

Comment

Instructions for repair

Gregory A Petsko

Address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA.
Email: petsko@brandeis.edu

Published: 28 April 2006

Genome Biology 2006, **7**:106 (doi:10.1186/gb-2006-7-4-106)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2006/7/4/106>

© 2006 BioMed Central Ltd

Last month I wrote about the sharp decline in the success rate for scientific research proposals submitted to the US National Institutes of Health (NIH) and other agencies. That column provoked numerous responses from both administrators of the funding organizations and life scientists. The administrators, while not denying some of the problems I discussed, argued that things aren't quite as bad as they seem, and that a large part of the difficulty stems from sizeable increases in the number of grant applications and the amounts requested, rather than from poor choices in managing the doubling of the NIH budget that took place not long ago. The scientists, on the other hand, all said that things were even worse than I had claimed.

Care has to be taken in drawing conclusions from either of these sources. I'm sure that people who have experienced difficulty in obtaining funding are more likely to respond to that column than those who've had success. And administrators probably feel the need to defend themselves, and their agencies, from what they might, with some justification, see as an attack by someone who doesn't know the whole story the way they do.

Nevertheless, although I think both sets of comments are useful, I also think both largely missed the point. People who wrote to me were all concerned, in one way or another, with the amount of money available for research and how it is being allocated. That's what seems to be on everybody's minds, and it's certainly worth talking about. Whether or not we're allocating the available funding sensibly is something that ought to be engaging officials as well as researchers in an ongoing dialog about priorities in science. (But that dialog isn't taking place. Somehow it just seems easier to keep asking for more money.) Yet, that wasn't the main point of the column. What concerns me is that, whether there really is a crisis in scientific funding or not, the perception that there is – and believe me, that is the perception on the part of just about every researcher I have talked to – has crippled the peer-review system.

Peer review is the foundation of quality in science. It prevents widespread cronyism and slowly weeds out unproductive lines of inquiry. But it requires that reviewers be both fair and wise. When the perception is that there's not nearly enough money to fund even all of the highest-quality proposals, a defensive turf-protection replaces a spirit of curiosity and egalitarianism. When it seems as if the primary job of a reviewer is to eliminate most proposals rather than to fight for the good ones, nit-picking replaces generosity. When the feeling is that every dollar counts so much that no risks dare be taken, conservatism and incremental advances get rewarded at the expense of bold new ideas. And when all of these things happen – and I believe they are happening, now, in the US – then the system is broken.

Societies based on scarcity tend to become hierarchical, with a well-fed elite and starving masses. As can be seen from publicly available data (http://grants1.nih.gov/grants/financial/QA_Doubling_Period.doc), during the recent doubling of the NIH budget over a seven-year period, the number of investigators getting funded changed very little. Where did the money go? Besides a very large increase in the funding for NIH's own intramural research program, it seems to have gone to large increases in funding for established investigators who renewed their grants successfully during this period, or wrote additional ones. Instead of bringing lots of new people into the system, we ended up with more money for roughly the same set of grant holders. Now that funding is tight, those bloated operations are under tremendous pressure to at least maintain their size, which makes it even more difficult for new investigators – or new ideas – to enter the system. The average age at which a scientist receives his or her first NIH grant in the US is currently 42 for PhDs (even older for MDs), and in this time of perceived scarcity a broken peer-review system is not likely to change that.

What's the best way to fix things? It could be argued that the problem is temporary, and that when funding loosens up again, as it always has in past boom-bust cycles, peer review

will recover along with everything else. After all, that's what happened in the 1970s. No need to tamper with the system. Time will take care of the problem.

I have my doubts. There's one big difference between peer review in 1975 and peer review today: the number of senior investigators participating in the process. Back then most review panels had a preponderance of such scientists, who provided the system with institutional memory of the way things were supposed to work. Nowadays, most established investigators feel they are too busy to put in the considerable time required to deal with the glut of proposals that every panel faces. The result is that less experienced scientists, with no history of a different gestalt, are being fed into a system where fault-finding and conservatism are the norm, so when the funding situation improves, there's no guarantee that the peer-review system will improve with it. (If you doubt this, consider the former Soviet Union. When it collapsed in 1989, newer Soviet-block countries like Poland and Hungary and Czechoslovakia, where there was a generation of people who still had a memory of how a market-based economy should work, did much better than Russia, where no one alive had experienced any system but communism.) In addition, the insistence that the composition of the panels must satisfy a requirement for geographic and institutional balance means that it's hard to have a large number of top scientists on any panel, even if they wanted to serve.

