Effects of Editorial Peer Review
A Systematic Review

Tom Jefferson, MD
Philip Alderson, MBChB
Elizabeth Wager, MA
Frank Davidoff, MD

Context Editorial peer review is widely used to select submissions to journals for publication and is presumed to improve their usefulness. Sufficient research on peer review has been published to consider a synthesis of its effects.

Methods To examine the evidence of the effects of editorial peer-review processes in biomedical journals, we conducted electronic and full-text searches of private and public databases to June 2000 and corresponded with the World Association of Medical Editors, European Association of Science Editors, Council of Science Editors, and researchers in the field to locate comparative studies assessing the effects of any stage of the peer-review process that made some attempt to control for confounding. Nineteen of 135 identified studies fulfilled our criteria. Because of the diversity of study questions, methods, and outcomes, we did not pool results.

Results Nine studies considered the effects of concealing reviewer/author identity. Four studies suggested that concealing reviewer or author identity affected review quality (mostly positively); however, methodological limitations make their findings ambiguous, and other studies' results were either negative or inconclusive. One study suggested that a statistical checklist can improve report quality, but another failed to find an effect of publishing another checklist. One study found no evidence that training referees improves performance and another showed increased interrater reliability; both used open designs, making interpretation difficult. Two studies of how journals communicate with reviewers did not demonstrate any effect on review quality. One study failed to show reviewer bias, but the findings may not be generalizable. One nonrandomized study compared the quality of articles published in peer-reviewed vs other journals. Two studies showed that editorial processes make articles more readable and improve the quality of reporting, but the findings may have limited generalizability to other journals.

Conclusions Editorial peer review, although widely used, is largely untested and its effects are uncertain.

JAMA. 2002;287:2784-2786 www.jama.com

THE USE OF PEERS TO ASSESS THE work of fellow scientists goes back at least 200 years.[1] It is usually assumed to raise the quality of the end product and to provide a mechanism for rational, fair, and objective decision making. Despite the fact that peer review has such a long history and is so well established, research into its effects is a recent phenomenon. However, the body of original research on the effects of peer review has been growing, and systematic review and synthesis may now be possible.[2] This review assesses the effects of processes undertaken as part of editorial peer review of original research studies submitted for paper or electronic publication in biomedical journals.

METHODS

We used Cochrane methods to carry out our review, which will also be published in the Cochrane Database of Methodology Reviews.[3] A detailed description of the search strategy, sources, and terms used is available in online Table A (http://www.jama.com).[4-6]

We considered for inclusion all types of comparative studies in which some attempt to control for confounding had been made. This was an attempt to widen the criteria to collect as much evidence as possible, without collecting studies in which obvious confounding might explain any observed effect. We considered studies assessing the effects of any stage of the peer review process except those solely concerned with technical editing, which was the subject of another review.[7] For example, we looked for studies of the effects of different ways of assigning submissions to assessors and eliciting opinions, decision-making procedures, and methods of feedback to authors.

Two reviewers (T.J. and P.A.) examined all retrieved citations. Studies for possible inclusion were retrieved in full. The same 2 reviewers then examined studies independently against the selection criteria. The same reviewers extracted information on study design and outcomes. We collected descriptive information on study quality, as reported herein. We identified 135 reports of studies that could possibly fulfill our inclusion criteria. Nineteen of these fulfilled our criteria and we excluded the remaining 116 studies from our review.

We did not pool the results of similar studies into a formal meta-analysis because no 2 studies were alike and all asked a slightly different study question or used different designs or outcome measures. We decided to group studies by the broad issues they addressed.

