Review

Peer review of grant applications: what do we know?

Simon Wessely

Resources for scientific research are scarce. Since the 1940s and 1950s, peer review has become the principal mechanism that guides the use of these resources.[1] For many years, peer review was shrouded in secrecy, but the past 10 years has seen the emergence of empirical data on this subject. However, most of these data relate to scientific publication (Kostoff R Research programme peer review: principles, practices, protocols. http://www.dtic.mil/dtic/kostoff/index.html 1997). For example, at the 1997 International Congress on Peer Review in Prague, there were 92 abstracts about journals, but only one about grants. Nevertheless, the peer review of grant proposals may be more relevant to the health of science than publication practices. Good papers will get published somewhere, and so will bad papers, whereas applications for grants that do not succeed represent science that is not done.

I do not discuss the results of grant allocation, but focus on peer review as it applies to awarding grants. I obtained information from an extensive literature search and included all publications with empirical data on any feature of the peer-review process as applied to grant allocation. 121 publications were assessed, of which 61 met the inclusion criteria. For reasons of space, details of search strategy and the full list of papers are available from *The Lancet* and the journal website (http://www.thelancet.com). A full-length version of the paper is published elsewhere.[2]

Criticisms of peer review of research proposals address three issues: equity, efficiency, and failure to promote the best science. I shall address each in turn.

Is peer review fair?

Do researchers believe that peer review is a fair system? Surveys of grant applicants have been carried out at the National Cancer Institute (NCI), National Science Foundation (NSF), National Institutes of Health (NIH), Australian Research Council, and the Indian National Institute of Science. The most common complaint is of bias against lesser known institutions and unorthodox research, particularly from an old-boys network. The quality of the review is less contentious—at the NCI only 20% of grant applicants thought reviewers were "incomplete, inacurrate, and/or shoddy". All surveys find a link between satisfaction and grant success, but most contain a bias because successful applicants are more likely to respond than those who fail. Two surveys carried out at the NIH and the Australian Research Council looked at unsuccessful applicants only, who were more likely to complain that reviewers' comments were inconsistent, inadequate, or unfair. Even these applicants showed overwhelming support for the principle of peer review. Thus, applicants endorse this principle, although many have practical criticisms. But what evidence exists to support these criticisms?

Applicants frequently complain that reviewers are not specialists in the relevant field—ie, they are not true "peers". There is empirical evidence to support this perception,[3,4] particularly in narrow specialties,[5] but is this a deficiency? Olsson and Kennedy[6] showed that reviewers usually consult relevant references whilst reviewing grants.[6] It might be sufficient to be aware of the methodological issues in scientific research, rather than be an authority in the particular area of the grant. This feature of reviewer expertise has not been empirically tested. Some institutions routinely assess and record the quality of referees' reviews, but how and with what consequences is not clear.

Even if referees are not always authorities in the required specialty, does this introduce bias? NIH reviewers gave the most favourable ratings to proposals closest to their own area of interest, whereas reviewers from two other agencies did the opposite.[7] Some agencies, such as the Dutch Technology Foundation, encourage non-peers to take part in the process to give a wider view of "how this research may contribute to society".[8]

Is there bias against lesser known individuals and institutions? In the published work on grants, there is limited evidence that the choice of reviewers reflects this bias.[9] In 1974, reviewers at the NSF were more likely than applicants to come from top-ranked departments,[10,11] but 10 years later, the situation was reversed.[7] Even at the NSF, there was no evidence to link reviewer affiliation and institutional bias.[12] Cole[11] found no evidence of bias against lesser known institutions or younger investigators. Another study found that the rating of well known and unknown applicants were similar at the NIH (but not at the NSF).[7]

A related issue is that of "cronyism". Fuhrer and Grabois[3] reported that applicants who were funded in one grant round subsequently recognised more of the referees' names than unsuccessful applicants, which might suggest cronyism. Similarly, a Swedish study of grants for postdoctoral work found that applicants who were affiliated with a committee member were more likely to be successful, even though reviewers from the applicant's host institution were not allowed to take part in the assessment process.[13] A Brazilian study found that when principal applicants had similar measures of productivity, their chance of success increased if the funding board contained a member of the same institution.[14] Cronyism is also an issue in countries with small scientific communities;[15,16] for example, in Australia some panels are dominated by a few local academics.[17] On the other hand, there is scant evidence that reviewers from top-rated departments favour proposals from similarly prestigious

