NIH Tunes Up Peer Review

To relieve stress on an overburdened grant system, NIH has begun asking members of its study sections to reject up to half the applications without any panel discussion

If you're waiting for your grant proposal to be reviewed by the National Institutes of Health (NIH) in a handful of areas, you could soon be in for a nasty shock. Normally, you would expect your grant to get an exhaustive peer review, and even if it was rejected, you would expect a long critique from the reviewers—a critique you could use to modify the proposal and resubmit it. Now, however, you may be unlucky enough to receive a quick rejection and a short written explanation. At least you won't be kept waiting for more than a month to get the bad news, and you can take some comfort from the fact that you won't be alone.

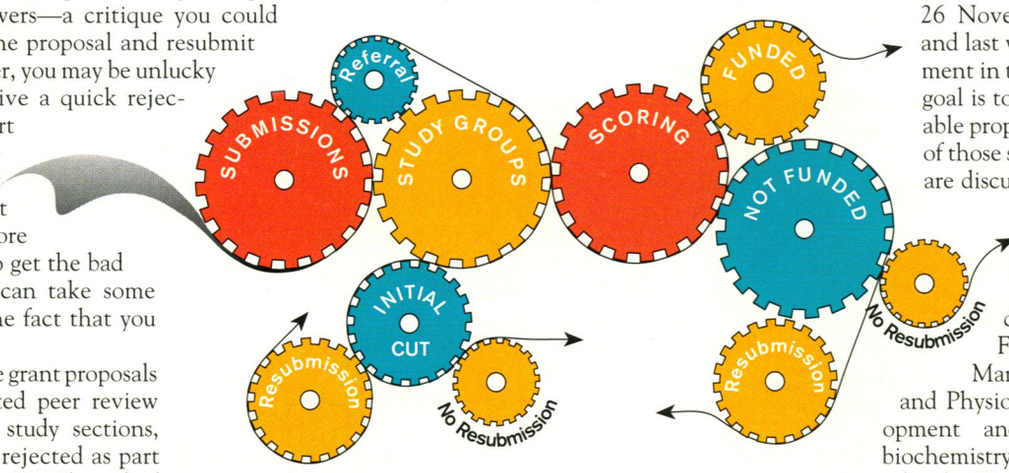
Up to half the grant proposals routed to selected peer review committees, or study sections, may be quickly rejected as part of an experiment NIH launched last week as it tries to relieve the pressure on its overloaded peer-review system. If the experiment, which NIH calls triaging, is successful, it could eventually be extended to all study sections. The goal is to allow peer reviewers to spend more time on top proposals and less effort reviewing—and re-reviewing—grants that are unlikely ever to get funded. And this could be just the first step in overhauling NIH's peer-review system, for NIH officials are also talking about refocusing study sections and relying on electronic data submission.

These changes are signs that NIH has recognized that, after nearly half a century of reliable service, the engine of peer review has begun to show signs of wear and tear and needs a tune-up. Harold Varmus, director of NIH, says reviewers complain about the "arbitrary" decisions they must make, the welter of "low-priority applications" they must review, and the number of repeat submissions they're seeing. Microbiologist Charles Moran of Emory University, who recently ended a stint on an NIH review panel, calls the experience "incredibly frustrating" and says he doesn't want to repeat it.

At its core, the problem is a fiscal one: Too many good ideas are chasing too few dollars. The heavy demand—only 21% of proposals are funded, most at less than the original request—forces study section members to pick and choose among equally good proposals. What's worse, they must make such life-or-death decisions without time to consider fully the merits of each proposal.

Moran says that NIH's system is mired in paperwork, much of it pointless. In 1993, for example, NIH received 19,072 new and competing requests for small, investigator-initiated proposals that NIH calls RO1 grants; it funded only 4121 of them. Virtually all the 14,951 applications that didn't make it were fully reviewed, critiqued, and sent back with extensive comments—even if they stood no chance of being funded. Moreover, the system is now swamped with second- and third-time visitors: The number of revised grants among the submissions has increased from 25% in 1980 to 34% in 1992. AIDS researcher Jay Levy of the University of California, San Francisco, for one, not only bemoans the time spent reviewing resubmissions but also questions the wisdom of some of the final decisions. He says people are "coming back again and again" until the review panel finally surrenders and says, "'My god, give it to him!'"

Moran and Levy—like many others—find the process particularly discouraging because of the difficulty of making distinctions among grants in the middle range—those beyond the top 10% but within the top quartile. Yet this is exactly where the ax falls when NIH draws a "pay line" dividing the fundable from nonfundable in what many reviewers consider to be an arbitrary, and, Moran adds, "depressing" process.



A grinding sound. The current peer-review machine spits out very few proposals before study sections meet and assign scores and percentile ratings to most submissions.

