4: The effectiveness of journal peer review

ROBERT H FLETCHER, SUZANNE W FLETCHER

Peer reviewed journals throughout the world have adopted similar review practices in an effort to select the best among submitted manuscripts and to improve their quality before publication. These practices have been justified mainly by tradition and argument. Considering the high stakes of publication, peer review practices should be supported by scientific evidence that they improve outcomes (such as better published manuscripts) and by sound ethical reasoning that they are justifiable. A small but growing number of scientific studies of peer review can help guide the choice of peer review practices, but their generalisability is limited by the great variety of journals, reviewers, and editors, and the evidence for the overall effectiveness of current peer review practices is not clear cut. Nevertheless, there is an empirical basis for such practices as selecting and instructing reviewers, masking them to the author's identity, and asking them to sign reviews. Participation in peer review also has important beneficial effects on the medical community as a whole, whether or not the manuscript is published, by providing a vehicle for communication among scholars and by reinforcing ethical standards in the conduct of research. The cost of peer review is only a small proportion of the total budget of the journal. In the absence of conclusive evidence that usual peer review practices are best, variation in peer review practices is defensible, but should be accompanied by vigorous debate about their value and strong research into their effectiveness.

When editors of biomedical journals initiate peer review of a manuscript, they set in motion a chain of events that has far reaching consequences. Authors' reputations and livelihoods depend on whether their work is published. Reviewers invest precious time in the belief that they are making important contributions to the scientific process. Readers believe that peer review helps them manage information by affirming the scientific validity of published articles. The knowledge base for the biology of disease and the care of patients depends on the accuracy of published research.

Peer review has become standard practice for biomedical journals in the belief that it is the best way to accomplish such worthwhile goals. However, critics of peer review say its effects are not worth the costs. They deride the myth of "passing peer review",[1] and suggest that electronic publication without prior review would have several

Box 4.1 Conventional peer review

- One to three reviewers per manuscript
- Reviewers selected from a database, organised by expertise
- Reviewers are not required to sign their reviews
- A time deadline for returning reviews
- Authors do not know reviewers' names and institutions
- Simple, general instructions to reviewers
- Ask reviewers to provide:

  Comments for authors
  Comments for editors
  Grades on individual components and overall quality of the manuscript
- Reviewers are not paid for their work

advantages. It would be possible to reduce the lag time from submission to publication, to provide a more complete report of the work, and perhaps to facilitate a more effective and well documented form of self correction, through public exchanges of views and ongoing corrections, than is possible by one time peer review.[2]

Therefore, it is incumbent on editors, who establish peer review practices, to get it right. Are they doing so? And how would we know?

Usual peer review practices

Most peer reviewed biomedical journals have similar practices (Box 4.1). These have tended to become more uniform with time because of widely disseminated policy statements published by influential journals, and committees of these journals,[3] about the best way to carry out this work, coupled with growing efforts to share views about editorial practices – in four international congresses on peer review in biomedical publications and in professional societies (Council of Biology Editors, European Association of Scientific Editors, and World Association of Medical Editors).[4–7]

In this chapter, we describe the rationale and evidence base for peer review practices. We examine the effects of these practices on the outcomes of review and publication, their costs, and their effects on the medical profession.

The rationale for peer review practices

Peer review has been promoted, and defended against its critics, mainly by appeal to tradition and by arguments for why it *ought* to get

good results. A rich set of writings taking this approach is available to guide peer review practices.[8–11] More recently, some members of the scientific community have subjected peer review practices to more rigorous examination. The rationale for peer review can be established in two general ways. One can argue that peer review practices should be promoted or discouraged according to the scientific evidence of their effects on the main outcomes of review, selecting the best manuscripts and improving them. Alternatively, one can make the ethical arguments for peer review practices, asserting that editors are in a position of public trust and should choose processes that represent the values of the society in which they work.

Scientific evidence

It is all very well to argue that one or another peer review practice ought to get better results. But does it really? Specifically, everything else being equal, does peer review achieve better results compared with no peer review? And does peer review done one way get better results than peer review done in another?