Lancet 1998; **352:** 301–05

Department of Psychological Medicine, King's College School of Medicine and the Institute of Psychiatry, London SE5 8AF, UK
(Prof S Wessely FRCP)

sources, the opposite commonly occurs,[11] which suggests jealousy rather than cronyism. In a survey of NSF reviewers, some referees admitted that favouritism and professional jealousy occasionally crept into the review process.[18]

There is no simple relationship between reviewer and reviewed. Even if most reviewers at the main US agencies did not have direct knowledge of the applicants, most had an acquaintance in common[7]—the so-called invisible college of science. That reviewers and applicants have links is not necessarily a problem. Indeed, some applicants objected when panel members familiar with their work were excluded—ie, that the old-boy system did not operate.[1] If there is reviewer bias, it may be recognised. One observational study indicated that review panels are aware of potential conflicts of biases among referees.[19] It is also unclear whether cronyism is an artefact of the better applicants coming from the better institutions which supply the better reviewers. One study showed that the more prestigious the applicant's department, the better the rating, but adjustment for the individual's track record removed this association.[7] The fundamental dilemma is the trade off between choosing reviewers who are peers and the increased risk of a conflict of interest that results from that choice.

Is peer review biased against women? In the past, sexism was not a common criticism of peer review, raised by only 4% of the dissatisfied NSF applicants.[20] Similarly, Cole[11] found no evidence of sex bias. If anything, selection of panel members at the main US institutions is biased towards women, although they are less likely than men to be external reviewers.[7] Men, however, received higher scores than women both at the NIH and the NSF, with proposals from men more than twice as likely as those from women to be funded at the NSF.[7] The strongest evidence of sex bias was provided by a Swedish study of applications for postdoctoral fellowships.[13] Women applicants received lower scores than their male counterparts after peer review of their respective applications. This scoring did not reflect the quality of the women's work, as judged by the total impact of their publications, which was a composite measure derived from the number of publications of each applicant, summed with the impact factor of the journals in which the publication appeared. There are strong reservations about using the impact factor of journals to measure an individual's productivity, although that is what the Swedish panellists did,[13] but these findings were also robust when other measures were used. The implications of these findings are not clear. A 1997 survey by the Wellcome Trust found no evidence of sex bias—successful and unsuccessful female applicants had similar publication profiles to their male counterparts, although there was evidence that women were less likely than men to apply for grants in the first place. Men and women were equally likely to receive research and career development fellowships at the UK Medical Research Council, although there is no information about whether the fellowships were of equal scientific merit. There are no studies of grant reviews blinded to the sex of the applicant.

Does the peer-review system operate against younger researchers?. In the 1970s the reverse was the case,[21] but since then the number of younger (≤36 years) scientists who apply for support from the NIH has decreased and the age at which a person gets their first grant is increasing (NIH Report of the Committee on Rating of Grant Applications. http//www.nih.gov/grants/peer/rga.txt 1996.) The importance of this observation is not clear without knowledge of the structure of the base population. Age is also confounded by experience. Overall, studies confirm that age has a minor role in grant success at the NIH, the NSH, and the Swedish Research Council for Engineering Science, and none at all at the Science and Engineering Research Council or Wellcome Trust.

