Although there is a large literature on biomedical journal peer review, most of it is in the form of editorials, commentaries, reviews, and letters rather than original research. From the viewpoint of medical research, the evidence base for the effects of peer review is sparse and there are considerable gaps in our knowledge. This chapter aims to provide an overview of the literature on journal peer review and to direct readers to other parts of the book for detailed assessments of the research and distillation of the many opinions.

Background

There is no shortage of publications about journal peer review. A Medline search of this term (from the earliest entry to February 2003) reveals over 7800 articles, yet the yield drops to less than 50 if the search is limited to randomised controlled trials. This chapter is designed as an introduction to the peer review literature and a commentary on the state of the evidence rather than an account of the evidence itself which is analysed in detail in other parts of this book.

This chapter started out with a literature review performed by one of us (J.O.).\(^1\) The review covered the period from 1966 to July 1997 and the output of the 3rd International Congress on Biomedical Peer Review and Global Communications held in September 1997 in Prague. Since then, the output of the 4th International Congress on Peer Review in Biomedical Publication held in September 2001 in Barcelona has been added (by EW) but the exhaustive literature searches have not been updated.\(^1\)\textsuperscript{2}–\textsuperscript{5}\textsuperscript{1} However, since the first edition, three Cochrane reviews on peer review have been published, so readers interested in a recent systematic review can consult them.\(^2\textsuperscript{5}\textsuperscript{2}–\textsuperscript{5}\textsuperscript{4}\)

Box 3.1 shows the sources that were used to perform the original literature review. The bibliographies of all original articles retrieved were also studied. A book by BW Speck (\textit{Publication peer review: an annotated bibliography}. Westport, CT: Greenwood Press, 1993) and a special issue of \textit{Science and Engineering Ethics} were used to find articles
Box 3.1 Sources used for the literature review of journal peer review

Primary sources

- Editorial peer review in biomedical publication. The first international congress. *JAMA* 1990;263(10):1317–441
- Second international congress on peer review in biomedical publication. *JAMA* 1994;272:79–174
- Peer review congress IV. *JAMA* 2002;287:2745–71
- *Nederlands Tijdschrift voor Geneeskunde* (Dutch Journal of Medicine; 1984 to July 1997) was used as a source for a different language area

Medline

- The index terms “peer review”, “decision”, “quality”, “referee”, “acceptance”, and “rejection” were used to perform a Medline search for original articles on these subjects published in the period 1966 to July 1997

Social SciSearch

- The search strategy used for Medline was applied to the “Social Science Citation Index” (Social SciSearch) for the period 1973 to July 1997

Embase

- The same search strategy was used with “Embase” for the period 1974 to July 1997

European Science Editing/CBE Views

- *European Science Editing*, the journal of the European Association of Science Editors, issues for 1988–97
- *CBE Views*, the periodical of the Council of Biology Editors, issues for 1988–97

Additional Sources

- A special issue of *Science and Engineering Ethics* was used as an additional source for non-medical disciplines (Advances in peer review research, *Science and Engineering Ethics* 1997;3:1–104)

(Continued)
The quality of the evidence

Despite the four international congresses on peer review, which have added considerably to our knowledge base, the number of robustly designed published studies of the effects of journal peer review remains small. Of the 229 original articles identified in our search, only 36 reported prospective studies. The Cochrane reviews, using stricter selection criteria, included only 21 studies on editorial peer review and 18 on technical editing (of which only 2 were randomised controlled trials).\(^{232,233}\)

One of the reasons for the very small numbers of prospective randomised studies of journal peer review may be that such methods are not appropriate for studying a complex human behaviour. Most research on peer review in biomedical journals has been carried out by journal editors with a background in medical research. They therefore tend to use methods that have been developed for clinical trials. These methods (both for original research and for research synthesis) are widely accepted as valid ways of producing robust evidence about medicines and medical techniques, but they may be less appropriate for the complex psychosocial interactions involved in peer review. From the viewpoint of biomedical research, there is sparse evidence about peer review after the “real” research has been separated from the descriptive studies, opinions, and commentaries. Yet, to a social scientist, this might represent a rich literature base.

Another difficulty of research in this area is that peer review appears to fulfil a number of different functions and there is no consensus on its primary aim. In contrast, medicines are usually tested in strictly defined indications, with well established endpoints. Thus, studies of antihypertensive agents will recruit hypertensive patients and

---

**Box 3.1 Continued**

**The International Congresses on Biomedical Peer Review and Global Communications**

- The results of the studies presented at this meeting in Prague, September 1997 and in Barcelona in September 2001 were added

**Systematic reviews**

- A set of systematic reviews on peer review was presented in Barcelona.\(^{190,240}\) These have now been published in the Cochrane library.\(^{252-254}\)
measure the effects of treatment on their blood pressure, or on outcomes such as cardiovascular mortality. When the function of treatment has been clearly defined, and similar outcomes are used in large numbers of studies, it is relatively easy to compare the effectiveness of different treatments even when direct comparisons are not available. However, since the function of peer review has not been clearly defined, it is very difficult to measure the effectiveness of different interventions.255

Given the lack of consensus about the primary function of peer review, it is perhaps not surprising that a review of the literature reveals little agreement about the outcomes that should be measured. As with clinical trials, outcomes may be measured directly or via surrogates. For example, the ultimate aim of treating hypertension is to reduce the incidence of strokes and heart attacks. Yet new antihypertensives are usually licensed on the basis of evidence about their effects on blood pressure (measured in millimetres of mercury) since there is good epidemiological evidence linking this surrogate outcome to the more clinically meaningful ones. Sadly, for peer review research, there is little evidence linking proxy outcomes (such as reviewer agreement) to broader definitions of effectiveness (such as the effect on the importance or comprehensibility of published reports). However, our literature review, and those of others, reveal that nearly all studies of peer review concentrate on surrogate or process-based outcomes.255

**Themes in peer review research**

Despite the patchiness of the evidence about journal peer review, certain themes emerge. To avoid repetition, we suggest that readers interested in particular aspects consult the reference sections of the relevant chapters (Table 3.1).