The results of an increasing number of scientific studies of peer review are available. Most are descriptions of what reviewers currently do. Such studies are valuable to colleagues who wish to follow usual practices, but they do not provide evidence for whether or not these practices achieve their intended results. For this, editors need rigorous studies of the actual effects of editorial practices. A small but growing body of research on peer review and editing is beginning to clarify what the effects, or absence of effects, of usual peer review practices are.

The most valuable studies for this purpose are those that provide strong tests of hypotheses about the effects of peer review practices on the intended end results of review such as better selection of manuscripts and better published manuscripts.[12] Randomised controlled trials are the standard of excellence for studies of interventions, and there are a small but growing number of such trials of peer review practices. Less directly useful are non-randomised comparisons of peer review practices and studies of effects on intermediate steps in the review process, such as of what reviewers do and find (Figure 4.1).

Ethical rationale

A peer review practice may also be chosen because it is "the right thing to do", whatever the results on review and manuscript quality.[13] Here, the arguments are about values, not practical consequences. Practices can be justified on ethical grounds whether or not they

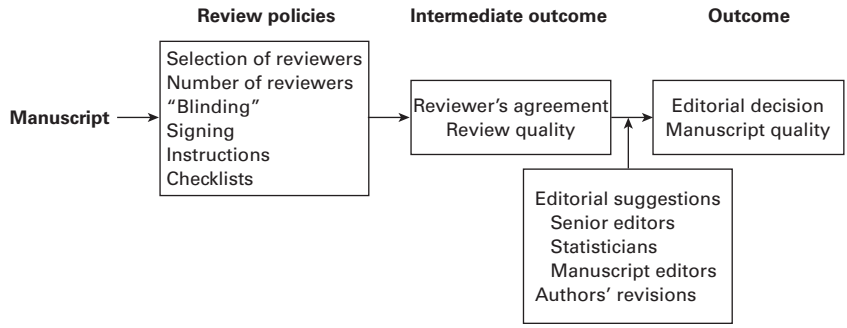

Figure 4.1 The process and outcome of peer review

achieve specific, practical outcomes such as better manuscripts. They can be framed in the language of professional ethicists (Table 4.1), though often they are not. The ethical bases for peer review practices stand or fall to the extent that they are sustained in the course of careful examination, vigorous exchange of views, sound argument, and connections to the values of the society in which they occur.

Effectiveness of specific peer review practices

In the following section we examine both the ethical rationale for several common peer review practices and what is known from research about their effects.

Selecting good reviewers

It is conventional wisdom that the best reviewers are senior, accomplished scholars because they have the experience and wisdom to give good advice. However, studies of reviewer performance suggest that this is not necessarily the case. Two North American studies[14,15] found that the best reviews (measured by both quality and promptness) were on average provided by relatively junior academicians. In one,[15] reviewer characteristics independently associated with review quality were age < 40, coming from a top institution, and being well known to the editor. In a European study,[16] younger reviewers, and those with training in epidemiology and statistics produced better reviews. A study at *Annals of Emergency Medicine* showed that editors' ratings of peer reviewers were only modestly correlated with the ability of a blinded reviewer to detect flaws deliberately introduced into manuscripts.[17]

Table 4.1 The ethical bases for peer review practices (examples)

Ethical principle	Example
Fairness	External reviewers are included in the review process because important scientific decisions should be made by one's peers and not just the editorial staff
Confidentiality	Reviewers' identities are not revealed because they have a right to anonymity while they are doing sensitive work and reviewers are asked not to share what they have learned about the work so that it remains the property of the authors
Conflict of interest	Reviewers and editors are asked to withdraw from the process if they have a financial interest in the question at hand
Full disclosure	Reviewers may decide on their own whether to sign their reviews because they are the ones who must bear the consequences (self determination). All editorial policies, including the elements of peer review, should be fully described to participants (authors, reviewers, and readers) so they know what they are getting into

The results of these studies suggest that editors should not have fixed views of what kinds of reviewers might return good reviews. Because the characteristics of good reviewers might vary from one setting to another, it seems editors should continue the common practice of grading their own reviewers but recognise that this is an imperfect predictor of their future performance.

Number of reviewers

Each journal has its own policy on the usual number of reviewers per manuscript. Most choose two, on the grounds that two is a reasonable trade-off between the need for external opinion on the one hand and the wish to use a scarce resource, reviewers' time, parsimoniously on the other.

