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 Risborough, HP27 9DE, UK, (e-mail: liz@sideview.demon.co.uk)

known as peer review, as part of the editorial process.

Editorial peer review is therefore an extension of the basic principles of science and scholarship. It has existed for more than 200 years[1] and has achieved near universal application for assessing research reports before publication. Despite its wide acceptance, peer review has been subjected to a variety of criticisms,[2] and, indeed, surprisingly little is known about its effects on the quality and utility of published information,[3] much less about its beneficial or adverse social, psychological, or financial effects.

The same can be said about critical appraisal in scholarly work generally. However, uncertainty about the effects of peer review is not simply a matter for academic concern. Clinical decisions must be made on the best available evidence, usually systematic reviews and meta-analyses, but these can be misleading if they are based on invalid, incomplete, inaccurate, or duplicate information, or if the review articles themselves are poorly done. Any process affecting the assessment and dissemination of clinical evidence therefore has a direct bearing on patient care.

In this article we review the criteria used by others to measure the effects of peer review, consider what this implies about the aims of peer review, especially in relation to clinical evidence, and suggest ways in which its effects might be measured more rigorously.

METHODS

In 2 systematic reviews of the effects of editorial review and technical editing, we identified published articles that evaluated the peer review process and identified the criteria used in those studies to evaluate peer review. Our first review considered processes that occur between submission of a paper and a decision on publication; the second considered the processes that occur between acceptance and publication. Both systematic reviews were performed using Cochrane methodology. The methods and primary findings of the reviews are published elsewhere.[3-6]

RESULTS

We included 19 studies in our systematic review of the effects of peer review; these are described separately.[3] Two studies were identified from our review on technical editing since they included information about changes that occur to papers between submission and acceptance or did not distinguish the preacceptance and postacceptance processes.[5] We identified 8 other studies that measured the quality of papers or reviews but did not compare peer-review processes. The outcome measures used in these studies are shown in online Table A (http://www.jama.com).[7-35] Brief descriptions of the 8 studies not described in companion papers are also shown.

Ten studies measured various aspects of the quality of papers that had undergone peer review. Only one study[8] compared the quality of articles published in peer-reviewed and non–peer-reviewed journals, but it used a nonrandomized design and the findings may have been confounded by other factors, such as differences in the quality of studies submitted to the different journals. The other studies 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 in a particular journal. A major limitation of most studies is that they assessed the quality only of accepted papers, and measured the changes that took place between submission or acceptance and publication. Only the studies of economic submissions[12] and statistical quality[9] included papers that were rejected by the target journal. Of the 10 studies, only 2[13,14] used journal readers to assess the quality of papers, the others were based on editors' assessment. Virtually every study used its own rating instrument. These included between 7 and 36 items rated using 2- to10-point scales. Most scales appeared to be unvalidated but, in 1 case when the scoring system was tested, it was found to have low reliability.[10] The 2 studies of readability[7,15] used published scales that have not been validated for use in this setting.

Eighteen studies measured the quality of reviewers' reports under different conditions such as blinding or training.[17-34] Three of these included an assessment by the authors whose work had been reviewed.[27,28,33] The others used editors to judge the quality of reviews. Instruments used to rate review quality ranged from 2- to 10-item scales, most were rated using a 2- to 5-point system, but 1 used a visual analog scale, and 1 used ratings from 1 to 100. One of these scales had a published validation.[36] Four studies examined the amount of agreement among reviewers,[30,31] between reviewers and editors[19] or between reviewers and readers.[27] One study compared the time and cost of different review processes.[35]

The aspects of reviews most commonly rated were those relating to the methodological soundness of the reviewed study, its importance, originality and presentation. Several studies also attempted to assess the tone or courteousness of the review. One study measured the number of errors that a review detected.[25] Three considered the speed of review.[11,23,35] Aspects of articles examined were more wide ranging, including quality assessments of each section (introduction, methods, results, and discussion) and also subjective measures of the article's relevance, overall quality, readability, and comprehensibility.

COMMENT

Analysis of published studies on editorial peer review reveals the diversity of study questions and end points. This suggests that peer review is expected to have a wide range of effects, that its true effects have not been determined, or that the aims of peer review have not been identified properly. Our review also showed that the term *peer review* is used to describe a number of processes, most commonly gathering opinions from external experts, but also review by in-house editors, and that it may not always be possible to make a clear distinction between peer review and technical editing.

