



## In Search Of an Optimal Peer Review System

[Richard W. Smith](#)

[Commentary](#) | [Opinion](#) | [Vol. 1, 2009](#) | October 21, 2009

### Abstract

**Summary:** After 30 years of practicing peer review and 15 years of studying it experimentally, I'm unconvinced of its value. Its downside is much more obvious to me than its upside, and the evidence we have on peer review tends to support that jaundiced view. Yet peer review remains sacred, worshiped by scientists and central to the processes of science — awarding grants, publishing, and dishing out prizes. It would be a bold funding body or journal that abandoned peer review, but could we at least do better? I want here to explore peer review — from a rather personal point of view — and ask questions about what would be the best system for the *Journal of Participatory Medicine*.

**Competing Interests:** The author has declared that no competing interests exist.

After 30 years of practicing peer review and 15 years of studying it experimentally, I'm unconvinced of its value. Its downside is much more obvious to me than its upside, and the evidence we have on peer review tends to support that jaundiced view. Yet peer review remains sacred, worshiped by scientists, and central to the processes of science—awarding grants, publishing, and dishing out prizes. It would be a bold funding body or journal that abandoned peer review, but could we at least do better? I want here to explore peer review—from a rather personal point of view—and ask questions about what would be the best system for the *Journal of Participatory Medicine*.

### The Misery of Peer Review

Let me begin with my immediate frustrations. For 25 of the past 30 years I've been editor (of the [BMJ](#)), and now I'm a reviewer and an author. I have just completed a review for the BMJ—of a paper that had interesting new data on an important topic but doubtful methods. Despite my scepticism about peer review, I usually accept requests to review, although I always wonder why. It's time-consuming and unpaid and usually my comments disappear into the void. The BMJ, as I expected, rejected the article, primarily on the grounds that the paper wasn't right for its audience. Like most major journals, the BMJ rejects over 90% of the studies it receives, many of them after hours of scrutiny and comment by reviewers. The reviewers' time is largely wasted because many authors, recognizing the arbitrariness of the "publishing game," simply send the paper elsewhere without revision.

I think that it would make much more sense simply to publish the paper—on a university website or in an electronic journal with a low threshold—with my comments and those of the other reviewer, and let the world decide what it thinks. That is anyway what happens in that many peer-reviewed

papers disappear without trace after publication, some are torn to pieces, and a few flourish and are absorbed into the body of science. The paper rejected by the BMJ, which may well not surface for another year, contained data that would fascinate some and inform a current and important debate. I can't see that any harm would result from it being available to all.

At the moment, I'm also waiting for an opinion on a paper that tells a complicated story of what we see as scientific misconduct on the part of a publisher. Four of us on three continents wrote the paper rapidly because it suddenly became topical after a major news story. We asked the BMJ to fast track the paper, and it was rapidly rejected with some thoughtful reviews. We did, unusually, revise the paper and submit it to another journal with a request for rapid review. That was about two months ago, and the only thing we've heard has been from a reviewer, who happens both to be a friend and to have written a review on our paper for the BMJ. He wanted to know if he could simply send the same review, but I told him that—perhaps unfortunately for him—we had revised the paper in the light of his opinion. So he'll have to review it again. Our chances of getting published in the second journal are perhaps 30%. If the paper is rejected, we'll either get fed up and abandon it or continue our way down the food chain—because you can get virtually anything published if you persist long enough.

Again, I think that much would be gained and really nothing lost if our paper was simply posted on a website with the reviewers' comments attached.

## Peer Review Is a Deeply Flawed System

[The Sixth International Congress of Peer Review in Biomedical Publication was held in September 2009](#), and dozens of scientific studies were presented on the subject. The First Congress was in Chicago in 1989, when many of the presentations were opinion rather than new data—but that was about the beginning of studies of peer review. Until then, it was unstudied despite being at the core of how science is conducted. Sadly, in my experience, most scientific editors know little about the now [large body of evidence](#) on peer review. So paradoxically, the process at the core of science is based on faith rather than experimental evidence.