The opinions of two reviewers, even if chosen at random from all possible reviewers, are too few in themselves to yield a statistically stable basis for deciding whether or not the manuscript should be published. Indeed, one would need to have at least six reviewers, all favouring publication or rejection, for their votes to yield a statistically significant conclusion ($P < 0.05$). This is one reason why it is not a good idea to "count votes" when deciding whether to publish

a manuscript. Reviewers' advice about whether to accept a manuscript is also limited because they cannot know about all of the factors that go into the decision to accept or reject a manuscript, such as other manuscripts that may have been submitted on the same topic and the editor's plans for the overall balance of articles in the journal. Reviewers are valuable consultants, providing second opinions and a rich array of insights. But they are not "referees" – that is, they should not, on their own, decide how the game is played.

Therefore, editors who choose only one reviewer, and those who choose several, have simply made different trade-offs between the value of additional information and their wish to spare reviewers, and themselves, work. It is not known how much additional, useful information is actually gained, on the margin, by choosing additional reviewers.

Instructing reviewers

Most journals provide simple instructions for reviewers. For example: "You should note on this sheet [Suggestions for Transmittal to Authors] your questions and comments about the manuscript and how it might be improved" (*New England Journal of Medicine*), or "A brief appraisal for the editor … should give a frank account of the strengths and weaknesses of the article" and "a detailed appraisal for the author … should be divided into major and minor comments, with annotations given according to page and paragraph" (*The Lancet*). Given the brevity of these instructions, it is not surprising that most first time reviewers are unclear about how this work is done. Experienced reviewers have learned the game by doing it, but there is no reassurance that they have learned how to play it well. A major method for improving peer review, so far little explored, may be for editors to be more forthcoming in what they want from reviewers.

At the opposite extreme are checklists for reviewers, designed to detect specific shortcomings in research design,[18] statistical analyses or written presentation.[19] These checklists remind reviewers to give attention to all the particulars of the manuscript. In fact, such lists do reveal many technical lapses that could be corrected before publication.[20] Structured abstracts[21] are a variation on this approach; by requiring authors to include a description of all essential components of their work, such as setting, patients, and design, valuable information is less likely to be omitted.

Despite their advantages, checklists are rarely used. Some editors believe that good reviewing, like good science, is not a mechanical process, but the careful use of knowledge, experience, and reasoning. There is also value in insight and imagination, which can be smothered by excessive structure. Perhaps a good compromise would be for peer review to be undertaken first with only broad guidelines, then be backed up by a checklist. In any case, there is little evidence on

whether the end result of peer review, such as better reviews or better editorial decisions about acceptance and revisions, is achieved by better instructions to reviewers or by the use of checklists.

Studies of efforts to teach reviewers have shown little effect.[22,23] In any case, although this approach might be helpful if promulgated in research training programmes, it is not feasible for the large number of geographically dispersed reviewers involved in most journal reviews.

Many editors send to reviewers a copy of the other reviewer's comments and their letter to authors. Small randomised controlled trials of this practice, both in the same journal, showed no effect on subsequent review quality.[24] This study alone is not sufficient to rule out an effect on review quality in other journals. In any case, sharing reviews may be worth doing for other reasons such as respect for reviewers' efforts and recognition of their interest in what others thought of the manuscript.

"Blinding" reviewers

Withholding from reviewers information on the authors' identity and their institutional affiliation ("blinding" or "masking") is believed to improve the quality of reviews by removing bias. For example, reviewers might give famous authors from prestigious institutions the benefit of the doubt and be more critical of the work of obscure authors from less respected places. On the other hand, blinding removes information that might be used to good purpose by reviewers. For example, one might want to be less forgiving of an experienced author who turns in a sloppy manuscript than of a junior author, or one working in a second language.

Some biomedical journals (especially in public health and the social sciences) do have a policy of blinding, and others (mainly those in the clinical and laboratory sciences) do not. Blinding is accomplished either by asking authors to submit manuscripts without identifiers or by having the editorial staff block out the information after the manuscript is received, a process that takes several minutes. It is not always successful. In one multi-journal study, 42% of reviewers were able to identify authors or institutions even after efforts to blind.[25] Reviewers may know about the work, or figure out its origins from references in the manuscript.

