

Comment

## The system is broken

Gregory A Petsko

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

Published: 30 March 2006

*Genome Biology* 2006, **7**:105 (doi:10.1186/gb-2006-7-3-105)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2006/7/3/105>

© 2006 BioMed Central Ltd

It's not true that things have never been this bad before. They were about the same in the early 1970s. But it is true that they have never been worse. When a scientist doing work in genomics, or cell biology, or biochemistry, or immunology submits a grant proposal to the US National Institutes of Health (NIH), the largest supporter of life science research in the world, his or her chance of it being funded are at historic lows. And this situation is threatening to destroy the jewel in the crown of US science, the system of competitive peer review of research applications.

In contrast to the hierarchical system in many other countries, where research funds are often distributed to heads of departments or centers, who then dole them out to their component research groups, in the United States most university research faculty are independent entrepreneurs, who compete with one another for funding on the basis of the quality of their proposals. The competition is judged by the applicant's peers - scientists in the same general area of research. This Darwinian selection system has, for over half a century, largely guaranteed that merit, not cronyism, determines what science is supported by the federal government. The procedure is straightforward, and until now has worked remarkably well.

But I think the procedure has stopped working well because of the perception that financial support for science in the US is drying up. Thanks to the war in Iraq and tax cuts mostly for the richest Americans, federal funding for life science research, which doubled over a seven year period not long ago, has remained flat in real dollars and declined in inflation-adjusted dollars during the last few years. To make matters worse, scientists from all disciplines flocked to the NIH for support like pigs to a trough during the budget-doubling period, resulting in a huge increase in the number of submitted research proposals. And NIH administrators didn't help matters either. They seem to have assumed that the big increases in their budget would go on forever, and rather than engineer a soft landing for when the inevitable

crash came, they spent like sailors on shore leave, mostly for big new programs that benefited only a small number of investigators (Hello, Structural Genomics Initiative). And since new programs are like living creatures and fight for survival with the ferocity of a cornered wolverine, the chance that we could rid ourselves of these white elephants when budgets got tight has, of course, turned out to be zero.

With chance for support dwindling, individual investigators, the lifeblood of creative scientific research, are beginning to flee the field. I personally know of many young research students who are either going into industry or leaving science altogether because they believe that they have little possibility of being able to obtain funding were they to set up their own laboratory. And I know of an equal number of senior scientists who are going into administration or taking early retirement, not because they want to, but because they have become discouraged about the prospects for continued support.

The Bush administration and our own greed are to blame for this situation, but the immediate cause of the problem from the perspective of the individual investigator is what I see as a breakdown of the peer-review system. Unless that can be fixed, the likelihood of a turnaround, even if budget levels improve, is not good.

Peer review of applications submitted to NIH takes place in two steps. Applications for support from the NIH are evaluated initially by peer-review groups of scientists who are assigned grants to review on the basis of their expertise. The objective of this initial peer review is to determine the scientific and technical merit of the proposed research project. If the project represents a continuation of one funded previously, the productivity during that period is also considered in evaluating the competing renewal. The panels that review the proposals are called Scientific Review Groups and are managed by Scientific Review Administrators, employees of the Center for Scientific Review, one of the approximately 27

institutes and centers that are the components of the NIH. Approximately half of the proposals considered at a particular Scientific Review Group meeting will be triaged as being not competitive for funding at all. The top half are discussed in detail and are assigned priority scores: numerical ratings of scientific merit from 100 (best) to 500 (worst). The scores are converted into percentile rankings that indicate, for example, whether a grant is in the top 20% of all grants scored by that group (the 20th percentile). After the conclusion of the meeting, the Scientific Review Administrator prepares a summary statement for each discussed proposal that includes the reviewers' written comments, recommendations of the group and the priority score and percentile ranking. The summary statement is sent to the program staff of the awarding institute and to the applicant. (The second level of peer review is carried out by the NIH National Advisory Councils. These councils are composed of scientists from the extramural research community and public representatives. They are meant to ensure that the NIH receives advice from a cross-section of the US population in the process of its deliberation and decisions. Councils don't usually overturn the funding decisions of the Science Review Groups, but they do have that power.)

There is some confusion about the meaning of the percentile score awarded by Science Review Groups as compared with the success rate for a grant being funded. The success rate is the total number of grant applications that are funded in a given fiscal year divided by the number of grant applications that were peer-reviewed. The percentile is a ranking that shows the relative position of each application's priority score among all scores assigned by that particular Scientific Review Group at its last three meetings.