If editors were to examine this body of literature, they would discover that evidence on the upside of peer review is sparse, while evidence on the downside is abundant. We struggle to find convincing evidence of its benefit, but we know that it is slow, expensive, largely a lottery, poor at detecting error, ineffective at diagnosing fraud, biased, and prone to abuse.[\[1\]\[2\]\[3\]](#) Sadly we also know—from hundreds of systematic reviews of different subjects and from studies of the methodological and statistical standards of published papers—that most of what appears in peer-reviewed journals is scientifically weak.[\[4\]\[5\]\[6\]\[7\]](#)

The evidence on peer review has been gathered together in a book specifically on peer review,[\[1\]](#) and I have summarized the evidence on its many problems in a book and an article.[\[2\]\[3\]](#) Let me quote here just three studies.

Two of them are Cochrane reviews of the evidence to support peer review in both scientific publishing and grant giving.[\[8\]\[9\]](#) Cochrane reviews, as most readers probably know, are widely regarded as the highest-quality systematic reviews. The paper on peer review in journals concludes: "At present, little empirical evidence is available to support the use of editorial peer review as a mechanism to ensure quality of biomedical research.[\[8\]](#)" And the review on grant giving: "There is little empirical evidence on the effects of grant-giving peer review. No studies assessing the impact of peer review on the quality of funded research are presently available.[\[9\]](#)" Both reviews point out that "absence of evidence" is not the same as "evidence of ineffectiveness," and the first paper says that studying peer review is methodologically difficult.[\[8\]](#) So there may be benefit in peer review, as most scientists believe, but we haven't yet been able to show it convincingly in empirical studies.

One way that we studied peer review at the BMJ was by inserting deliberate errors into short papers

and then asking reviewers to review the papers without telling them that they contained the inserted errors.<sup>[10]</sup> These studies consistently showed that reviewers spotted only a minority of errors and that many reviewers spotted none. The table shows how few of 607 reviewers spotted nine major errors and five minor errors that had been inserted into papers describing randomized trials, which are arguably one of the easiest types of study to review because expected standards are so explicit.

**Table: Proportion of reviewers identifying each error by group for the three papers.<sup>[10]</sup>**

	Paper 1	Paper 2	Paper 3
	Control	Self-taught	Face-to-face
<b>MAJOR ERRORS</b>			
Poor justification for study	31	36	36
Biased randomization procedure	49	58	53
No sample size calculation	21	24	21
Unknown reliability and validity of outcome measure	13	19	21
Failure to analyze the data on an intention-to-treat basis	22	18	22
Poor response rate	34	36	37
Unjustified conclusions	43	40	41
Discrepancy between abstract & results	23	25	28
Inconsistent denominator	38	45	53
<b>MINOR ERRORS</b>			
No ethics approval	18	14	14
No explanations for ineligible or non-randomized cases	50	48	58
Inconsistency between text & tables	5	2	2
Word reversal	13	9	10
No mention of Hawthorne effect	21	12	19

## Improving Peer Review

As people have understood better the many defects of peer review, they have tried ways to improve it. One of the first developments was to try blinding reviewers to the identity of authors. A randomized trial conducted by Bob and Suzanne Fletcher with others before they became editors of the *Annals of Internal Medicine*, using the quality of the opinion as the outcome measure, showed that blinding did improve quality.<sup>[11]</sup> But then two much bigger trials found no evidence of improvement, and about 10% to 20% of reviewers could anyway identify the authors.<sup>[12][13]</sup>

We conducted one of those trials at the *BMJ* and then decided that we would try the opposite—allowing the authors to know the identity of the reviewers. In a large trial of open peer review, we found no difference in the quality of the reviewers' opinions.<sup>[14]</sup> At this point we introduced open peer review—with the authors but not the readers knowing the identity of the reviewers—on ethical grounds, arguing that reviewers should be accountable for their judgments and receive credit. A judgment by an unknown judge seemed totalitarian.