Based on our reviews of studies and the larger literature of opinion about

peer review, we suggest that its aims may be categorized as (1) selecting submissions for publication (by a particular journal) and rejecting those with irrelevant, trivial, weak, misleading, or potentially harmful content, and (2) improving the clarity, transparency, accuracy, and utility of the selected submissions.

The selection of submissions depends on assessment of their quality and how well they match the journal's scope and aims. The quality criteria may be categorized as the importance, relevance, usefulness, and methodological and ethical soundness of the research and the clarity, accuracy, and completeness of the report.

The main purpose of medical research is to improve health or the delivery of health care. If peer review is regarded as one stage in this process, it might be expected to have measurable effects on health status. However, outcomes such as this are difficult to assess because they are affected by numerous other factors. Surrogate outcomes, such as process measures, are much easier to assess, but may not provide a reliable measure of more meaningful indicators of success.

In the TABLE we summarize the possible effects of editorial peer review on the quality of reports of clinical research, provide definitions for these, and suggest indicators that could be used to assess them.

Research so far has measured only certain aspects of peer review, largely focused on variations in the processes used rather than comparing the effects of peer review with those of other systems. Given the resources spent on peer review and the importance placed on it, this is unsatisfactory even though it may reflect the fact that no part of the process of scientific evaluation has been rigorously studied. Ironically, however, the fact that peer review is so well entrenched makes it harder to study, since scientists and editors may be unwilling to take part in randomized studies if they believe that the current system serves them well. How, therefore, should the scientific community proceed in its evaluation of peer review?

Ideally, this would be assessed by large-scale, long-term research into 2 cohorts of studies, randomized to undergo either peer review or an alternative method of assessment, such as random selection for publication. Given the complexity of factors at play a multivariate analysis may be necessary. However, researchers might not be prepared to accept such randomization, and knowledge of the trial could bias the results. It would be important to ensure that both groups of studies were of equal average quality—for example by examining submissions to a single journal. The follow-up period would have to be lengthy to allow for changes in health status or health care delivery to occur as a consequence of publication.

Another difficulty in studying the effects of peer review is that the quality components of a manuscript are often interlinked, and it is meaningless to study them in isolation. For example, a methodologically flawed study or incomplete report will detract from the publication's usefulness.

Table. Indicators of the Quality of the Output of Editorial Peer-Review of Clinical Studies and Methods to Assess Them

Outcome and Definition	Ideal Indicator	Surrogate Indicators
Important		
Study findings have a major impact on health or health care	Changes in health status Changes in health care delivery	Citation rates Media coverage Correspondence
Useful		
Study contributes significantly to the scientific debate or knowledge on a subject	Contributes significantly within a systematic review of the topic Narrows confidence intervals around estimates of effect	Contributes to nonsystematic reviews or guidelines Citation rates Correspondence
Relevant		
Topic is relevant to the journal's aims and readers	Topic is relevant and consistent with the aims and readership of the journal confirmed by survey	Citation rates Correspondence Internet hit rates
Methodologically sound		
Methods used are able to answer the study question	Study findings are replicated several times across different settings	Closeness of fit between methods and "evidence-based" methodological checklist Correspondence
Ethically sound		
Unnecessary harm to humans or animals has been avoided Study has been carried out and reported honestly	No divergence between reality and the report Rights of humans and animals safeguarded Privacy and informed consent maintained throughout Raw data match presented data Number preference check is negative	Study received ethical clearance No complaints from participants No duplicate publication
Complete		
All relevant information is presented	There is no selective presentation of data All relevant references are cited	The text is complete The publication is complete (ie not salami-sliced)
Accurate		
Presented information is a true reflection of what went on	Measurements truly reflect magnitude of findings Raw data match presented data References are accurate	The figures add up Corrections

*All other studies are described in Jefferson et al[2] and Wager and Middleton,[5] which are published in this issue of *JAMA*.

If the scientific community could agree on the objectives of peer review and collaborate in the assessment of its effects, we could start identifying some of the practices for which evidence of effect is better than that of controls. We propose that the following questions should be tested collaboratively across journal settings:

- Does peer review identify submissions of higher quality than other selection methods (or chance)?
- Does peer review improve the clarity, transparency, accuracy, and utility of published papers meaningfully beyond that of the submitted version?

