

SHORT COMMUNICATION

Classical peer review: an empty gun

Richard Smith*

'If peer review was a drug it would never be allowed onto the market,' says Drummond Rennie, deputy editor of the *Journal Of the American Medical Association* and intellectual father of the international congresses of peer review that have been held every four years since 1989. Peer review would not get onto the market because we have no convincing evidence of its benefits but a lot of evidence of its flaws.

Yet, to my continuing surprise, almost no scientists know anything about the evidence on peer review. It is a process that is central to science - deciding which grant proposals will be funded, which papers will be published, who will be promoted, and who will receive a Nobel prize. We might thus expect that scientists, people who are trained to believe nothing until presented with evidence, would want to know all the evidence available on this important process. Yet not only do scientists know little about the evidence on peer review but most continue to believe in peer review, thinking it essential for the progress of science. Ironically, a faith based rather than an evidence based process lies at the heart of science.

What is peer review?

Peer review is not easily defined, and every grant giving body and journal will have a process that is unique in some way. It is clearly something to do with an external, third party reviewing a grant proposal or manuscript. But how many external reviewers should there be? And under what conditions should they review? Should they be anonymous or identified to authors and readers? And who is a peer? Somebody who also researches on the subject of the proposal or manuscript or somebody who is simply in the same discipline? Should reviewers be trained? Different answers to these questions and many others lead to wide variation in systems of peer review.

One useful way of classifying peer review of completed studies is into 'pre-publication' and 'post-publication.' When people speak and write about peer review they usually mean pre-publication review, the process that

takes place before a study is published. But what happens after publication can also be called peer review, and that, I believe, is the peer review that really matters - the process whereby the world decides the importance and place of a piece of research. Arthur Balfour, a British prime minister, might have been speaking of science when he famously said that 'nothing matters much and few things matter at all.' Many studies are never cited once, most disappear within a few years, and very few have real, continuing importance.

And the correlation between what is judged important in pre-publication peer review and what has lasting value seems to be small. Fabio Casati, professor of computer science at the University of Trento, the holder of 20 patents, and the founder of a 'liquid journal' that had dispensed with prepublication peer review, says: 'We've...found that peer review doesn't work, in the sense that there seems to be very little correlation between the judgement of peer reviewers and the fate of a paper after publication. Many papers get very high marks from their peer reviewers but have little effect on the field. And on the other hand, many papers get average ratings but have a big impact' [1].

Indeed, the correlation could even be inverse in that peer review may well be biased against the truly original. I return to this point below.

But what is peer review for? (And from now on I shall mean pre-publication peer review when I write just 'peer review.' I will also be writing mostly about peer review of manuscripts for publication rather than of grants because that is what has been studied the most, it is what I know best, and it does have a clear alternative - simply publishing the manuscript and letting the world decide.) I see four main objectives for peer review: selecting what should be published, improving what is published, detecting errors, and detecting fraud.

Is peer review effective?

The Cochrane Collaboration, the organization that through its systematic reviews produces the most reliable evidence in medicine and health care, has reviewed the evidence on peer review of manuscripts and of grant proposals. This is its conclusion on peer review of manuscripts: 'At present, little empirical evidence is available to support the use of editorial peer review as a

*Correspondence: richardswsmith@yahoo.co.uk
35 Orlando Road, London SW4 0LD, UK

mechanism to ensure quality of biomedical research' [2]. And here is its conclusion on peer review of grant proposals: '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' [3].

Of course the absence of evidence and evidence of absence of effect are not the same thing, and many, particularly the many with a vested interest in peer review, continue to believe that peer review is beneficial but that it has not been studied in the right way. Many can also tell anecdotes of how a study they published was much improved by peer review. Many can also, however, tell anecdotes of bad experiences of peer review, and particularly of huge delays caused by peer review with no benefit. Everybody could perhaps agree that it is shameful that a process so central to science should have no evidence to support its effectiveness - even if in reality it is effective.

If peer review is to be thought of primarily as a quality assurance method, then sadly we have lots of evidence of its failures. The pretentiously named medical literature is shot through with poor studies. John Ioannidis has shown how much of what is published is false [4]. The editors of *ACP Journal Club* search the 100 'top' medical journals for original scientific articles that are both scientifically sound and important for clinicians and find that it is less than 1% of the studies in most journals [5]. Many studies have shown that the standard of statistics in medical journals is very poor [6].