RESULTS

A summary of the 19 included studies is shown online Table B (http://www.jama.com). A more detailed

Author Affiliations: Health Reviews Ltd, Rome, Italy (Dr Jefferson); UK Cochrane Centre, Oxford, England (Dr Alderson); Sideview, Princes Risborough, England (Ms Wager); and *Annals of Internal Medicine*, Philadelphia, Pa (Dr Davidoff).
Corresponding Author and Reprints: Tom Jefferson, MD, Reparto Epidemiologia Clinica, Istituto Superiore di Sanità, Viale Regina Elena, 299-00161, Rome, Italy (e-mail: toj1@aol.com).

description will be available in our Cochrane review.[3]

We included 9 studies assessing the effect of blinding/masking or revealing authors' and/or reviewers' identities in some way (eg, exchanging reviews between peer reviewers) on the quality of external reviews.[8-16] All of the studies in this group were randomized controlled trials. Five studies[11-15] reported no apparent effect in quality of reviews, time taken to review, or tone of reviews. The 4 studies[8-10,16] that suggested that concealing reviewer or author identity affected review quality had methodological limitations that make their outcomes ambiguous. The most interesting methodological issue we identified in the design of the studies was the consistent difficulty in ensuring robust blinding procedures. This was probably a reflection of the relatively small world of editorial peer review, especially when applied to specialist areas of knowledge. Overall, most of the studies were inconclusive because blinding was ineffective and confidence intervals were not sufficiently narrow to allow firm conclusions to be made.

We included 2 studies that assessed the effects of submission checklists on outcomes.[17,18] The study by Gardner and Bond [18] shows some benefit from using a statistical checklist, although this finding may be of limited generalizability given the setting of the study (*BMJ*) and its small size. The other study found that publication of a checklist for economic studies did not appear to improve submissions.[17]

Two studies assessed the effects of the media used by journals to communicate with reviewers. One found that posting to the Internet may benefit authors but does not appear to affect the quality of reviews.[19] The randomized controlled trial by Neuhauser and Koran[20] showed that warning reviewers of an impending review by telephone lengthened total turnaround time, although it shortened review time.

Two studies assessed the effects of training reviewers.[21,22] The results were ambiguous in the study by Callaham et al,[21] showing no obvious effect, in contrast with the results of Strayhorn et al,[22] which showed that interrater reliability increased. The open design of both studies makes minimization of biases impossible, and results are therefore difficult to interpret. Generalizability of results beyond the context of the specific interventions tested is questionable.

A randomized controlled trial by Ernst and Resch[23] assessed the presence and effects of reviewer bias. They found no evidence of reviewer bias toward an unconventional treatment. Despite its randomized design, bias may distort the results of this study (the response rate was 61%).

We included 1 study assessing the effects of peer review on study validity,[24] which failed to show any difference, although the small size and nonrandomized design of the study again makes generalizability of its results uncertain.

Two studies assessed the effects of peer review on study report quality.[25,26] Both studies showed a beneficial effect, but results may again have limited generalizability because of atypical settings (both journals studied are well resourced and keen on improving quality).

In summary, the quality of randomized studies was superior to that of trials with open design. (Quality refers to the ability of a study to answer its research question.) Overall, the small numbers of reviews and reviewers involved, potentially atypical settings (with a few major journals being the object of many studies), and many methodological weaknesses make both the internal and external validity of these studies difficult to interpret.

COMMENT

Our review was limited to evidence about peer review as practiced by biomedical journals. We discovered remarkably few well-designed studies assessing the effects of this process. Of the studies identified, the majority were focused on specific editorial processes, such as masking of reviewers and authors or use of checklists. Very few examined the broader effects of peer review. We assess the implications of this in a separate study.[27] The only study that attempted to address broader issues[24] is difficult to interpret because of methodological weaknesses and the likelihood of limited generalizability. The results of our systematic review are consistent with those of an earlier descriptive review.[2]

Given the widespread use of peer review and its importance, it is surprising that so little is known of its effects. However, the research needed to address these questions would require a well-funded and coordinated effort involving several sectors of the scientific community as well as the cooperation of large numbers of authors and editors, and the methodological issues in conducting proper studies of the subject are daunting.