This would involve assessing the quality of both published and unpublished submissions using well-validated instruments. We suggest that measures of quality should include the importance, relevance, usefulness, and methodological and ethical soundness of clinical studies. Such research would also involve tracking submissions between journals since submissions rejected by one journal often go on to be published by another.[37]

Although there is some evidence that peer review and editing improves articles between submission and publication, the effectiveness of its selection and filtering functions remains virtually untested.[3,5] Yet, despite this lack of evidence, peer review is well established in most academic disciplines. It is therefore possible that peer review is retained for different reasons than those stated. For example, it may serve to protect journals' reputations or to provide acceptability for commercially-funded studies. Using unpaid reviewers probably reduces some aspects of the work of in-house editors, although it also carries administrative costs. Peer review is also so well established that it has become part of the system for assessing academic merit in appointments and promotions. The broader functions of peer review, including its social and psychological effects such as increasing the credibility and prestige of published work, are rarely acknowledged and have not, to our knowledge, ever been seriously studied.

CONCLUSIONS

Given the widespread use of peer review, it is surprising that so little is known of its aims or effects although the same might be said of several other, well-established processes of scientific appraisal. The financial costs of peer review to the scientific community are difficult to estimate but should not be ignored.[38] There is also anecdotal evidence that peer review has shortcomings and may even have harmful effects.[2] Yet, until we have properly defined the aims of peer review it will remain almost impossible to estimate the effectiveness of the process or to improve it systematically.

The research needed to evaluate the effects of peer review poses many methodological problems and would require the cooperation of large numbers of authors and editors. The growth of electronic publishing has increased the urgency of establishing an effective and efficient system for evaluating scientific information but may also offer opportunities to explore alternatives to the current peer-review system.[39] Until such research is undertaken, the ability of peer review to improve the quality of published research and, ultimately, improve the dissemination of reliable health information will remain uncertain.

Author Contributions: *Study concept and design:* Jefferson.
Acquisition of data: Jefferson, Wager.
Analysis and interpretation of data: Jefferson, Wager, Davidoff.
Drafting of the manuscript: Wager, Davidoff.
Critical revision of the manuscript for important intellectual content: Jefferson, Davidoff.
Study supervision: Jefferson
Disclosure: Dr Tom Jefferson co-edited the book *Peer Review in Health Sciences*, Ms Wager wrote 2 chapters in the book, and they are co-authors of a book entitled *How to Survive Peer Review*. Dr Davidoff was the editor of a peer-reviewed journal. All authors are active peer reviewers and have published articles in peer-reviewed journals.
Acknowledgment: We thank Philip Alderson and Philippa Middleton for their contributions to the original systematic reviews, Iain Chalmers and John Overbeke for helpful comments on early drafts of the manuscript, and Fiona Godlee for her review of the submitted version.

REFERENCES

1. Kronick DA. Peer-review in 18th-century scientific journalism. *JAMA*. 1990;263:1321-1322.
2. Wager E, Jefferson T. The shortcomings of peer review. *Learned Publishing*. 2001;14:257-263.
3. Jefferson T, Alderson P, Wager E, Davidoff F. Effects of editorial peer review: a systematic review. *JAMA*. 2002;287:2784-2786.
4. Alderson P, Davidoff F, Jefferson TO, Wager E. Editorial peer review for improving the quality of reports of biomedical studies [protocol for Cochrane Methodology Review on CD ROM]. Oxford, England: Cochrane Library, Update Software; 2001;issue 3.
5. Wager E, Middleton P. Technical editing in biomedical journals. *JAMA*. 2002;287:2821-2824.
6. Wager E, Middleton P. Technical editing of research reports in biomedical journals [Protocol for a Cochrane Methodology Review on CD ROM]. Oxford, England: The Cochrane Library Update Software; 2001;issue 3.
7. Biddle C, Aker J. How does the peer review process influence *AANA Journal* article readability? *AANA J*. 1996;64:65-68.
8. 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.
9. Gardner MJ, Bond J. An exploratory study of statistical assessment of papers published in the *British Medical Journal*. *JAMA*. 1990;263:1355-1357.
10. 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.
11. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, Moher D. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. *JAMA*. 1998;280:278-280.
12. Jefferson T, Smith R, Yee Y, Drummond M, Pratt M, Gale R. Evaluating the *BMJ* guidelines for economic submissions: prospective audit of economic submissions to *BMJ* and *The Lancet*. *JAMA*. 1998;280: 275-277.
13. Justice AC, Berlin JA, Fletcher SW, Fletcher RH, Goodman SN. Do readers and peer reviewers agree on manuscript quality? *JAMA*. 1994;272:117-119.
14. 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.
15. Roberts JC, Fletcher RH, Fletcher SW. Effects of peer review and editing on the readability of articles published in *Annals of Internal Medicine*. *JAMA*. 1994; 272:119-121.
16. Rochon PA, Gurwitz JH, Cheung CM, Hayes JA, Chalmers TC. Evaluating the quality of articles published in journal supplements compared with the quality of those published in the parent journal. *JAMA*. 1994;272:108-113.
17. Bingham CM, Higgins G, Coleman R, Van der Weyden MB. The *Medical Journal of Australia* Internet peer-review study. *Lancet*. 1998;352:441-445.
18. Blank RM. The effects of double-blind versus single-blind reviewing; experimental evidence from the *American Economic Review*. *Am Econ Rev*. 1991;81: 1041-1067.
19. Callaham ML, Wears RL, Waeckerle JF. Effect of attendance at a training session on peer reviewer quality and performance. *Ann Emerg Med*. 1998;32(3 pt 1):318-322.
20. Das Sinha S, Sahni P, Nundy S. Does exchanging comments of Indian and non-Indian reviewers im-