Several randomised controlled trials (RCTs) have established that blinding has at most a small effect on the quality of reviews. An early RCT showed a small, statistically significant effect (0·4 on a 5-point scale) on the quality of reviews.[26] Subsequent randomised trials did not find statistically significant effects on detection of weaknesses introduced into the manuscript,[27] or on review quality.[25,28,29] Other controlled studies have shown that blinded reviewers for an economics journal were more critical and less likely to recommend

acceptance[30] and that blinded reviewers for a paediatrics journal gave better scores to authors who had published more articles previously.[31]

As a consequence, journal editors might reasonably choose to blind or not. There appears to be little at stake in their choice.

Signing reviews

Some journals encourage reviewers to sign their reviews and many others do not. Proponents of signing reason that if reviewers take personal responsibility for their work they will do a better job. Opponents argue that reviewers who are compelled to sign might hold back useful information for fear of reprisal when their own work is up for judgement, perhaps by the very author that they had judged harshly. Research suggests that signing is not associated with large differences in review quality. In one study, those who chose to sign were more often judged constructive and courteous by editors and fairer by authors.[26] In two others, randomising reviewers to sign or not sign their reviews had no effect on the quality of the review.[27,28]

At present there is not enough evidence to require reviewers to sign or to ask them not to sign. Rather, the decision should be up to the individual reviewer. Whether the editor encourages or discourages signing depends on the kinds of personal interaction the editor wants contributors to his or her journal to have with each other.

Detecting scientific misconduct

It is generally agreed that peer reviewers cannot be relied on to detect misconduct in science,[32] defined as "fabrication, falsification or plagiarism, in proposing, performing or reporting research".[33] Reviewers are simply too far removed from the data and how they were collected to recognise inaccuracies in the original observations. Rarely, reviewers might notice inconsistencies in the results that suggest problems with the data. But, as a general rule, peer review is no protection against careless or fraudulent data collection.

Similarly, peer review is an unreliable way of detecting duplicate publication. Sometimes reviewers have by chance encountered a similar manuscript by the same authors. But judging from the frequency of duplicate publication, and how infrequently it is detected during the review process, it appears that traditional review practices pick up no more than the minority of potential duplicate publications before they occur.

Agreement among reviewers

The extent to which reviewers agree in their evaluation of a manuscript is an example of an intermediate step in the review

process, not an end in itself (Figure 4.1). Studies have shown that different reviewers of the same manuscript generally do not agree with each other.[34–36] Accurate measurement (validity) in science depends on reliability (reproducibility), which means multiple measurements agreeing with each other. But is disagreement among peer reviewers really bad? Some editors believe that reviewers should be chosen because they bring different kinds of expertise to bear on a manuscript. For example, one might be an expert in the content area, such as congestive heart failure, and another on the research methods, such as randomised trials. The reviewers can then complement each other and more information is available than there would be if they held similar views about the issues dealt with in the manuscript. Individual reviewers also tend to be consistently more positive or negative ("assassins or zealots").[37] Under these circumstances, one would expect reviewers to disagree. If reviewers are advisers to editors, then that advice is richer if their reviews reflect different expertise and values, and as a result disagree on the overall strength of the manuscript. Only if the reviewers' votes directly decided whether a manuscript should be accepted (which they should not) would lack of agreement among reviewers be a liability.

Overall effects of peer review on manuscript quality

Peer review is not the only reason why manuscripts change between submission and publication. Input from the editorial office – from senior editors, statisticians, and manuscript editors – can all affect the end result, as might the authors themselves, when they are given occasion to reconsider their manuscript after constructive suggestions for improvement and some time to reflect on how they have described their work. Because all of these inputs occur for all published manuscripts, it is difficult to separate out the effects of one (such as peer review) from the others.

How does the process as a whole affect manuscripts? Goodman and colleagues described changes, from submission to publication, in manuscripts published in *Annals of Internal Medicine*.[19] Of 34 items reflecting the quality of the research report (not the research itself), 33 changed in the direction of improvement, with the largest improvements for items related to aspects of the manuscript considered especially important to the editors: discussion of study limitations, generalisations, use of confidence intervals, and tone of the conclusions. Improvement was greatest in the manuscripts that had, on submission, the greatest room for improvement.