Sadly we have many examples of studies published in medical journals that are not only scientifically poor but also have done great damage. The most famous example is the *Lancet* paper that suggested that the MMR (measles, mumps, rubella vaccine) caused autism: the result was a drop off in the number of children vaccinated, epidemics of measles, and more than a decade of fruitless argument [7]. Another example is the *New England Journal of Medicine* article that seemed to show that a new drug for arthritis, rofecoxib, was safer than the traditional non-steroidal anti-inflammatory drugs because it was less likely to cause gastrointestinal bleeding [8]. Unfortunately, the flawed paper hid the increase in myocardial infarctions. The paper was important in the new drug being widely used and in causing thousands of patients to have heart attacks.

Doug Altman, perhaps the leading expert on statistics in medical journals, sums it up thus: 'What should we think about researchers who use the wrong techniques (either wilfully or in ignorance), use the right techniques wrongly, misinterpret their results, report their results selectively, cite the literature selectively, and draw unjustified conclusions? We should be appalled. Yet numerous studies of the medical literature have shown

that all of the above phenomena are common. This is surely a scandal' [9].

While Drummond Rennie writes in what might be the greatest sentence ever published in a medical journal: 'There seems to be no study too fragmented, no hypothesis too trivial, no literature citation too biased or too egotistical, no design too warped, no methodology too bungled, no presentation of results too inaccurate, too obscure, and too contradictory, no analysis too self-serving, no argument too circular, no conclusions too trifling or too unjustified, and no grammar and syntax too offensive for a paper to end up in print.'

The downside of peer review

We have little or no evidence that peer review 'works,' but we have lots of evidence of its downside.

Firstly, it is very expensive in terms of money and academic time. At the *British Medical Journal* we calculated that the direct cost of reviewing an article was, on average, something like £100 and the cost of an article that was published was much higher. These costs did not include the cost of the time of the reviewing academics, who were not paid by the journal. The Research Information Network has calculated that the global cost of peer review is £1.9 billion [10]. The cost in time is also enormous, and many scientists argue that time spent peer reviewing would be better spent doing science.

The cost in time and money is much increased by studies working their way down the food chain of journals. A study may be submitted to *Nature* and rejected, then sent to the *New England Journal of Medicine* and rejected, and so on through the *Lancet*, *British Medical Journal*, and several specialist journals before ending up in a local journal. Often the same reviewers will be consulted repeatedly. And we know that if authors persist long enough, you can get anything published.

This expensive and time consuming process might be acceptable if it sorted the information effectively, with the most important studies being in the most important journals. Not only does this not happen (see below) but this ineffective sorting of information introduces an important bias - because the 'sexier' articles end up in the 'top' journals. The many people who read these journals because they think that they are reading what is most important are actually being presented with a distorted view of science.

Secondly, peer review is slow. The process regularly takes months and sometimes years. Publication may then take many more months. A friend of mine, a fellow of the Royal Society, has written a paper that I think very important for global health. As I write, it is still unpublished after two years of being reviewed by several 'top' journals. None of the reviewers have raised a major flaw with the study.

Thirdly, peer review is largely a lottery. Multiple studies have shown how if several authors are asked to review a paper, their agreement on whether it should be published is little higher than would be expected by chance [11]. A study in *Brain* evaluated reviews sent to two neuroscience journals and to two neuroscience meetings [12]. The journals each used two reviewers, but one of the meetings used 16 reviewers while the other used 14. With one of the journals the agreement among the journals was no better than chance while with the other it was slightly higher. For the meetings the variance in the decision to publish was 80 to 90% accounted for by the difference in opinions of the reviewers and only 10 to 20% by the content of the abstract submitted.

A fourth problem with peer reviews is that it does not detect errors. At the *British Medical Journal* we took a 600 word study that we were about to publish and inserted eight errors [13]. We then sent the paper to about 300 reviewers. The median number of errors spotted was two, and 20% of the reviewers did not spot any. We did further studies of deliberately inserting errors, some very major, and came up with similar results.