prove the quality of manuscript reviews? *Natl Med J India*. 1999;12:210-213.
21. Ernst E, Resch KL. Reviewer bias against the unconventional? a randomized double-blind study of peer review. *Complement Ther Med*. 1999;7:19-23.
22. Ernst E, Resch KL. Reviewer bias: a blinded experimental study. *J Lab Clin Med*. 1994;124:178-182.
23. Feurer ID, Becker GJ, Picus D, Ramirez E, Darcy MD, Hicks ME. Evaluating peer reviews: pilot testing of a grading instrument. *JAMA*. 1994;272:98-100.
24. Fisher M, Friedman SB, Strauss B. The effects of blinding on acceptance of research papers by peer review [published correction appears in *JAMA*. 1994; 272:1170]. *JAMA*. 1994;272:143-146.
25. Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding peer reviewers and asking them to sign their reports: a randomized control trial. *JAMA*. 1998;280:237-240.
26. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17:1-12.
27. Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D. Does masking author identity improve peer review quality? a randomized controlled trial. *JAMA*. 1998;280:240-242.
28. McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer-review. a randomized trial. *JAMA*. 1990;263:1371-1376.
29. Nylenna M, Riis P, Karlsson Y. Multiple blinded reviews of the same two manuscripts: effects of referee characteristics and publication language. *JAMA*. 1994;272:149-151.
30. Oxman AD, Guyatt GH, Singer J, et al. Agreement among reviewers of review articles. *J Clin Epidemiol*. 1991;44:91-98
31. Strayhorn J Jr, 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.
32. van Rooyen S, Godlee F, Evans S, Smith R, Black N. Effect of blinding and unmasking on the quality of peer review: a randomized trial. *JAMA*. 1998;280: 234-237.
33. van Rooyen S, Godlee F, Evans S, Black N, Smith R. Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *BMJ*. 1999;318:23-27.
34. Walsh E, Rooney M, Appleby L, Wilkinson G. Open peer review: a randomised controlled trial. *Br J Psychiatry*. 2000;176:47-51.
35. Neuhauser D, Koran CJ. Calling *Medical Care* reviewers first: a randomized trial. *Med Care*. 1989;27: 664-666.
36. van Rooyen S, Black N, Godlee F. Development of the review quality instrument (RQI) for assessing peer reviews of manuscripts. *J Clin Epidemiol*. 1999; 52:625-629.
37. Ray J, Berkwits M, Davidoff F. The fate of manuscripts rejected by a general medical journal. Am J Med. 2000;109:131-135.
38. Donovan B. The truth about peer review. *Learned Publishing*. 1998;11:179-184.
39. Bingham C. Peer review on the internet: are there faster, fairer, more effective methods of peer review? In: Godlee F, Jefferson TO, eds. *Peer Review in Health Sciences*. London, England: BMJ Books; 1999; 205-223.

Author Perception of Peer Review

Impact of Review Quality and Acceptance on Satisfaction

Ellen J. Weber, MD
Patricia P. Katz, PhD
Joseph F. Waeckerle, MD
Michael L. Callaham, MD

Context To determine author perception of peer review and association between quality of review and author satisfaction.