Many other biases have been claimed. Reviewers' responses were more likely to be favourable when they dealt with their own specialty,[22] just as reviewers are more likely to cite their own discipline within the context of general reviews, a possible interdisciplinary bias. There is scant evidence, however, to suggest bias against clinical, as opposed to molecular, research.[23] Another bias, supported by my own observation, is that grants discussed early in a review session tend to be assessed more thoroughly and critically than those reviewed later. The NIH seems to have made a similar anecdotal observation, but there are no data to confirm this.[24]

Theft of ideas by reviewers from the grants they review is another criticism. The only quantitative assessment was done in 1978, and found that five of 126 failed applicants to the NIH alleged some theft of ideas.[25] There are no recent data. In the USA, the courts will probably have an increasing role in the arbitration of future claims.[26]

Reliability of peer review

Are the ratings of peer reviews reliable? Several studies observed weak correlations between most reviewers' scores of grant proposals.[27] Reliability is usually better in straightforward decisions (fund or not fund) than for more complex decisions. In Cole's study of peer review at the NSF,[11] there was agreement between genuine reviewers and a surrogate panel—the sham panel supported 75% of the decisions made by the real panel. There was substantial variation in the reviewers' ratings, however, emphasised by Cicchetti's reanalysis[28] which showed only slight agreement in the scores allocated by the reviewers.[20] As with journals, there was most agreement for reviews of poor-quality rather than good-quality grants.[28] Cicchetti concluded that reviewers show greater reliability in the decision to reject rather than accept,[28] although the opposite conclusion was reported in an earlier study by the American Heart Association.[6]

Cole[11] found no evidence of systematic bias or inequity in peer review, but highlights the degree of randomness that characterises the process which reflects honest disagreement among reviewers. Cole and others argue that high reliability is possible only when there is a "single, agreed upon dogma",[11] an undesirable situation. Hence consensus at the frontiers of science is as elusive as in other forms of human judgment.[29]

What are the costs of peer review?

Much time is spent writing and reviewing grant proposals. Over 160 000 reviews are provided by 50 000 reviewers to the NSF. The NIH receives 40 000 applications every year, with panel reviewers who devote 30–40 days to the task of reviewing. The 1989 Boden Report estimated that the UK research councils used 25 477 days of reviewer time, or 115 reviewer years. In 1961, Leo Szilard imagined a situation in which the combination of the time required to prepare an application, and the chance of its success, meant that a scientist's time would consist solely of writing applications with no time left for actually doing the research itself. This dystopian vision has moved ever closer, since nearly all funding bodies report that the proportion of funded grants has decreased. Many grant-giving bodies encounter difficulties in persuading scientists to give up time to review grant applications; the Australian Research Council now makes agreement to undertake reviewing duties a condition of awarding a grant.

Does peer review serve the best interests of science?

Peer review is frequently censured for its inherent conservatism and bias against speculative or innovative research. Those who write grant proposals agree,[18,25] and may deliberately underplay the innovative parts of their proposals.[19,20] The most trenchant critic, David Horrobin,[30] accepts that the peer-review system is generally fair, and by implication agrees that overall it avoids wasting resources on bad science. But he argues that such considerations are irrelevant, since it is the fate of the minority of innovative research projects that provides the true test of the peer-review system. This view is supported by Yalow,[29] a Nobel laureate, who contended that the need to promote scientific revolutions is opposed to the outcome of peer review. But substantiating this claim is difficult, and rests on retrospective case studies. 10% of the authors of papers that became "citation classics" reported initial difficulties with publication,[32] but that they are seminal papers implies that their findings were not suppressed. There is no public data on similar difficulties with grant peer review. We do not know about research that was important yet never done, and we can only guess about how innovative research might have prospered under different scenarios. In a simulation study with NIH personnel, grant applications rated as "solid and well tried" received slightly more favourable scores than "experimental procedures innovative but untested".[33] Similarly, at the principal research council in Germany new proposals received both the longest and most critical reviews.[27]

Another approach is to look at the fate of research proposals rejected by peer review. A proportion of rejected proposals are still completed. At the NCI, two-thirds of unsuccessful applicants pursued the same line of research and most were eventually published.[34] In a follow-up of a sample of projects rejected by the NIH in 1970–71, 22% of proposals were carried out without substantial changes, whereas 43% were abandoned.[25] Similarly, in the NSF survey, 48% of researchers who were not awarded funds said they stopped that line of research.[1] Further cohort studies of unfunded proposals are needed. Such studies will, however, always be difficult to interpret—do they show how peer review prevents resources from being wasted on bad science, or do they reveal the blinkered conservative preferences of senior reviewers who stifle innovation and destroy the morale of promising younger scientists? We cannot say.