The fifth problem with pre-publication peer review is bias. There have been many studies of bias - with conflicting results - but the most famous was published in *Behavioural and Brain Sciences* [14]. The authors took 12 studies that came from prestigious institutions that had already been published in psychology journals. They retyped the papers, made minor changes to the titles, abstracts, and introductions but changed the authors' names and institutions. They invented institutions with names like the Tri-Valley Center for Human Potential. The papers were then resubmitted to the journals that had first published them. In only three cases did the journals realise that they had already published the paper, and eight of the remaining nine were rejected - not because of lack of originality but because of poor quality. The authors concluded that this was evidence of bias against authors from less prestigious institutions. Most authors from less prestigious institutions, particularly those in the developing world, believe that peer review is biased against them.

Perhaps one of the most important problems with peer review is bias against the truly original. Peer review might be described as a process where the 'establishment' decides what is important. Unsurprisingly, the establishment is poor at recognizing new ideas that overturn the old ideas. It is the same in the arts where Beethoven's late string quartets were declared to be nothing but noise and Van Gogh managed to sell only one painting in his lifetime. David Horrobin, a strong critic of peer review, has collected examples of peer review turning down hugely important work, including Hans Krebs's description of the citric acid cycle, which won him the Nobel

prize, Solomon Berson's discovery of radioimmunoassay, which led to a Nobel prize, and Bruce Glick's identification of B lymphocytes [15].

Finally, peer review can be all too easily abused. Reviewers can steal ideas and present them as their own or produce an unjustly harsh review to block or at least slow down the publication of the ideas of a competitor. These have all happened. Drummond Rennie tells the story of a paper he sent, when deputy editor of the *New England Journal of Medicine*, for review to Vijay Soman [16]. Having produced a critical review of the paper, Soman copied some of the paragraphs and submitted it to another journal, the *American Journal of Medicine*. This journal, by coincidence, sent it for review to the boss of the author of the plagiarised paper. She realised that she had been plagiarised and objected strongly. She threatened to denounce Soman but was advised against it. Eventually, however, Soman was discovered to have invented data and patients and left the country.

Improving peer review

Peer review is often compared with democracy in being the least bad system available, and attempts have been made to improve peer review - by blinding reviewers to the identity of authors, opening up the process so that authors and possibly even readers know the identity of the reviewers, and training reviewers. In summary, none of these methods have made much difference [17,18].

Alternatives to pre-publication peer review

For journal peer review the alternative is to publish everything and then let the world decide what is important. This is possible because of the internet, and Charles Leadbeater has illustrated how we have moved from a world of 'filter then publish' to one of 'publish then filter' and a world of 'I think' to one of 'We think' [19]. The problem with filtering before publishing, peer review, is that it is an ineffective, slow, expensive, biased, inefficient, anti-innovatory, and easily abused lottery: the important is just as likely to be filtered out as the unimportant. The sooner we can let the 'real' peer review of post-publication peer review get to work the better.

Fabio Casati puts it thus: 'If you and I include this paper in our journals [our personal collections], we are giving it value....When this is done by hundreds of people like us, we're using the selection power of the entire community to value the contribution. Interesting papers will rise above the noise.' This is 'we think' rather than what a few arbitrarily selected reviewers think.

The problem of finding an alternative to peer review of grants is more difficult - because clearly there are not the resources to fund every grant proposal. But it may be more important to try and find an alternative - such as giving highly successful scientists funds to pursue what

they want - because the anti-innovatory nature of peer review may mean that important science does not get done.

Barriers to change

I recently debated peer review in front of around 80 people from the Association of Learned and Scholarly Publishers. Unsurprisingly, I was arguing against peer review. Nobody agreed with my position before my talk - and nobody agreed with me afterwards. These editors and publishers were 100% in favour of peer review. The majority of scientists are also strongly in favour of peer review, although it is less than 100%.

Why are people so strongly in favour of peer review? One argument is that we have to have a mechanism, albeit an imperfect one, to sort science - otherwise people will be overwhelmed with information, much of it poor. My responses are this is the case already and that far from sorting studies into the important and unimportant the present system delivers misleading signals by giving excessive prominence to the 'scientifically sexy' [20]. I am in favour of sorting, but I think that this works better after publication when hundreds of minds and publications rather than just one or two decide what they think important.

Another argument in favour of peer review, particularly in medicine, is that it stops people being misled. Unfortunately, it does not, as I have illustrated. Furthermore, many results are made available first through conferences and the mass media - so that even if peer review was effective it could not prevent the dissemination of misleading results and conclusions.