Methods Survey between May 1999 and October 2000 of 897 corresponding authors of manuscripts under consideration by the *Annals of Emergency Medicine* and had received final editorial decisions during the study period. A total of 576 authors (64%) returned the survey. Using a 5-point Likert scale, the survey assessed differences in satisfaction between authors whose manuscripts were accepted, reviewed and rejected, and rejected without full review. The association of author satisfaction with editor's assessment of review quality, publication decision, author sex, specialty, and publication experience were also assessed.

Results Overall mean (SD) satisfaction score, indicated by agreement with "My experience with the review process will make me more likely to submit to *Annals* in the future," was 3.1 (1.0) and was significantly higher among authors of accepted papers (3.7 [0.9]) than among either group of rejected papers (rejected/reviewed, 2.8 [1.0]; rejected/no review, 3.0 [0.9]; *P*.05). Authors whose manuscripts were reviewed and rejected were the least satisfied with the time to decision (rejected/reviewed, 3.0 [1.2] vs accepted, 3.7 [1.0] and rejected/no review, 3.9 [0.9]; $P<.05$). Those whose papers were rejected without review were the least satisfied with the letter explaining the editorial decision (rejected/no review, 2.8 [1.2] vs accepted, 4.2 [0.7] and rejected/reviewed, 3.1 [1.2]; $P<.05$). Among respondents whose manuscripts underwent full review (accepted and rejected/reviewed), overall satisfaction was highly associated with acceptance of the manuscript for publication (odds ratio [OR], 6.12; 95% confidence interval [CI], 3.43-10.91) but not with quality rating of reviews (OR, 1.26; 95% CI, 0.84-1.90).

Conclusion Contributor satisfaction with peer review was modest. Authors of rejected manuscripts were dissatisfied with the time to decision and communication from the editor. Author satisfaction is associated with acceptance but not with review quality.

JAMA. 2002;287:2790-2793 www.jama.com

PEER REVIEW IS A RESOURCE-intensive process relying on considerable, chiefly volunteer, effort to evaluate manuscripts for publication and craft objective and constructive reviews. However, little is known about how author's experience the peer-review process and, in particular, whether the quality of the reviews affects their satisfaction. Previous studies suggest that prestige and circulation are the factors frequently used by researchers in determining the journal to which they submit their work, while other aspects of the process—quality of the journal's peer-review panel, likelihood of acceptance, turnaround time, and biostatistical review—have less influence on their choice.[1,2]

The *Annals of Emergency Medicine* has conducted a number of studies and initiatives to monitor and improve the quality of its review process.[3,4] We conducted a survey of authors who submitted manuscripts to *Annals* to un-

Author Affiliations: Division of Emergency Medicine (Drs Weber and Callaham) and Institute for Health Policy Studies (Dr Katz), Department of Medicine, University of California, San Francisco; and Department of Emergency Medicine, Baptist Medical Center, Menorah Medical Center, School of Medicine, University of Missouri, Kansas City (Dr Waeckerle).
Corresponding Author and Reprints: Ellen J. Weber, MD, Emergency Department, University of California San Francisco, 505 Parnassus Ave, Box 0208, San Francisco, CA 94143-0208 (e-mail: weber@medicine.ucsf.edu).

Table A. Quality Criteria and Rating Methods Used in Published Studies of Editorial Peer Review