The opposite perspective is provided by research funded by grants awarded after peer review of the proposed research, which lies outside the scope of this review. As one might expect, three studies note that successful grant applicants are more productive than unsuccessful applicants, with one exception.[14] The limitations of this perspective are obvious, and shed little light on the effectiveness of peer review. There are no direct studies that link peer review and the specific outcomes and impact of research funded or unfunded.

There have been many attempts to respond to the charge of failure to support innovation. For example, in the past, the NIH opposed separate ratings of creativity on the grounds that all research is innovative and creative, but because of repeated criticisms creativity will be a specific item in the Division of Research Grants study sections. Several institutions have set up programmes to offer short-term support for high-risk proposals, though their impact is not known.

Can peer review be improved?

Could the blinding of referees to applicants and their institutions improve quality? This is the most frequently researched topic in journal peer-review studies, but is contentious, not least because it is difficult to achieve in practice. There are even greater objections to blinding becoming the norm in grant review. Track record is a predictor of grant success,[7,35] but it is unclear how much it actually influences reviewers. Cole's simulation studies[11] confirmed that the characteristics of the proposal were more important than those of the applicant, and his comparison of blinded and unblinded proposals showed that past publication affected the overall rating, but only to a minor degree. In another experimental simulation, in which various characteristics of hypothetical grant reviews were manipulated, funding decisions were influenced by information on the scientific relevance of the proposal, but not by whether the researchers were rated as "highly respected in the field" rather than "new but promising".[33] Overall, the omission of information on the reputation of the researcher slightly reduced the chances of funding. Cole[11] concluded "if reviewers are being influenced at all by past performance and reputations of principal investigators, the influence does not seem to be very strong". If this conclusion is correct, blinding is unnecessary. Equally important, the reviewers found the blinded proposals unreadable, and firmly rejected the scheme;[11] most applicants to the NCI were also against the blinding.[34]

Should reviewers sign their reports? This question is the subject of controlled trials in the publications fields, and has occasionally been suggested for grant reviewing. The strongest objection comes from Hull's study of the scientific process.[36] He showed that reviewers in a specialist field used the cloak of anonymity to voice appropriate criticisms of the work of friends and colleagues, concerns which would otherwise have remained unexpressed. No major institution has plans to alter the current status quo.

If reliability is a problem, can it be improved? Increasing the number of referees is one method,[28] but would also increase the likelihood of descrepant reviews, and, in turn, may lead to application dissatisfaction. This approach would also reduce efficiency. Another way is to make rating criteria more explicit. After consultation with psychometric experts and reviewing the literature on decision making, a 1996 NIH panel concluded that global judgments of grant quality were unreliable, and that instead reviewers should make separate ratings of various criteria (significance, approach, and feasibility). However, after further discussion among panellists, global ratings were retained. Restricting NIH reviewers to fewer incremental categories in their ratings was thought to

Limitations of bibliometric methods

Citation counts are a measure of quantity, not quality.
Publication practices differ between disciplines.
Bias towards English language publications.
Influenced by self citation and salami publications.
Do not distinguish between positive and negative citations.
Time lag between work completed, work published, and citation.
Complications with group and papers and multiauthorship.
Sensitive to mistakes in spelling, initials, &c.
Ignores non-journal (electronic) methods of communications.
Patterns of communication change with time.

prevent "bunching" of scores, but did not.[37] There also seems to be little difference in outcome between the conventional method of adding item scores and multiplying them.[38]

Few studies have addressed the relative merits of different peer-review procedures, such as the internal (panel) versus external (ad hoc or postal) reviewers. One study reported that detailed external reviewing did not alter the rank order ascribed by the immediate "in house" screening operated by an arthritis grant-giving body,[39] and two other studies found that external mail reviewers did not alter the rating given by panel reviewers[40] or contributed far less to the final decision.[41] By contrast, Das and Froehlich[42] found a high level of agreement among reviewers of programme and centre grants when the rating was preceded by discussion.[42] Some grant-giving bodies, such as the Australian National Health and Medical Research Council, use interviews as well as written evaluations. In the only empirical study, reviewers who were randomly allocated to meeting the principal investigator tended to give more favourable ratings than those who had not.[43]