The 2 principal functions of peer review, filtering out incorrect or inadequate work and improving the accuracy and clarity of published reports, directly reflect the 2 virtues that Francis Bacon in 1605 attributed to doubt: guarding "against errors" and causing issues that "would have been passed by lightly without intervention" to be "attentively and carefully observed." As such, the doubts contributed by peer review are an intrinsic and essential part of science, layered over the critical reviews that take place during the process of doing the work. Future research in biomedical peer review needs to recognize these basic concepts and must consider a number of specific underlying methodological issues. Among these are both the positive contributions of peer review and its abuses, the complexity of biomedical science, and the incremental nature of scientific progress. This is based on rare major discoveries and very frequent minor ones, clarifying patterns of life and disease and society's response to them. Moreover, most studies of biomedical journal peer review have been concerned with manuscripts that were accepted for publication. A number of important questions about peer review can only be answered, however, by studying rejected manuscripts as well as those that are accepted.

Until such research is undertaken, peer review should be regarded as an untested process with uncertain outcomes.

Author Contributions: *Study concept and design:* Jefferson, Alderson, Wager, Davidoff.
Acquisition of data: Jefferson, Alderson.
Analysis and interpretation of data: Jefferson, Alderson, Wager, Davidoff.
Drafting of the manuscript: Jefferson, Alderson.
Critical revision of the manuscript for important intellectual content: Jefferson, Wager, Davidoff.
Obtained funding: Jefferson.
Administrative, technical, or material support: Alderson.
Study supervision: Jefferson.

Financial Disclosure: Dr Jefferson coedited the book *Peer-Review in Health Sciences*. All authors are active peer reviewers and have published articles in peer-reviewed journals.

Disclaimer: This article is based on a Cochrane protocol published in the Cochrane Library; 2002, issue 2 (see http://www.CochraneLibrary.net for information). Cochrane reviews are regularly updated as new evidence emerges and in response to comments and criticisms, and the Cochrane Library should be consulted for the most recent version of the review. The results of a Cochrane review can be interpreted differently, depending on perspectives and circumstances. Please consider the conclusions presented carefully. They are the opinions of review authors and are not necessarily shared by the Cochrane Collaboration.

Acknowledgment: We thank Iain Chalmers, MD, for inspiration; Drummond Rennie, MD, for helpful discussion; and Philippa Middleton, BSc, and Sarah White for help with the literature searches and administration of the review.