Interestingly, when we asked a sample of reviewers whether they would review openly about half said yes and half no. When we conducted the trial, very few people declined to review openly and when we introduced the policy only a handful of reviewers in a database of around 5,000 refused to sign reviews. As a professor of economics said to me recently: "Economists have no interest in people's views but only in what they do."

Our next step was to conduct a trial of telling reviewers that if the papers they reviewed were published then all the background peer review material—including reviewers' and editors' comments and authors' responses—would be published. We completed the trial approximately five years ago, but the results have not yet been published, partly, ironically, because of delays in peer review. The trial did not find any meaningful difference in the quality of opinions. We've also tried training reviewers—but again, this had little impact on the quality of their reviews, perhaps because we were trying to teach "old dogs new tricks" or because the dose of training was inadequate.<sup>[15]</sup>

The plan at that stage at the *BMJ* was to proceed to open up the whole process—placing submitted papers online, asking reviewers to comment, and allow anybody, but particularly authors, to comment as the process proceeded. Peer review would thus be transformed from a black box to an open scientific discourse.

This development hasn't happened, and it seems to many an impossibly radical step. The main objection is that "low quality, possibly dangerous" material will be released. My response is that this happens already. We know that lots of poor quality research appears because of the hopelessness of peer review, because papers are regularly presented at major conferences with virtually no peer review, and because many researchers present the results of their studies to the press. In the latter two cases there is usually no detailed description of methods and results, meaning that it's impossible for even the well-informed to evaluate the study. With the system we proposed, the full data would be available.

## Acquiring Better Evidence on Peer Review

Most scientists continue to believe in peer review despite the lack of evidence to support it. This is partly because most are unaware of the evidence, but some think that we are simply not studying peer review in the right way. Could it be that more sophisticated or different methods could show the utility of peer review? This question was raised at the end of the Fourth Congress on Peer Review in Biomedical Publications, but there has been progress with methods. Most of those who have studied peer review have been from epidemiological and statistical backgrounds. Perhaps social scientists using qualitative methods could find evidence to support the belief of scientists that peer review is beneficial.

## "The Job of the Many Not the Few"

Web 2.0, the social web, may hold the key to the future of peer review. Peer review will become the job of the many rather than the few, and we know that the many can solve problems better than the few[16]—that must, indeed, be part of the philosophy of participatory medicine. We need to move, says Charles Leadbeater, one of the gurus of the Web, from "I think" to "we think" and find ways to harvest the thinking of thousands.[17] Instead of filtering and then publishing because publishing is expensive, we can now, says Clay Shirky, another thought leader on social media, "publish and then filter.[18]" It is this radical thinking that has created the magnificence of Wikipedia.

Peter Frishauf, founder of Medscape and another forward thinker, has suggested that we can use reputation systems—rather as eBay does—to filter material.[19] In a sense, this happens already: after publication a process ensues whereby most studies disappear but a few flourish and have consequences in the real world. It must in some way be the trusted, those with reputations, who drive this process.

Might we be able to automate the process, asking scientists and others to score studies—so allowing those that are the most important to rise to the top? This has been tried—for example, by the Public Library of Science (PLoS)—but so far scientists seem reluctant to score studies. PLoS has, however, recently added "article metrics" to all its studies. These metrics include article usage statistics (in graphic form), citations from scholarly literature, social bookmarks, comments left by readers, notes left within articles, blog posts, and ratings. When combined, these metrics will surely give a clear picture of which studies are of the most importance. Article metrics can be useful, however, only after publication.

**And** could we formalize and give validity to the process of attributing reputation? Currently, reputations are won dubiously and, as Mark Twain said, "Once you have a reputation for being an early riser, you can sleep in 'til noon every day." B. Thomas Adler and Luca de Alfaro have described a system whereby a reputation is built mathematically for Wikipedia contributors. It involves scoring them positively for making contributions that persist and negatively (lost points) for contributions that are rapidly removed.[20] Could we find an equivalent for peer review? (See [Frishauf sidebar](#) to this article.)