The most popular way of improving efficacy has been to introduce triage processes. At the NIH, a pilot study of reviewers suggested triage was still fair[44] and did not discriminate against ethnic minorities.[24] Statistical modelling and empirical data from the NCI showed that using a five-member subcommittee of a larger group to triage applications was efficient and effective. An alternative has been to remove the fixed closing dates required by nearly all organisations; anecdotal evidence from the Engineering and Physical Sciences Research Council suggests this approach has led to a decrease in numbers and an increase in quality. In another effort to reduce costs, this time to the applicants, the NIH proposes to simplify the amount of financial information requested until the grant is awarded.[24] Other suggestions include adjusting individual reviewers' scores according to their previous performance, paying referees, and restricting reviewers from receiving grants from the same source.

Should peer review be replaced?

Many alternatives to peer review have been suggested, the most common involves bibliometrics—the use of mathematical models of scientific productivity, based on the concept that scientific work results in scientific publication. Although this alternative brings apparent mathematical rigour to the process of review, it has major limitations (panel). Bibliometric methods are post-hoc measures, which can not be used until several years after the completion of a research project, let alone when a grant is awarded, and are at best only proxy measures of scientific excellence.

There are also concerns about national bias in citation practices. The results of citation analysis and peer review are not closely correlated, as shown by an analysis of 75 proposals submitted to the Dutch Technology Foundation. However, given that the two systems measure different constructs, this should not be expected. Citation analysis seems better suited for review of existing programmes or institutions, rather than individuals, and overzealous use of such measures will introduce bias against younger researchers and innovations. Despite these drawbacks, citation is likely to have an increasing role as an adjunct to peer review in the decision-making processes of many of the European funding organisations.

Other alternatives to replace peer review include awarding of grants at random or after a lottery, cash prizes to stimulate research in key areas, random selection of reviewers from a pool, or a system of professional reviewers. There is one historical precedent for funding by cash prizes.[45] Such suggestions tend to be expressions of disquiet with the current systems rather than practical suggestions for reform.

Conclusion

The main charge against peer review, that of institutional or sex bias, is generally unfounded, with a few exceptions. There seems to be no such thing as the perfect referee. Those too close to the subject may be influenced by jealousy or cronyism, whereas more distant referees may not have the required expertise. The use of international referees is frequently proposed as a way to reduce conflict of interest and jealousy, but off-the-record observations suggest that they tend to produce more favourable and less rigorous evaluations. Perhaps some competition is a spur to critical appraisal. There seems no substitute for a grants officer who knows the strengths and weaknesses of their referees.

Lack of reliability is a limited problem. Some unreliability may be due to lack of reviewer expertise. If the work published in journals is to be believed, some may also result from the increasing age of the reviewer, but much results from the lack of consensus in areas at the frontiers of knowledge. Efforts by institutions to ensure a wide range of reviewers and viewpoints, usually agreed to be desirable, will also reduce statistical reliability.

Triage processes to reduce time spent on uncompetitive proposals should become routine. Asking for general estimates of costs and requiring detailed financial data only when a grant is awarded, seems another sensible reform. Electronic forms and electronic refereeing would also be improvements.

Randomised controlled trials are needed to assess the role of blinding, feedback, and the balance of external and internal reviewers, as well as sex and institutional bias. The absence of controlled trials in this area of scientific decision making is ironic. Until such trials are conducted, proposals not supported by the research community, such as the withdrawal of referee anonymity, should not be implemented. Similarly, increased use of bibliometric data could create more problems than it solves.

This review does not substantiate claims that peer review is so flawed, biased, or corrupt that it needs to be replaced in its entirety. As with all human systems, however, perceptions of bias are common, and individual injustices do occur. Some institutions now recognise the shortcomings of peer review and have improved the transparency of the process and introduced a limited appeals procedure. Such systems are likely to reduce unsubstantiated allegations of bias and improve confidence in the system, but will inevitably increase costs. No doubt, *The Lancet* will inform us, in due course, of the effectiveness of their own ombudsman.