REFERENCES

1. Kronick DA. Peer-review in 18th-century scientific journalism. *JAMA*. 1990;263:1321-1322.
2. Overbeke J. The state of the evidence: what we know and what we don't know about journal peer review. In: Godlee F, Jefferson T, eds. *Peer Review in Health Sciences*. London, England: BMJ Books; 1999:32-45.
3. Alderson P, Davidoff F, Jefferson TO, Wager E. Editorial peer review for improving the quality of reports of biomedical studies [Protocol for a Cochrane Methodology Review]. Oxford, England: Cochrane Library, Update Software; 2001; issue 3.
4. Rennie D. Editorial peer review in biomedical publication. *JAMA*. 1990;263:1317.
5. Rennie D, Flanagin A. The second International Congress on Peer Review in Biomedical Publication. *JAMA*. 1994;272:91.
6. Rennie D, Flanagin A. Congress on Biomedical Peer Review. *JAMA*. 1998;280:213.
7. Wager E, Middleton P. Effects of technical editing in biomedical journals: a systematic review. *JAMA*. 2002; 287:2821-2824.
8. McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer review. *JAMA*. 1990;263:1371-1376.
9. Fisher M, Friedman SB, Strauss B. The effects of blinding on acceptance of research papers by peer review. *JAMA*. 1994;272:143-146.
10. Jadad AR, Moore A, Carroll D, et al. Assessing the quality of reports of randomized clinical trials. *Control Clin Trials*. 1996;17:1-12.
11. van Rooyen S, Godlee F, Evans S, et al. Effect of blinding and unmasking on the quality of peer review. *JAMA*. 1998;280:234-237.
12. Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding peer reviewers and asking them to sign their reports. *JAMA*. 1998;280:237-240.
13. Justice AC, Cho MK, Winker MA, et al. Does masking author identity improve peer review quality? *JAMA*. 1998; 280:240-242.
14. van Rooyen S, Godlee F, Evans S, et al. Effect of open peer review on quality of reviews and on reviewers' recommendations. *BMJ*. 1999;318:23-27.
15. Das Sinha S, Sahni P, Nundy S. Does exchanging comments of Indian and non-Indian reviewers improve the quality of manuscript reviews? *Nat Med J India*. 1999;5:210-213.
16. Walsh E, Rooney M, Appleby L, Wilkinson G. Open peer review. *Br J Psychiatry*. 2000;176:47-51.
17. Jefferson T, Smith R, Yee Y, et al. Evaluating the *BMJ* guidelines for economic submissions. *JAMA*. 1998;280:275-277.
18. Gardner MJ, Bond J. An exploratory study of statistical assessment of papers published in the *British Medical Journal*. *JAMA*. 1990;263:1355-1357.
19. Bingham CM, Higgins G, Coleman R, Van der Weyden MB. The *Medical Journal of Australia* Internet peer-review study. *Lancet*. 1998;352:441-445.
20. Neuhauser D, Koran CJ. Calling *Medical Care* reviewers first. *Med Care*. 1989;27:664-666.
21. Callaham ML, Wears RL, Waeckerle JF. Effect of attendance at a training session on peer reviewer quality and performance. *Ann Emerg Med*. 1998;32:318-322.
22. Strayhorn J, McDermott JF Jr, Tanguay P. An intervention to improve the reliability of manuscript reviews for the *Journal of the American Academy of Child and Adolescent Psychiatry*. *Am J Psychiatry*. 1993;150:947-952.
23. Ernst E, Resch KL. Reviewer bias against the unconventional? *Complement Ther Med*. 1999;7:19-23.
24. Elvik R. Are road safety evaluation studies published in peer reviewed journals more valid than similar studies not published in peer reviewed journals? *Accid Anal Prev*. 1998;30:101-118.
25. Goodman SN, Berlin J, Fletcher SW, Fletcher RH. Manuscript quality before and after peer review and editing at *Annals of Internal Medicine*. *Ann Intern Med*. 1994;121:11-21.
26. Pierie JP, Walvoort HC, Overbeke AJ. Readers' evaluation of effect of peer review and editing on quality of articles in the *Nederlands Tijdschrift voor Geneeskunde*. *Lancet*. 1996;348:1480-1483.
27. Jefferson T, Wager E, Davidoff F. Measuring the quality of editorial peer review. *JAMA*. 2002;287: 2786-2790.

Measuring the Quality of Editorial Peer Review

Tom Jefferson, MD
Elizabeth Wager, MA
Frank Davidoff, MD

Context The quality of a process can only be tested against its agreed objectives. Editorial peer-review is widely used, yet there appears to be little agreement about how to measure its effects or processes.

Methods To identify outcome measures used to assess editorial peer review as performed by biomedical journals, we analyzed studies identified from 2 systematic reviews that measured the effects of editorial peer review on the quality of the output (ie, published articles) or of the process itself (eg, reviewers' comments).

Results Ten studies used a variety of instruments to assess the quality of articles that had undergone peer review. Only 1, nonrandomized study compared the quality of articles published in peer-reviewed and non–peer-reviewed journals. The others measured the effects of variations in the peer-review process or used a before-and-after design to measure the effects of standard peer review on accepted articles. Eighteen studies measured the quality of reviewers' reports under different conditions such as blinding or after training. One study compared the time and cost of different review processes.

Conclusions Until we have properly defined the objectives of peer-review, it will remain almost impossible to assess or improve its effectiveness. The research needed to understand the broader effects of peer review poses many methodologic problems and would require the cooperation of many parts of the scientific community.