There is also evidence that readers recognise and value the changes made during peer review and editing.[38,39] A study of the *Nederlands*

Tijdschrift voor Geneeskunde found that readers recognised improvement in manuscripts after peer review (a comparison of submitted and accepted manuscripts) and editing (a comparison of accepted and published versions).[38]

More recently a series of systematic reviews have found that the evidence of an effect of peer review on manuscript quality (which was defined by the authors in a separate study) is thin and scattered.[40,41] Evidence of the effectiveness of technical editing however is firmer.[42]

The evidence suggests that peer review (perhaps) and editing (certainly) lead to better reports of research results. Whether the magnitude of improvement is worth the effort is a separate question. It is clear that even after the peer review process, published articles still have plenty of room for improvement.[19]

Effects of peer review on the profession

Requests for peer review set in motion a cascade of events in which authors, reviewers, and editors (including statisticians and manuscript editors) communicate with each other about written descriptions of scientific work. The multilateral conversation includes the full range of issues that matter in science: is the question important, are the methods sound, is the description clear, are the conclusions based on the results, etc. Participants both applaud and challenge each other's efforts. All of this takes place whether or not the manuscript is accepted for publication.

The magnitude of this communication network is enormous. Several scholars participate in each manuscript's review and there are many manuscripts per journal and many journals. In aggregate, journal peer review is the occasion for a massive programme of communication among scholars on important issues.

We believe that all participants benefit from the review process. Authors receive advice from other scholars in their field. When the manuscript is sent to another journal, and is reviewed by a second set of reviewers before it is published, the advice has a wider, and possibly sounder, base. For young reviewers who aspire to be successful researchers, reviewing is part of their initiation into the profession. They learn how successful manuscripts are crafted and how the "give and take" between authors and editors is carried out. More experienced reviewers may value the opportunity to see new work before it is published. All reviewers can improve their critical appraisal skills by putting themselves in a position where they must examine a research report in depth and by receiving the comments of other reviewers and editors, who have also examined the same manuscript carefully.

Table 4.2 The cost of peer review

	% of total budget (range)
Reviewer database	< 0·1–1·9
Staff time	1·3–1·9
Correspondence	< 0·1–2·6
Payment to reviewers	0·0–0·8
TOTAL	2·6–7·5

Based on estimates in 1998 from: *Annals of the Academy of Medicine, Singapore; British Medical Journal; Canadian Medical Association Journal; Journal of the American Medical Association; Journal of the Norwegian Medical Association; The Lancet*

The costs of peer review

Peer review is not without cost, both in financial and human terms. These costs need to be weighed against effectiveness when deciding whether peer review is worth while.

The financial costs of peer review have not been well described. To assess the size of the issue, we asked the editors of several peer reviewed journals to estimate the proportion of their journal's budget that could be attributed to peer review itself. That is, if they had not included peer review but otherwise carried out their work in the same way, how much smaller would their total budget have been? Items of peer review costs borne by journals are: creation and maintenance of a reviewer database; staff time for identifying reviewers and tracking reviews and manuscripts; correspondence with reviewers and with authors about reviews; editors' time in dealing with external reviews; and payment (if any) to reviewers.

Peer review appears to account for about 2·6–7·5% of total journal costs (Table 4.2), This percentage was generally higher for the smaller journals. Clearly, other aspects of publishing such as staff, printing, and distribution, which are present if there is to be a journal at all, take the lion's share of the total budget. This estimate is from the journal's perspective and does not take into account reviewers' work when it is uncompensated. In one study, the mean time per review was 3 hours (range 1/2–16 hours).[26]

Peer review has human costs too. Sometimes reviewers are discourteous or make unfounded suggestions, causing authors anger and frustration. This cost of reviewing can be minimised by the editors simply not sending hurtful or incompetent reviews to authors. Peer review also delays publication of research findings that might improve clinical and public health practices. There is great variation in how long journals take to carry out peer review, ranging from a few

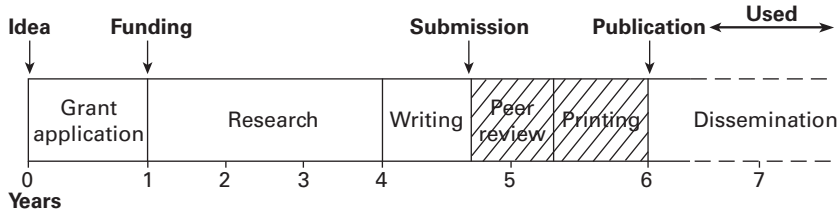

Figure 4.2 The time taken for peer review of randomised trials in relation to the total time from research idea to use of the results (data from *JAMA* 1998;**279**:284)

weeks for unusually important work, to 10 weeks in a well run journal, to many months in journals that have not made a special effort to see that all of the many steps in the peer review process take no longer than they need to.