My fear is that the real barrier to change is vested interest. That £1.9 billion cost of peer review is a great many jobs, and, more importantly, it is seen as an essential part of the £24 billion industry of publishing, distributing, and accessing journal articles, which itself is 14% of the costs of undertaking, communicating, and reading the results of research. This is not only a great many jobs but also considerable revenue and profits for commercial publishers and scientific societies that own journals.

But just think what might be done if we were to liberate the nearly £2 billion spent on peer review.

Competing interests

RS was the editor of the *British Medical Journal* and the chief executive of the BMJ Publishing Group. He is now a member of the board of the Public Library of Science and the editor of *Cases Journal*, which uses a minimal peer review system. Richard Smith is the director of the UnitedHealth Chronic Disease Initiative.

Acknowledgements

This article has been published as part of *Breast Cancer Research* Volume 12 Supplement 4, 2010: Controversies in Breast Cancer 2010. The full contents

of the supplement are available online at <http://breast-cancer-research.com/supplements/12/S4>

Published: 20 December 2010

References

1. Smith R: Enter the 'liquid journal', quoting F Casati [<http://blogs.bmj.com/bmj/2010/08/05/richard-smith-enter-the-%E2%80%9Cliquid-journal%E2%80%9D/>]
2. Jefferson T, Rudin M, Brodny Folse S, Davidoff F: **Editorial peer review for improving the quality of reports of biomedical studies.** *Cochrane Database Syst Rev* 2007:MR000016.
3. Demicheli V, Di Pietrantonj C: **Peer review for improving the quality of grant applications.** *Cochrane Database Syst Rev* 2007:MR000003.
4. Ioannidis JPA: **Why most published research findings are false.** *PLoS Med* 2005, **2**:e124.
5. Haynes RB: Where's the meat in clinical journals? *ACP J Club* 1993, **119**:A22-A23.
6. Altman DG: **Poor-quality medical research: what can journals do?** *JAMA* 2002, **287**:2765-2767.
7. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, Berelowitz M, Dhillon AP, Thomson MA, Harvey P, Valentine A, Davies SE, Walker-Smith JA: **Ileal-lymphoid-nodular hyperplasia, non-specific colitis and pervasive developmental disorder in children.** *Lancet* 1998, **351**:637-641.
8. Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, Davis B, Day R, Ferraz MB, Hawkey CJ, Hochberg MC, Kvien TK, Schnitzer TJ; VIGOR Study Group: **Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis.** VIGOR Study Group. *N Engl J Med* 2000, **343**:1520-1528.
9. Altman DG: **The scandal of poor medical research.** *BMJ* 1994, **308**:283-284.10. Research Information Network: **Activities, costs, and funding flows in the scholarly communications system** 2008 [<http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/activities-costs-and-funding-flows-scholarly-commu>]
11. Lock S: *A Difficult Balance: Editorial Peer Review in Medicine.* London: Nuffield Provincials Hospital Trust; 1985.
12. Rothwell PM, Martyn C: **Reproducibility of peer review in clinical neuroscience - is agreement between reviewers any greater than would be expected by chance alone?** *Brain* 2000, **123**:1964-1969.
13. 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.
14. Peters D, Ceci S: **Peer-review practices of psychological journals: The fate of submitted articles, submitted again.** *Behav Brain Sci* 1982, **5**:187-255.
15. Horrobin DF: **The philosophical basis of peer review and the suppression of innovation.** *JAMA* 1990, **263**:1438-1441.
16. Rennie D, Gunsalus CK: **Regulations on scientific misconduct: lessons from the US experience.** In *Fraud and Misconduct in Biomedical Research*. 3rd edition. Edited by Lock S, Wells F, Farthing M. London: BMJ Books; 2001:13-31.
17. Smith R: *Peer Review: a Flawed Process at the Heart of Science and Journals. The Trouble With Medical Journals.* London: RSM Press; 2006.
18. Smith R: **Peer review: a flawed process at the heart of science and journals.** *J R Soc Med* 2006, **99**:178-182.
19. Leadbeater C: *We Think: Mass Innovation Not Mass Production: the Power of Mass Creativity.* London: Profile; 2008.
20. Young NS, Ioannidis JPA, Al-Ubaydli O: **Why current publication practices may distort science.** *PLoS Med* 2008, **5**:e201.

doi:10.1186/bcr2742

Cite this article as: Smith R: Classical peer review: an empty gun. *Breast Cancer Research* 2010, **12**(Suppl 4):S13.