JAMA. *2002;287:2786-2790* www.jama.com

A FUNDAMENTAL TENET OF ALL scientific and scholarly work is that every aspect of it must be subjected to critical appraisal; only those findings and principles that withstand such appraisal become established. Although much appraisal occurs as work is in progress (and some after it has been published), work that is submitted for publication undergoes critical appraisal,

Author Affiliations: Health Reviews Ltd, Rome, Italy (Dr Jefferson); Sideview, Princes Risborough, England (Ms Wager); *Annals of Internal Medicine*, Philadelphia, Pa (Dr Davidoff).
Corresponding Author and Reprints: Ms Liz Wager, Sideview, Station Road, Princes Risbor-ough, HP27 9DE, UK, (e-mail: liz@sideview.demon.co.uk)

Table A. Search Strategy for Identification of Studies

Databases Searched*

Australasian Medical Index, 1980-2000 (searched November 2000)
Best Evidence 4, 2000 (searched July 2000)
BIOETHICSLINE, 1973-2000 (searched July 2000)
CINAHL, 1997-1999 (searched July 2000)
Cochrane Library, 2000, issue 2 (searched November 2000, including Cochrane Controlled Trials, Cochrane Database of Systematic Reviews, Cochrane Methods Register, and Database of Abstracts of Reviews of Effectiveness)
Current Contents, 1999-2000 (searched February 2000)
Dissertation Abstracts, 1861-2000 (searched February 2000)
EMBASE, 1989-1993; January 1998–January 2000 (searched July 2000)
HealthSTAR, 1975-1999 (searched December 1999)
MEDLINE, 1966–February 2000 (searched by Anne Lusher, United Kingdom Cochrane Center, for the Cochrane Methodology Register, February 2000)
National Research Register (searched March 2000)
PsycLIT, 1887-2000 (searched February 2000)
PubMed, 1998-2000 (searched February 2000); 1997 (searched March 2000)
PubSCIENCE, 1998–January 2000 (searched February 2000)
SIGLE, 1980–June 1999 (searched March 2000)

Other Sources Searched

Reference lists of relevant articles
JAMA special issues on peer review[4-6]
Posting on European Assocation of Science Editors and World Association of Medical Editors Web sites (http://www.ease.org.uk and http://www.wame.org)
Locknet Web site (last searched September 2000)

*MEDLINE was searched for Medical Subject Heading exploded and text term *peer review*; all other databases were searched for text words *peer review*.