The time for peer review should be viewed in context (see Figure 4.2). It is small in relation to the total time from research idea through funding and conduct of the research, to writing results up and submitting them for publication. It is not even a large part of the time from submission to publication. An even longer delay can occur from the time of publication to the time the new information is incorporated into practice. For example, at a time when there was conclusive evidence that β-blocking drugs reduce death rates following myocardial infarction, only 21% of eligible patients in the United States received these drugs in 1987–91.[43] Perhaps a clearer message in the original manuscript, as a result of peer review and editing, might actually shorten the duration from research idea to use in practice, the most meaningful interval.[44]

Should peer review be standardised?

Peer review practices may have served us well. We would argue that the weight of evidence suggests that current peer review practices do more good than harm, although others, using the more demanding standards of the Cochrane Collaboration, have described peer review as "largely untested and its effects are uncertain".[41] However, neither ethical arguments nor scientific evidence is currently so decisive as to suggest that peer review practices should be standardised across journals. Rather, there is every reason for journal editors to be open minded in their peer review policies, to develop imaginative ways to improve them, to continue to debate whether they are fair and to measure whether they actually achieve their desired effects.

References

1 Altman LK. The Ingelfinger rule, embargoes, and journal peer review – Part 1. *Lancet* 1996;**347**:1382–6.
2 Judson HF. Structural transformation of the sciences and the end of peer review. *JAMA* 1994;**272**:92–4.
3 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *JAMA* 1993;**269**:282–6.
4 *Guarding the guardians: research on editorial peer review*. Selected proceedings from the First International Congress on Peer Review in Biomedical Publication (1990) *JAMA* 1990;**263**:1309–456. (Also *Peer review in scientific publishing*. Papers from the First International Congress on Peer Review in Biomedical Publication. Chicago: Council of Biology Editors, 1991.)
5 The Second International Congress on Peer Review in Biomedical Publication. *JAMA* 1994;**272**:91–170.
6 The Third International Congress on Biomedical Peer Review in Global Communications. *JAMA* 1998;**280**:203–306.
7 Rennie D. Fourth International Congress on Peer Review in Biomedical Publication. *JAMA* 2002;**287**:2759.
8 Lock S. *A difficult balance. Editorial peer review in medicine*. Philadelphia: ISI Press, 1985.
9 Relman AS. Peer review in scientific journals – what good is it? *West J Med* 1990; **153**:520–2.
10 Kassirer JP, Campion EW. Peer review. Crude and understudied, but indispensable. *JAMA* 1994;**272**:96–7.
11 Chubin DE, Hackett EJ. *Peerless science. Peer review and U.S. science policy*. Albany, NY: State University of New York Press, 1990.
12 Fletcher RH, Fletcher SW. Evidence for the effectiveness of peer review. *Sci Eng Ethics* 1997;**3**:35–50.
13 Scheidermayer DL, Siegler M. Believing what you read. Responsibilities of medical authors and editors. *Arch Intern Med* 1986;**146**:2043–4.
14 Stossel TP. Reviewer status and review quality. Experience of the *Journal of Clinical Investigation*. *N Engl J Med* 1985;**312**:658–9.
15 Evans AT, McNutt RA, Fletcher SW, Fletcher RH. The characteristics of peer reviewers who produce good-quality reviews. *J Gen Intern Med* 1993;**8**:422–8.
16 Van Rooyen S, Godlee F, Evans S, Smith R, Black N. What makes a good reviewer and what makes a good review? *JAMA* 1998;**280**:231–3.
17 Callaham ML, Waeckerle J, Baxt W. Relationship of editorial ratings of reviewers to reviewer performance on a standardized test instrument. The International Congress on Biomedical Peer Review and Global Comunications. (abstract book) Chicago: JAMA, 1997:32.
18 Begg C, Cho M, Eastwood S, *et al*. Improving the quality of reporting of randomized controlled trials. The CONSORT Statement. *JAMA* 1996;**276**:637–9.
19 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.
20 Gore SM, Jones G, Thompson SG. *The Lancet's* statistical review process: areas for improvement by authors. *Lancet* 1992;**340**:100–2.
21 Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. *Ann Intern Med* 1990;**113**:69–76.
22 Callaham ML, Wears RL, Waeckerle JF. Effects of attendance at a training session on peer reviewer quality and performance. *Ann Emerg Med* 1998;**32**:318–22.
23 Strayhorn J, McDermott JF Jr, Tanquay 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**:947–52.
24 Callaham ML, Knopp RK, Galligher EJ. Effect of written feedback by editors on quality of reviews: two randomized trials. *JAMA* 2002;**287**:2781–3.
25 Van Rooyen S, Godlee F, Evans S, Smith R, Black N. The effect of blinding and unmasking on the quality of peer review: a randomized controlled trial. *JAMA* 1998;**280**:234–7.