For a given NIH Institute, the success rate usually differs from the percentile ranks. The percentile ranks are calculated using all applications reviewed by that initial Review Group, which includes applications assigned to other NIH institutes and centers. If grants assigned to one institute tend to receive better priority scores than the NIH average, then that year more than, say, 10 percent of its grant applications will rank better than the 10th percentile. Applications that are amended and resubmitted during the same fiscal year are also only counted once in the success-rate calculations, whereas all applications, both original and amended versions, are included when the percentiles are calculated. Therefore, funding all applications with ranks better than, say, the 20th percentile will result in a success rate greater than 20 percent when revised versions of some projects are removed from the success-rate base.

For 2006 the percentile cut-off for a grant to be funded by the National Institute of Allergy and Infectious Diseases is the 14th percentile. It's the 10.5th percentile in the National Institute of Aging, the 11th percentile for the National Cancer Institute, and the 12th percentile for the National

Institute of Neurologic Diseases and Stroke. These translate into success rates in the order of slightly above 20% for most institutes, which can be compared with success rates close to 40% 7-10 years ago. (Most institutes try to give young investigators a break by setting the 'payline' about 2-5 percentile points higher for their proposals, resulting in a slightly higher success rate for first-timers.)

A drop in success rate of 50% is nothing to be happy about. But the number that really matters for peer reviewers is the percentile ranking, because this is what the Scientific Review Group members are aware of when they review a proposal. If they know that the payline is around the 10th percentile, as it is now, then they also know that out of 100 proposals that might be reviewed at that meeting, only about 10 will get funded. And that knowledge is the problem.

Ten years ago, when grants scoring better than the 25th or sometimes even the 30th percentile were being funded, reviewers knew that most good proposals would be supported, and that if they made a mistake about a grant at the margin, they were not making a mistake about the very best science. Consequently, the tone in review-group discussions was that of constructive criticism. Reviewers tried hard to find reasons to support work, particularly by young investigators, and their comments were often encouraging and guiding. No one was afraid that if someone else were funded, it would hurt their own chances of being funded; the pie was large enough that everyone felt they had a fair chance at a slice.

Not any more. When the percentile cut-off is around 10%, reviewers are being asked to do the impossible. They have to make choices from among research proposals that they themselves have evaluated as being better than 90% of all other grants in the field. No human being can make objective distinctions between grants at that level of quality. Because, since they must, subjectivity inevitably creeps in. Now Scientific Review Group members must try to find reasons not to fund proposals. The tone of reviewing is one of nit-picking. Increasingly silly criteria are being used to distinguish between applications: one of my proposals lost points because I did not give enough detail about how I was planning to carry out a particular experimental technique. Forgive me if I was a trifle starry-eyed about it, but I really didn't think I needed to demonstrate my competence in using a method that I had invented some fifteen years before.

Of course, when funds are this tight, generosity of spirit is in danger of being replaced by unenlightened self-interest. Every funded proposal now is a direct threat to one's own grants being funded. This mentality inevitably leads to turf protection, as reviewers in a subfield look after one another's applications, even if these are not of the best quality. To the

credit of most reviewers, I haven't seen too much of this, but I've sure seen more than I saw a few years ago.

And if good grants are not funded simply because they just miss the cut-off, for whatever reason, including pure bad luck, it's not likely that there are many, if any, substantive criticisms that the investigators can address in a resubmission. Imagine how discouraging it must be to write a good proposal and see it not funded, and not to have any idea how to improve it because there's really nothing to improve. Who wants to roll the dice again with those odds?

But I think it's equally discouraging for the reviewers. If you're given 20 proposals to evaluate out of a crop of, say, 100, and you determine that 6 are of excellent quality, but you know that the probability that more than 2 of these will actually get funded is nil, how can you feel good about what you're doing? Or about your own prospects for getting funded? Or about the future of your profession? Also, with a payline this low there's a significant chance that nothing you review will get funded, making the whole, time-consuming exercise one of futility. Good people won't serve on study sections under these circumstances.

When the payline hovers around the 10th percentile, when fewer than a quarter of submitted grants are funded, and when the process of peer review has become one of trying to make judgments among things of equal quality, the system is broken. But I don't think it's broken beyond repair, at least not yet. Next month, I'll tell you how I think it can be fixed.