So my first repair instruction is simple: Do away with the misguided concept of balance, and require that all holders of research grants serve at least one year on a reviewing panel for every five years of funding they receive, regardless of seniority. Renewal of funding would be contingent on fulfillment of this service. If there is a surplus of available talent, then grants administrators could forgive the obligation for any given five-year cycle, but the requirement would kick in again when a grant was renewed. There would need to be a mechanism to deal with people who hold multiple grants - perhaps they would only incur a single one year debt for every five years of total funding, or the length of service could scale with the total budget; these details can be worked out. The important point is to create a pool of the best researchers, and to make sure that they represent the majority on all peer-review panels. As a dyed-in-the-wool advocate of personal freedom, the coercive aspects of this suggestion do trouble me somewhat, but it isn't really all that different from the way things work in the other main form of peer review - the jury system.

My second idea for how fix things is meant to address the problem of reviewer morale. When someone is given twelve grants to review, and knows that there is only a small probability that even the best one is actually going to be funded, he or she rapidly becomes discouraged. It's even more depressing when some less knowledgeable reviewer nit-picks one's best proposals, and depression is not the best mindset from

which to make judgments. I suspect the program officers at the funding agencies must feel equally demoralized: it's no fun having to say "no" all the time, and to watch conservative study sections pass over the most exciting new ideas in favor of more of the usual. The solution, I think, is to give the program officers more autonomy in funding decisions. Some NIH institutes and centers claim that they do this, but in practice I have found that program officers rarely go against the recommendations of the reviewing panels. I suggest taking at least 10% of the budget of each institute or center and allowing the program officers to use it to fund grants that they believe to be exciting but that would otherwise miss the payline cut-off. They would need to justify each decision to the council, of course, but this suggestion would empower them to rectify some of the worst mistakes of the panels. In my experience, funding officials tend to be bright, committed individuals with a good broad knowledge of their field; I have no hesitation in giving them more autonomy. This is the way things actually work at the National Science Foundation, a funding agency that many believe has a better long-term history of supporting innovative research than does the NIH.

There also needs to be a way to improve the judgments coming out of the panels. Having more experienced reviewers would help, but it's hard to deal thoroughly and fairly with each proposal when the number being reviewed has increased so greatly. The way to solve the problem of proposal overload is to reduce proposal size. NIH proposals now are limited to 25 pages for the scientific description (that includes background and significance, progress during the past budget period, and the plan for future research). I think that should be shortened to 15 pages. If you can't describe clearly in 15 pages what you've already done, what you intend to do, and why it's important, you probably can't do it in 50.

But I think the proposals should be structured differently for different investigators. Scientists submitting their first proposal need to spend more space detailing how they are planning to carry out the work than established investigators should. In fact, I would argue that established investigators shouldn't have to describe their proposed methods in any detail at all, except if these are novel. To ask someone who has demonstrated for years that they can deliver the goods to prove that they know what they're doing is silly and borders on insulting. It also provides the nit-pickers with extra ammunition. People who tout stocks are constantly warning investors that past performance is no guarantee of future returns. But there is one area where it is: scientific research. The best predictor I know of as to whether a project will work is the track record of the principal investigator. Someone who has been consistently successful is not likely to fail, even when doing something risky. We need to stop pretending that isn't true. Most organizations that award pre- and post-doctoral fellowships spend very little time picking over the details of the applicant's research proposal,

because they know that these young people haven't had any experience writing proposals and anyway usually end up doing something different, or in a very different way, from what they propose. Instead, fellowship reviewers tend to consider the qualities of the individual to be the most important factor on which to base their judgments. I think that makes sense at all levels of science. We need to be much less concerned with the details of projects, and put our bets on people and ideas.

While we're waiting for funding levels to improve, we need additional mechanisms to get young people started. The observation that, while investigator funding went way up during the NIH budget doubling, the number of investigators changed very little, suggests that we should consider putting a cap on the size of each award so as to make more money available for funding new projects and people. This is a serious matter, because it potentially has an impact on current employees, so if we implement a cap we will need to phase it in gradually. I am not proposing that we limit the total amount of funding that an individual can have - I think if someone can justify the need for millions of dollars to do first-rate science they should be able to obtain it. But I do think that we should exercise more scrutiny in such cases, and one way to do that is to force someone who claims to need, say, a million dollars to support a project to submit two or three proposals instead of one. I also think we will do better science, as a community, if we have more individual investigator-initiated projects and fewer mega-sized 'me-too' programs. Most innovation comes from small projects by relatively new people.