Neither ombudsmen nor randomised trials will address the more difficult, and arguably more important, question: does peer review help scientists make important discoveries that stand the test of time? The answer is a judgment which, by definition, can only be made with hindsight. Similarly, the most important charge made against peer review is that it impedes innovation, but this charge remains unproven, and possibly unprovable.

The scientific community's interest in peer review has a pragmatic basis—the links between grants and the structures of scientific careers. Obtaining grants is increasingly an end in itself, rather than a means to an end. Since grants are so important for scientists, it is proper to obtain further empirical data on questions such as equity and efficiency, and to continue to improve the transparency of the process. Such research, however, can only answer short-term questions rather than the real purpose of scientific endeavours.

I would like to acknowledge the various members of staff of the National Institutes of Health, Wellcome Trust, Medical Research Council (UK), NHS R & D Programme, European Science Foundation, and the Australian Research Council for assistance in answering questions and locating sources. I am grateful to Agnes Wold and Fiona Wood for additional information, John Overbeke for assistance with references, and Lindy Castell, Iain Chalmers, Bruce Charlton, David Horrobin, Ron Kostoff, Ron Stamp, and Fiona Wood for comments.

References

1 Chubin D, Hackett E. Peerless science: peer review and US science policy. Albany: SUNY Press, 1990.
2 Wessely S, Wood F. Peer review of grant applications. In: Godlee F, Jefferson T, eds. Peer review in health care. London: BMA Publications, 1998.
3 Fuhrer MJ, Grabois M. Grant application and review procedures of the National Institute of Handicapped Research: survey of applicant and peer reviewer opinions. *Arch Phys Med Rehabil* 1985; **66:** 318–21.
4 Glantz SA, Bero LA. Inappropriate and appropriate selection of "peers" in grant review. *J Am Med Assoc* 1994; **272:** 114–16.
5 Olsson C, Kennedy W. Urology peer review at the National Institutes of Health. *J Urol* 1995; **154:** 1866–69.
6 Weiner S, Urivetsky M, Bregman D, et al. Peer review: inter-reviewer agreement during evaluation of research grant evaluations. *Clin Res* 1977; **25:** 306–11.
7 Anon. Peer review: reforms needed to ensure fairness in federal agency grant selection. Washington, DC: United States General Accounting Office, 1994.
8 Van den Beemt F. The right mix: review by peers as well as by highly qualified persons (non peers). Peer Review Process: Australian Research Council commissioned report, no 54. Canberra: Australian Research Council, 1997: 153–64.
9 Westhead E. Peer review: distribution of reviewers. *Science* 1975; **188:** 204.
10 Groenveld L, Koller N, Mullins N. The advisers of the United States National Science Foundation. *Soc Stud Sci* 1975; **5:** 343–54.
11 Cole S. Making science: between nature and society. Cambridge, MA: Harvard University Press, 1992.
12 Anon. Report of the National Science Board to the Subcommittee on Science, Research and Technology of the Committee on Science and Technology, US House of Representatives, regarding peer review procedures at the National Science Foundation. Washington, DC: National Science Foundation, 1977.
13 Wenneras C, Wold A. Nepotism and sexism in peer-review. *Nature* 1997; **387:** 341–43.
14 Fonseca L, Rangel V, Lustosa P, et al. Productivity versus promised results: one of the dilemmas of biotechnology in Brazil. *Braz J Med Biol Res* 1994; **27:** 2709–20.
15 Pouris A. Peer review in scientifically small countries. *R & D Management* 1988; **18:** 333–40.
16 Anon. Peers on peers: allocation policy and review procedures at the Swedish Research Council for Engineering Sciences. Stockholm: Swedish Research Council for Engineering Sciences (TPR), 1997.