**Conclusions**

Much has been published about journal peer review but most of it is opinion or commentary. Although the international congresses in peer review have stimulated research, most published studies are descriptive and there are few prospective, randomised trials. From the viewpoint of clinical research methodology we conclude that there is little evidence about the effects of peer review. Since its primary objectives have not been well defined, there is no consensus about how the quality of peer review should be measured. Most published studies have therefore focused on surrogate or process-based outcomes. However, lack of evidence should not be taken to imply lack of effect. Furthermore, it is possible that we should approach the
study of peer review from a different angle, and, instead of applying methods designed for clinical research and the synthesis of clinical data, we should seek a deeper understanding of this complex phenomenon from the behavioural and social sciences.

Although we do not know much about the peer review process, it is the only system we currently have to assess the quality of scientific work. Large amounts of resources are spent on peer review by journals and grant giving bodies. The costs in time and resources to the broader scientific community have not been properly measured but are undoubtedly high.\textsuperscript{256} It is therefore important to develop standardised and validated methods for further research. An important component of this will be the development of agreed outcome measures to assess the effects of peer review.

References

1 Pierie JPEN, Hoogeveen JH, Overbeke AJPM. Peer review, een gestructureerd overzicht van de literatuur; het effect van blinderen en het gemis aan prospectief onderzoek [Peer review, a structured review of the literature; the effect of blinding and the lack of prospective studies]. Listing available from John Overbeke.

Primary sources

3 Barnard H, Overbeke AJPM. Dubbelpublikatie van oorspronkelijke stukken in en uit het 
Nederlands Tijdschrift voor Geneeskunde [Duplicate publication of original articles 
4 Bero LA, Glantz SA, Rennie D. Publication bias and public health policy on 
5 Bonjean CM, Hullum J. Reasons for journal rejection: an analysis of 600 
6 Cho MK, Bero LA. Instruments for assessing the quality of drug studies published in 
7 Cicchetti DV. The reliability of peer review for manuscript and grant submissions: 
8 Colaianni LA. Peer review in journals indexed in Index Medicus. JAMA 
10 Dickersin K, Min YI, Meinert CL. Factors influencing publication of research 
11 Easterbrook PJ, Berlin J, Copalan R, Matthews DR. Publication bias in clinical 
12 Evans AT, McNutt RA, Fletcher SW, Fletcher RH. The characteristics of peer 
13 Evans JT, Nadjari HI, Burchell SA. Quotation and reference accuracy in surgical 
14 Feurer ID, Becker GJ, Picus D, Ramirez E, Darcy MD, Hicks ME. Evaluating peer 
15 Fisher M, Friedman SB, Strauss B. The effect of blinding on acceptance of research 
16 Gallagher EB, Ferrante J. Agreement among peer reviewers for a middle-sized 
biomedical journal. In: Peer review in scientific publishing. Chicago: Council of 
17 Gardner MJ, Bond J. An exploratory study of statistical assessment of papers 
18 Garfunkel JM, Lawson EE, Hamrick HJ, Ulshen MH. Effect of acceptance or 
rejection on the author's evaluation of peer review of medical manuscripts. JAMA 
19 Garfunkel JM, Ulshen MH, Hamrick HJ, Lawson EE. Effect of institutional prestige 
20 Garfunkel JM, Ulshen MH, Hamrick HJ, Lawson EE. Problems identified by 
21 Gilbert JR, Williams ES, Lundberg GD. Is there gender bias in JAMA's peer review 
22 Glantz SA, Bero LA. Inappropriate and appropriate selection of "peers" in grant 
23 Goodman SN, Berlin JA, 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.
25 Hobma SO, Overbeke AJPM. Fouten in literatuurverwijzingen in het Nederlands 
Tijdschrift voor Geneeskunde [Accuracy of references in the Dutch Journal of 
26 De Jong BCH, Overbeke AJPM. Peer review: is één of koning? [Peer review: is 
27 Justice AC, Berlin JA, Fletcher SW, Fletcher RH, Goodman SN. Do readers and peer 
28 Kan CC, Lockefer JHM, Overbeke AJPM. Redenen van afwijzing van artikelen voor 
publikatie bij het Nederlands Tijdschrift voor Geneeskunde in 1990 [Reasons of 
rejection at the Dutch Journal of Medicine in 1990]. Ned Tijdschr Geneeskd 
29 Laband DN, Piette MJ. A citation analysis of the impact of blinded peer review. 

Medline


60 Lloyd ME. Gender factors in reviewer recommendations for manuscript publication. *J Appl Behav Anal* 1990;23:539–43.


**Social SciSearch**


87 Wallmark JT, Sedig KG. Quality of research measured by citation method and by peer review – a comparison. *IEEE Tran Eng Manage* 1986;33:218–22.


---

**Embase**


---

**European Science Editing**


---

**Additional sources**


106 Fiske DW, Fogg L. But the reviewers are making different criticism of my paper! Diversity and uniqueness in reviewer comments. *Am Psychol* 1990;45:591–8.

**The International Congress on Biomedical Peer Review and Global Communications, Prague, 1997**


**Fourth International Congress on Peer Review in Biomedical Publication, Barcelona 2001**


219 Pitkin RM, Burmeister LF. Identifying manuscript reviewers. Randomized comparison of asking first or just sending. JAMA 2002;287:2795–6.


**Cochrane reviews**


**Other references**