Study, y	Quality Criteria	Rating Method
	Studies of the Quality of Papers After Peer Review	
Biddle and Aker,[7] 1996	Readability	Gunning and Flesch scores
Elvik,[8] 1998	Sampling technique, sample size, inclusion criteria, study design, control for confounders	7-Item, 5-point validity scale
Gardner,[9] 1990	Study design; features, analysis and presentation of statistical data, and recommendations	12- and 24-Item checklists for general papers and clinical trials respectively, scored: yes/no/ unclear
Goodman,[10] 1994	34 Items including study aims, design, blinding, drop-outs, statistics, discussion of limitations, and appropriate conclusions	34-Item, 5-point scale
Jadad et al,[11] 1998	No. of authors, patients and trials, sources of trials, description of outcomes, inclusion/exclusion criteria, pooled effect estimates, heterogeneity testing, assessment of trial quality, and language restrictions	Binary and continuous outcomes in10-item list
Jefferson et al,[12] 1998	Acceptance rate and quality	36-Item checklist based on economic principles
Justice et al,[13] 1994	18-Items including clarity, conciseness and readability; quality of abstract and title, enjoyability, and relevance of article; and importance of question, originality, and scientific validity	18- or 5-item, 5-point questionnaire for readers and peers respectively 10-point summary grade
Pierie et al,[14] 1996	23 Items including description of study design, inclusion criteria, statistics, discussion of limitations, understandability, importance, and general medical value	25-Item, 5-point questionnaire
Roberts et al,[15] 1994	Readability	Gunning and Flesch scores
Rochon et al,[16] 1994	Robustness of allocation, randomization, blinding, statistical analysis, handling of drop-outs, discussion of adverse effects	14-Item, 3-or 4-point scale
	Studies of the Quality of Reviewers' Comments	
Bingham et al,[17] 1998	Timeliness, completion of grade sheet, etiquette, section-by section review, offering supporting references, summary, new insights offered	7-Item, 2-point checklist
Blank,[18] 1991	Theoretical exposition, empirical analysis, contribution to the field, overall quality	4-Item, 5-point scale
Callaham et al,[19] 1998	Congruence of reviewers' recommendations with editorial decision and acceptance rate of advice	2-Item, binary scale
Das Sinha et al,[20] 1999	Importance of question, target key issues, methods, presentation and general impression	5-Items scored from 1 to 100
Ernst and Resch,[21] 1999	Relevance, hypothesis-formulation, randomization, inclusion/exclusion criteria, sample size, statistics, outcome choice, follow-up, clarity, and linguistic	10-Item VAS 3-point scale for overall scores
Ernst and Resch,[22] 1994	Adequacy of study design, patient description, statistical methods, end points and linguistic merit	5-Item questionnaire
Feurer et al,[23] 1994	Timeliness, grade sheet completion, etiquette, sectional narratives, citations, narrative summary, and insights	7-Item, 3-point scale
Fisher et al,[24] 1994	Reviewers' recommendation on acceptance (accept, optional revision, mandatory revision, reject with extra review, outright rejection)	5-Point scale
Godlee et al,[25] 1998	Ability of reviewers to identify 8 deliberate errors inserted in manuscripts	Proportion of errors identified
Jadad et al,[26] 1996	Randomization, blinding, withdrawals, and drop-outs	6-Item and 3-item instruments
Justice et al,[27] 1998	Importance of research question, methodologic soundness, courteousness, and reviewer's production of evidence to support their views	4-Item, 5-point scale
McNutt,[28] 1990	Editors rated reviews for covering importance of question, key issues, methods, and presentation Authors rated whether review was thorough, constructive, fair, courteous and knowledgeable and gave an overall grade	4- or 6- item, 5-point scale
Nylenna et al,[29] 1994	Importance, originality, methods, discussion, conclusion, structure, readability, length	8-Item, 5-point scale plus summary scores
Oxman et al,[30] 1991	Search methods, inclusion criteria, selection bias avoidance, validity criteria and assessment, combining studies and findings methods, conclusions supported, overall scientific quality	10-Item, 7-point scale Inter-rater agreement
Strayhorn et al,[31] 1993	Inter-rater reliability before and after training	Acceptance rates
Van Rooyen et al,[32] 1998; Van Rooyen et al[33] 1999	Importance of research question, originality, method, presentation, constructiveness of comments, substantiation of comments, interpretation of results	7-Item, 5-point scale with summary score time taken to review
Walsh et al,[34] 2000	As above plus tone of review	Tone of review

(continued)

Table A. Quality Criteria and Rating Methods Used in Published Studies of Editorial Peer Review (cont)

Study, y	Quality Criteria	Rating Method
	Studies With Other Outcomes	
Neuhauser and Koran,[35] 1989	Time and cost of editorial process	Time to review
	Quality Assessment Measures Used in Studies of Editorial Peer Review*	
	Study Topic	
Blank,[18] 1991	Effect of blinding on reviews	
Ernst and Resch,[22] 1994	Bias and interrater reliability of referees' judgments	
Feurer et al,[23] 1994	Validity of quality instrument for editors	
Jadad et al,[11] 1998	Quality of Cochrane and non-Cochrane reviews	
Justice et al,[27] 1998	Reader and reviewer agreement on manuscript quality	
Nylenna et al,[29] 1994	Effect of reviewing papers in native vs non-native language	
Oxman et al,[30] 1991	Consistency of quality index of systematic reviews	
Rochon et al,[16] 1994	Quality of randomized control trials published in supplements compared with those in the parent journal	

*Description of studies not mentioned in companion papers.