17 Wood F. The peer review process. Canberra: Government Printing Service, Australian Research Council, 1997.
18 Hensler D. Perceptions of the National Heart Foundation peer review process: a report on a survey of NSF reviewers and applicants. Washington DC: Committee on Peer Review, National Science Board, 1976.
19 Travis G, Collins H. New light on old boys-cognitive and institutional particularism in the peer review system. *Sci Technol Hum Values* 1991; **16:** 322–41.
20 McCullough J. First comprehensive survey of NSF applicants focuses on their concerns about proposal review. *Sci Technol Hum Values* 1989; **14:** 78–88.
21 Kalberer J. Treatment of young investigators in the National Cancer Program. *J Natl Cancer Inst* 1979; **63:** 1097–103.
22 Porter A, Rossini F. Peer review of interdisciplinary proposals. *Sci Technol Hum Values* 1985; **10:** 33–38.
23 Marshall E. Does NIH shortchange clinicians? *Science* 1994; **265:** 20–21.
24 McCullough J. Changes in the proposal review processes used by the National Science Foundation and the National Institutes of Health. In: Wood F, ed. The peer review process. Commissioned report no 54. Canberra: National Board of Employment, Education and Training, 1997: 125–51.
25 Carter G, Cooper W, Lai C, March D. The consequences of unfunded NIH applications for the investigator and his research. Santa Monica, CA: Rand Report R-2229-NIH, 1978.
26 McGarity T. Peer review in awarding federal grants in the arts and sciences. *High Technol Law J* 1994; **9:** 1–92.
27 Hartmann I, Neidhardt F. Peer review at the Deutsche Forschungsgemeinschaft. *Scientometrics* 1990; **19:** 419–25.
28 Cicchetti D. The reliability of peer review for manuscript and grant submissions: a cross-disciplinary investigation. *Behav Brain Sci* 1991; **14:** 119–86.
29 Kahneman D, Slovic P, Tversky A. Judgement under uncertainty. Cambridge: Cambridge University Press, 1982.
30 Horrobin DG. Peer review of grant applications: a harbinger for mediocrity in clinical research? *Lancet* 1996; **348:** 1293–95.
31 Yalow R. Is subterfuge consistent with good science? *Bull Sci Technol Soc* 1982; **2:** 401–04.
32 Campanario J. Compensation for the scientist—sometimes it is hard to publish papers that are later highly cited. *Soc Stud Sci* 1993; **23:** 342–63.
33 Norman K. Importance of factors in the review of grant proposals. *J Appl Psychol* 1986; **71:** 156–62.
34 Gillespie G, Chubin D, Kurzon G. Experience with the NIH peer review: researcher's cynicism and desire for change. *Sci Technol Hum Values* 1985; **10:** 44–53.
35 Abrams P. The predictive ability of peer review of grant proposals—the case of ecology and the United States National Science Foundation. *Soc Stud Sci* 1991; **21:** 111–32.
36 Hull D. Science as a process. Chicago: Chicago University Press, 1988.
37 Green JG, Calhoun F, Nierzwicki L, Brackett J, Meier P. Rating intervals: an experiment in peer review. *FASEB J* 1989; **3:** 1987–92.
38 Bunch DS, Cardus D, Fuhrer MJ, Thrall RM. When are additive models valid for evaluating proposed research? *Methods Inf Med* 1989; **28:** 168–77.
39 Russell AS, Thorn BD, Grace M. Peer review: a simplified approach. *J Rheumatol* 1983; **10:** 479–81.
40 Klahr D. Insiders, outsiders and efficiency in a National Science Foundation panel. *Am Psychol* 1985; **40:** 148–54.
41 Hodgson C. Evaluation of cardiovascular grant-in-aid applications by peer review: influence of internal and external reviewers and committees. *Can J Cardiol* 1995; **11:** 864–68.
42 Das NK, Froehlich LA. Quantitative evaluation of peer review of program project and center applications in allergy and immunology. *J Clin Immunol* 1985; **5:** 220–27.
43 Noble J. Peer review: quality control of applied social research. *Science* 1974; **185:** 916–21.
44 Gavaghan H. Cautious welcome to NIH peer review reforms. *Nature* 1994; **369:** 269.
45 Sobel D. Longitude. London: Fourth Estate, 1995.
46 Van Raan A. Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics* 1996; **36:** 397–20.