Two final ideas pertain to the machinery of the reviewing process itself. Turf protection is one of the biggest problems in peer review: as fields try to survive in a time of scarce resources, they often fight to fund their own mediocre science at the expense of quality in other areas. This largely stems from the personal and professional relationships that develop among members of a particular discipline. It's less of a problem when there's more money to go around, but right now we need to fight it. Here's a heretical and possibly crazy idea: I think we should consider not allowing people to review grants in their own field. Instead, they should only be allowed to comment on any questions of technical feasibility that come up during the review. This may seem absurd, but I'm not sure it is. If we follow my suggestion to bet on people rather than projects, detailed technical expertise isn't so important. And if we have the best, most experienced people back on our review panels, they usually will have a pretty broad knowledge of genomics, or biology, or whatever the main subject is. That will allow them to assess the importance of the proposed research and the impact of the applicant's previous contributions, which I maintain are the only two criteria that really should matter. Reviewing outside one's primary area of technical expertise happens all the time on fellowship panels, and they usually make pretty

Box 1

Repair instructions

Do away with the misguided concept of balance, and require that all holders of research grants serve at least one year on a reviewing panel for every five years of funding they receive, regardless of seniority.

Give the program officers more autonomy in funding decisions.

Take at least 10% of the budget of each institute or center and allow the program officers to use it to fund grants that they believe to be exciting but that would otherwise miss the payline cut-off.

Solve the problem of proposal overload by reducing proposal size.

The scientific description (that includes background and significance, progress during the past budget period, and the plan for future research) should be limited to 15 pages.

Established investigators shouldn't have to describe their proposed methods in any detail at all, except if these are novel.

Be much less concerned with the details of projects, and put our bets on people and ideas.

We should consider, as a community, the advisability of putting a cap on the size of each award so as to make more money available for funding new projects and people.

We should also consider trying the idea of not allowing people to review grants in their own field. Instead, they should only be allowed to comment on any questions of technical feasibility that come up during the review.

Guarantee the quality of the peer-review process by reviewing the reviewers.

good decisions. After all, you don't have to be able to lay an egg in order to tell a good one - or to smell a bad one.

The second procedural change we should consider is aimed at addressing the issue of possible bias in the system. In times of scarce dollars, reviewers worried about their own chances of obtaining funding have an incentive to prevent others from being funded. Even if we assume such a thing

rarely happens, we should want to ensure that proposals are reviewed wisely as well as fairly. I think the best way to guarantee the quality of the peer-review process is to review the reviewers. The data to do so exist, because there is a record of how every member of a panel has voted on every grant. Since most panelists only read their assigned proposals in detail, we need only be concerned with how a reviewer's scoring of such applications compares with the average score awarded to those same applications by the other assigned reviewers. Abnormally high or low scores would not be damning in and of themselves (there's plenty of room for legitimate differences of opinion in science) but a consistent pattern of low or high scores could indicate either poor judgment or bias. You may wonder how to be sure about such an evaluation, but it's actually easy, because we can compare each reviewer with him or herself. Bias or territoriality should be relatively easy to detect by examining how a suspect reviewer treats the same grants when they are resubmitted after revision. Since unsuccessful applicants try hard to answer the criticisms raised by the previous review, the scores of resubmitted grant proposals should improve, on average. If a reviewer's scoring on such resubmissions remains abnormally low compared with other reviewers who are also seeing the proposal for a second time, then there is reason to question the impartiality, or the judgment, of that reviewer, and they can be eased off the panel. There might even be no need to evaluate every reviewer all the time: random checking might be all that is needed to discourage trying to rig the game.

When fear and discouragement drive peer-review decisions, then the system is broken. But it's not broken beyond repair. The suggestions I've offered here can help mend it; at the very least, I hope they will start a dialog about what should be done. The worst thing we can do as a community is to throw up our hands in despair or pretend that everything will right itself magically when more money becomes available. Peer review is too important to give up on, or be left to chance. If we're serious about funding the best science and making our profession attractive to the brightest, most creative young minds, then we need to fix the system so that it once again serves those ends. Let's get to work.