26 McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer review. *JAMA* 1990;**263**:1371–6.
27 Godlee F, Gale CR, Martyn CN. The effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. *JAMA* 1998;**280**:237–40.
28 Van Rooyen S, Godlee F, Evans S, Black N, Smith R. Effect of open peer review on quality of reviewers and on reviewers' recommendations: a randomised trial. *BMJ* 1999;**318**:23–7.
29 Justice AC, Cho MK, Winkler MA, *et al*. Masking author identity in peer review: does it improve peer review quality? *JAMA* 1998;**280**:240–2.
30 Blank RM. The effects of double-blind versus single-blind reviewing: experimental evidence from the *American Economic Review*. *Am Econ Rev* 1991;**81**:1041–67.
31 Fisher M, Friedman SB, Strauss B. The effects of blinding on acceptance of research papers by peer review. *JAMA* 1994;**272**:143–6.
32 Lock S, Wells F, eds. *Fraud and misconduct in medical research*. London: BMJ Publishing Group, 1993.
33 National Academy of Science. *Responsible science*. *Ensuring the integrity of the research process*. Washington, DC: National Academy Press, 1992.
34 Cicchetti DV. The reliability of peer review for manuscript and grant submissions; a cross-disciplinary investigation. *Behav Brain Sci* 1991;**14**:119–86.
35 Ernst E, Saradeth T, Resch KL. Drawbacks of peer review. *Nature* 1993;**363**:296.
36 Rubin HR, Redelmeier DA, Wu AW, Steinberg EP. How reliable is peer review of scientific abstracts? Looking back at the 1991 Annual Meeting of the Society of General Internal Medicine. *J Gen Intern Med* 1993;**8**:255–8.
37 Siegelman SS. Assassins and zealots: variations in peer review. *Radiology* 1991;**178**:637–42.
38 Pierie JPEN, Walvoort HC, Overbeke AJPM. Readers' evaluation of effect of peer review and editing on quality of articles in the *Nederlands Tijdschrift voor Geneeskunde*. *Lancet* 1996;**348**:1480–3.
39 Sweitzer BJ, Cullen DJ. How well does a journal's peer review process function? A survey of authors' opinions. *JAMA* 1994;**272**:152–3.
40 Jefferson TO, Wager E, Davidoff F. Measuring the quality of editorial peer review. *JAMA* 2002;**287**:2786–90.
41 Jefferson TO, Alderson P, Davidoff F, Wager E. Effects of editorial peer review. *JAMA* 2002;**287**:2784–6.
42 Wager E, Middleton P. Effects of technical editing in biomedical journals: a systematic review. *JAMA* 2002;**287**:2821–4.
43 Soumerai SB, McLaughlin TJ, Spiegelman D, Hertzmark E, Thibault G, Goldman L. Adverse outcomes of underuse of beta-blockers in elderly survivors of acute myocardial infarction. *JAMA* 1997;**277**:115–21.
44 Fletcher SW. Research agenda for medical journals. In: Lock S, ed. *The future of medical journals*. London: BMJ Publishing Group, 1991.