Table B. Descriptive Summary of Studies Included in the Review*

Reference	Study Question	Sample	Design	Methods	Findings	Comments
Studies Assessing Effect of Blinding/Masking on Quality of External Opinions						
McNutt et al,[8] 1990	Effect of blinding on review quality	127 Submissions to *Journal of General Internal Medicine*	RCT	One reviewer for each MS knew authors' identity; the other was masked. Signing was optional. Editors and authors rated review quality.	No association between signing and review quality was found, but quality of reviews was higher for blinded manuscripts.	Success of blinding was tested by asking reviewers and editors to guess authors' identity.
Fisher et al,[9] 1994	Whether reviewers are biased toward well-known authors	57 Submissions to *Journal of Developmental and Behavioral Pediatrics*	RCT	Editors assessed quality of blinded and nonblinded reviews. Masked reviewers were asked to guess author identity.	Reviewers correctly guessed author identity in 46% of cases. Blinded reviewers gave better scores to the work of authors with longer publication records. The authors concluded that blinded review produces less-biased reviews.	Small RCT with robust randomization; the weakness of the blind raises questions about its feasibility in small specialties
Jadad et al,[10] 1996	Effect of blinding on rating of RCTs	36 Reports of RCTs in the field of pain control	RCT	Seven blinded reviewers made assessments and 7 did so under open conditions.	Blinded assessment produced more consistent results than open. Blinded reviewers rated articles more severely, resulting in lower-quality scores.	Small RCT with robust randomization but no attempt to measure success of blinding
van Rooyen et al,[11] 1998	Effect of knowing coreviewers' and authors' identities on review quality	527 Submissions to *BMJ*	RCT	Reviewers randomized to be blinded or unblinded to coreviewers' and authors' identities; review quality assessed on 7-item 5-point scale and time taken	Knowledge of one or both reviewers' or authors' identities made no difference to review quality, nature of recommendations, or time taken. Blinding was unsuccessful in 42% due to small research fields and authors referencing their own work.	Well-designed and -conducted study; hampered by difficulty of successful blinding
Godlee et al,[12] 1998	Effects of blinding reviewers to authors' identities and of signing reports	221 Reviews of submissions to *BMJ*	4-Arm RCT with different types of disclosure of reviewers and authors plus a control arm	Review quality based on ability to detect errors inserted in manuscript	No effect of open peer review on review quality was shown. Most reviewers failed to detect the majority of most errors inserted into manuscript.	Well-designed study but 53% reviewer response rate raises the issue of bias
Justice et al,[13] 1998	Effect of masking authors' identity on review quality	118 Submissions to *Annals of Emergency Medicine, Annals of Internal Medicine, JAMA, Obstetrics and Gynecology,* and *Ophthalmology*	RCT	Normal practice (blinded review) vs one reviewer blinded to author's identity and the other unblinded; review quality judged on 4-item, 5-point scale	Blinding had no effect on review quality. Submissions of well-known authors were more difficult to blind successfully (blinding success was 68%).	Well-designed study with robust randomization; relatively low blinding success rate cannot exclude possibility of bias
van Rooyen et al,[14] 1999	Effect of open peer review on review quality; reviewers' acceptance of open review.	113 Submissions to *BMJ*	RCT	The identity of consenting reviewers was revealed when sending the review to the authors.	Asking reviewers' consent to identification did not affect review quality, time taken, or recommendations to publish; however, a significant number of reviewers objected to open review.	Well-designed and -conducted study with a robust randomization process
Das Sinha et al,[15] 1999	Whether non-Indians produce better reviews	78 Submissions to *Nature Medicine Journal of India*	Single-blind RCT	Each submission randomly assigned to an Indian and a non-Indian reviewer; reviews assessed by editors unaware of reviewer nationality	Non-Indian reviewers produced better-quality reviews than Indians. Exchanging reviews among reviewers did not was not found to affect quality.	Robust randomization, but no details given on 22 submissions that could not be included in the study
Walsh et al,[16] 2000	Feasibility of open peer review	Reviews of 408 submissions to *British Journal of Psychiatry*	RCT of signed vs unsigned reviews	All reviewers were aware of authors' identity. Review quality and tone were assessed.	Signed reviews were of higher quality and more courteous. The signed group spent more time on reviews and were also less likely to recommend rejection. The authors concluded that an open system is feasible.	Randomization methods unclear (intervention and control groups are uneven; n = 222 vs n = 186); blinding appeared robust

(continued)

Table B. Descriptive Summary of Studies Included in the Review* (cont)