## Conclusion

We already have some provisional answers to these questions that have come from a discussion that is available as a podcast in this journal (see "[Peer Review and Reputation Systems: A Discussion](#)"). The overall conclusion was that the present system of peer review is badly broken and that something new is needed. Most of those in the discussion would favor moving to a system of "publish and let the world decide," preferably with systems of reputation and article metrics. We are at an early stage with these systems, and there was agreement that we should experiment—recognizing that experimentation inevitably means some "failures." "In all science, error precedes the truth, and it is better it should go first than last," said Hugh Walpole.

It does, however, feel very bold for editors to abandon prepublication of peer review—like walking into the street naked. But if the emperor has no clothes, what's to be lost? Nothing, but much is to be gained.

## References

1. Godlee F, Jefferson T. *Peer Review in Health Sciences. 2nd ed.* London: BMJ Books; 2003. [[Google Scholar](#)] ⇐
2. Smith R. Peer review: A flawed process at the heart of science and journals. *J R Soc Med* 2006;99:178-182. [[Google Scholar](#)] ⇐
3. Smith R. *The Trouble With Medical Journals.* London: RSM Press; 2006. [[Google Scholar](#)]

4. Altman DG. Poor-quality medical research: What can journals do? *JAMA*. 2002;287:2765–2767. [[Google Scholar](#)] ↔
5. Altman DG. Statistics in medical journals. *Stat Med*. 1982;1:59–71. [[Google Scholar](#)] ↔
6. Andersen B. *Methodological Errors in Medical Research*. Oxford: Blackwell; 1990. [[Google Scholar](#)] ↔
7. Altman DG. The scandal of poor medical research. *BMJ*. 1994;308:283–284. [[Google Scholar](#)] ↔
8. Jefferson T, Rudin M, Brodney Folsie S, Davidoff F. Editorial peer review for improving the quality of reports of biomedical studies. *Cochrane Database of Systematic Reviews* 2007, Issue 1. Art. No.: MR000016. DOI: 10.1002/14651858.MR000016.pub3 [[Google Scholar](#)] ↔
9. Demicheli V, Di Pietrantonj C. Peer review for improving the quality of grant applications. *Cochrane Database of Systematic Reviews* 2007, Issue 1. Art. No.: MR000003. DOI: 10.1002/14651858.MR000003.pub2 [[Google Scholar](#)] ↔
10. Schroter S, Black N, Evans S, Godlee F, Osorio L, Smith R. What errors do peer reviewers detect, and does training improve their ability to detect them? *J R Soc Med*. 2008;101: 507–514. [[Google Scholar](#)] ↔
11. 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. [[Google Scholar](#)] ↔
12. Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D. The PEER investigators. Does masking author identity improve peer review quality: A randomised controlled trial. *JAMA*. 1998;280:240–242. [[Google Scholar](#)] ↔
13. 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. [[Google Scholar](#)] ↔
14. 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. [[Google Scholar](#)] ↔
15. Schroter S, Black N, Evans S, et al. Effects of training on the quality of peer review: A randomised controlled trial. *BMJ* 2004;328:657–658. [[Google Scholar](#)] ↔
16. Surowiecki J. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few*. London: Abacus, 2005. [[Google Scholar](#)] ↔
17. Leadbeater C. *We-think, Mass Innovation, Not Mass Production*. London: Profile, 2008. [[Google Scholar](#)] ↔
18. Shirky C. [Here Comes Everybody: How Change Happens When People Come Together](#). London: Penguin, 2009. [[Google Scholar](#)] ↔
19. Frishauf P. The end of peer review and traditional publishing as we know it. Available at: <http://www.medscape.com/viewarticle/583316> Accessed October 3, 2009 (free registration required). [[Google Scholar](#)] ↔
20. Adler BT, de Alfaro, L. A content driven reputation system for the Wikipedia. In *WWW 2007, Proceedings of the 16th International World Wide Web Conference*, ACM Press, 2007. Available at: <http://users.soe.ucsc.edu/~luca/papers/07/wikiwww2007.html>. Accessed October 14, 2009. [[Google Scholar](#)] ↔