

Other research has illustrated how selective citation of previous research in the discussion sections of research articles can be biased. Studies that have yielded relatively dramatic results are more likely to be cited in reports of subsequent similar studies than previous studies yielding unremarkable point estimates of effects.⁴ In addition, authors from a particular country or specialty have been shown to selectively cite material generated from within that country^{5,6} or specialty.^{7,8}

Ideally, the discussion section of the report of a new trial should involve the presentation of an up-to-date systematic review,⁹ as was done, for example, in the 1986 article on the First International Study of Infarct Survival (ISIS-1).¹⁰ To show how closely reports of controlled trials reflect this ideal more than a decade after ISIS-1, we assessed the discussion sections of all reports of randomized trials published in May 1997 in 5 general medical journals. This month was chosen to allow as up-to-date an assessment as possible and without prior knowledge of the articles to be published. We concentrated on how well the discussion sections

From the Clinical Trial Service Unit, Oxford University (Dr Clarke), and the United Kingdom Cochrane Centre (Drs Clarke and Chalmers), Oxford, England.

Presented at the Third International Congress on Peer Review in Biomedical Publication, Prague, Czech Republic, September 20, 1997.

Reprints: Michael Clarke, DPhil, ICRF Clinical Trial Service Unit and Epidemiological Studies Unit, Radcliffe Infirmary, Oxford OX2 6HE, England (e-mail: mike.clarke@ctsu.ox.ac.uk).