Reference	Study Question	Sample	Design	Methods	Findings	Comments
Studies Assessing Effects of Submission Checklists on Outcomes						
Jefferson et al[17] 1998	Effect of publishing guidelines on reviewing economic submissions	192 Submissions to *BMJ* and *The Lancet*	Before-and-after study with economic submissions to *The Lancet* acting as a control	Submission quality was assessed using 36-item checklist and questionnaire to editors.	Publishing guidelines on economic submissions had no apparent impact on submission quality but helped editors manage submissions	The open design and number of untraceable submissions makes interpretation difficult.
Gardner and Bond,[18] 1990	Usefulness of checklist for statistical assessment of submitted manuscripts	45 Submissions to *BMJ*	Before-and-after study	12- And 24-item referee checklists	Only 5 submitted manuscripts (11%) were methodologically acceptable of 38 (84%) after publication. Using a checklist can increase the statistical quality of published studies.	Generalizability of results may be limited. Lack of a control arm means that bias or a Hawthorne effect cannot be excluded.
Studies Assessing Effects of Communication Media on Outcomes						
Bingham et al,[19] 1998	Effect of open Internet peer review on review quality	56 Submissions accepted by *Medical Journal of Australia*	Open comparative study	Comments were invited on manuscripts and referees' reports and posted on journal's Web site.	Few readers submitted comments to Web site, but authors changed 7 manuscripts in response. No change in quality of commissioned reviews was noted.	Descriptive study of uncertain generalizability
Neuhauser and Koran,[20] 1989	Effect of telephone warning regarding review turnaround time	177 Submissions to *Medical Care*	RCT	Submissions were randomized to reviewers who received phone warning before review was sent vs those who did not.	Calling reviewers before sending out manuscript increased costs and total turnaround time, although warned reviewers returned reviews more quickly.	Unclear randomization method; generalizability may be limited
Studies Assessing Effects of Training on Outcomes						
Callaham et al,[21] 1998	Effect of attending voluntary training workshops on review quality	39 Reviewers of *Annals of Emergency Medicine*	Controlled trial	Editors assessed quality of reviews from reviewers who attended a 4-hour training workshop, reviewers matched for pretraining review quality and experience, and 220 unmatched controls.	There was no significant change in any performance measure after reviewers attended a training workshop.	The self-selected nature of attendees at the training workshop and the absence of blinding raise the issue possibility of bias.
Strayhorn et al,[22] 1993	Effect of introducing new evaluation scales on review quality	568 Reviews for *Journal of the American Academy of Child and Adolescent Psychiatry*	Before-and-after study	New scales had more separate discrete items and training manuals.	Interrater reliability increased after the introduction of the evaluation scales.	The open nature of the design makes the interpretation of its findings difficult.
Studies Assessing Presence and Effects of Reviewer Bias on Outcomes						
Ernst and Resch,[23] 1999	Reviewer bias toward unconventional drugs	130 Reviews from physicians invited to review for a fictitious journal	Double-blind RCT	Two versions of a letter to the editor were reviewed. Versions were identical except for study drug (metoprolol or beef spleen cell extract).	No evidence of reviewer bias toward unconventional drug but poor interrater reliability overall	Robust randomization but few details of blinding; reviewer response rate was 61%
Elvik,[24] 1998	Whether peer-reviewed journals publish more valid reports than non–peer-reviewed journals	123 Road safety evaluation studies	Open comparison study	Quality of 44 studies in peer-reviewed journals was compared with 79 similar reports published in non–peer-reviewed journals.	No significant difference found between validity of studies published in peer-reviewed and non–peer-reviewed journals. Studies from authors affiliated with universities obtained higher average validity scores.	This is the only identified study addressing the effects of peer review on validity. The author did attempt to adjust for potential confounders.

(continued)

Table B. Descriptive Summary of Studies Included in the Review* (cont)

Reference	Study Question	Sample	Design	Methods	Findings	Comments
Studies Assessing Effects of Peer Review on Study Report Quality						
Goodman et al,[25] 1994	Effects of peer review and editing on quality	111 Manuscripts accepted by *Annals of Internal Medicine*	Before-and-after study	Expert reviewers assessed manuscript quality using a 34-item checklist.	Peer review and editing improved the quality of manuscripts overall and particularly in the use of confidence intervals, discussion of study limitations, and appropriateness of conclusions.	The study relates to a journal that invests heavily in editing. Expert reviewers' views may not reflect those of ordinary readers.
Pierie et al,[26] 1996	Effects of peer review and editing on quality	50 Submissions to the *Dutch Journal of Medicine*	Retrospective study	Volunteer assessors drawn from journal readers compared the quality of submitted, accepted, and published versions of published articles. Each assessor received a pair of typescripts and was unaware of their status.	Quality scores improved between submission and acceptance and between acceptance and publication, especially for items relating to style and readability.	Volunteer assessors (who responded to an advertisement in the journal) not to be representative of all readers

*RCT indicates randomized controlled trial.