ILLUSTRAION: C. FABER SMITH

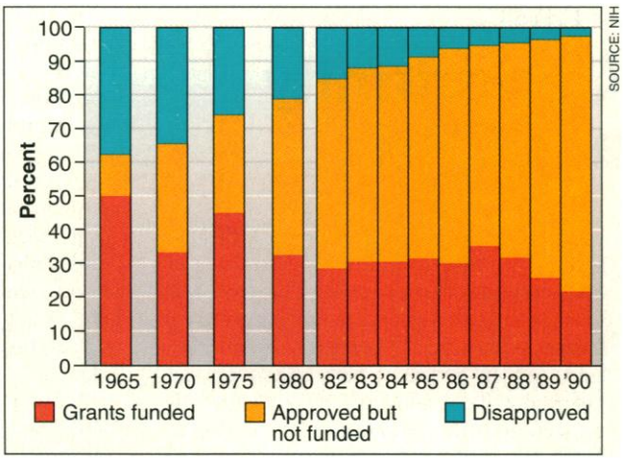
But relief—or at least an expression of sympathy—is on the way. Even before he came to NIH, Varmus had been talking about the need for improvements (*Science*, 26 November 1993, p. 1364), and last week he put the experiment in triaging into effect. The goal is to eliminate the least viable proposals—the bottom half of those submitted—before they are discussed at the meetings of the study sections. Triaging has already begun in four study sections that met during the last week of February and early March (Cellular Biology and Physiology-2, Human Development and Aging-1, Metallobiochemistry, and Experimental Virology). Within the next 2 weeks, the first letters bearing a new acronym—"NC" for "not competitive"—will be mailed to applicants who failed to make the cut. That's when Varmus and his staff will begin to hear how the community likes it.

A strong backbone

Triaging isn't entirely new to NIH. Already, according to Wendy Baldwin, who was confirmed last week as NIH's director of extramural research, the institutes routinely set aside weak proposals that arrive in response to targeted research efforts, called "requests for applications." But for the investigator-initiated grants, which NIH calls the "traditional" form of biomedical research, the blunt 50% cutoff would be new. Thirty years ago, reviewers dropped more than one-third of the applications as not viable, but in recent years the number of "disapproved" submissions has dropped below 10%. NIH staffers have tried to put more steel into reviewers' backbones, so far with little success. However, some reviewers who spoke with *Science*—including Lawrence Rothfield of the University of Connecticut and Barry Honig of Columbia—said their study sections do practice a form of triage, by agreeing not to waste time discussing the least competitive grants. This sensible approach, says

Rothfield, requires a strong chair; by endorsing its use, NIH is giving others a gentle push in the same direction.

To understand how triaging is meant to work, a quick tour through the Byzantine structure of NIH's granting system is needed. Researchers seeking new or renewed grants submit their proposals to NIH, which assigns them to one of about 100 study sections. Each proposal gets a primary and secondary reviewer, who prepare a written critique. A third reader may also examine the grant and help lead the discussion during the formal review session. The study section meets for 3 days near the NIH campus for a grueling session in which 80 proposals are typically dealt with. Every accepted proposal gets a score that is converted to percentile, after which it is forwarded to the advisory councils at each institute for funding or rejection. In 1993, only 21% of the reviewed RO1 proposals were funded.



Can't say no. Peer reviewers over the years have found it harder and harder to reject a proposal at the outset.

Within a couple of weeks after the study section's meeting, says Anthony Demsey, an official in NIH's division of research grants, applicants receive word of their score. It takes another 6 to 8 weeks for the NIH staff to compile a "pink sheet" that contains a summary of the reviewers' written critiques and comments made during the meeting. These summary sheets, according to Baldwin, were once brief, but they now run to five pages and have acquired a "tutorial" quality. Preparing them is one of the most arduous and time-consuming aspects of the review process. Triage—if it succeeds—will do away with some of the tutorial essays.

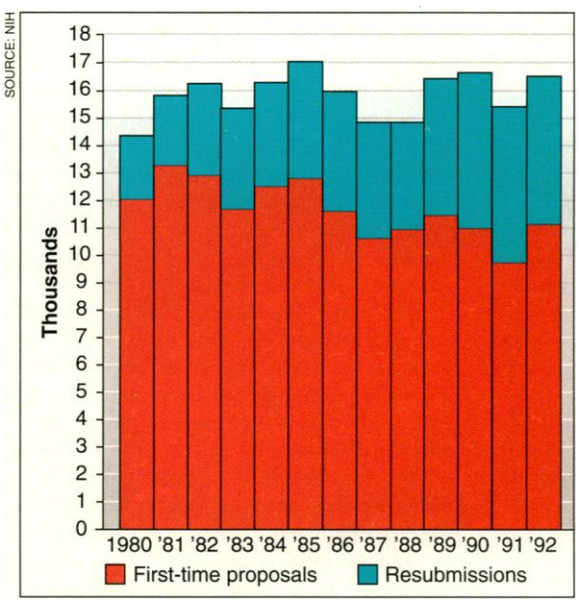
To streamline the process, Demsey has been instructing panel members in the four experimental study sections to regard any proposal that they would rank below the 50th percentile as being "not competitive." Before coming to the meeting, the reviewers draw up lists of NC proposals. It takes two to nominate a grant for this status, and just one objection would "bring it back to full review." Proposals remaining on the list at the first day of the meeting receive no further review. Other proposals are handled in the usual way.

The process is flexible, Demsey explains. "We're not saying, 'Hey, look, you have 100 applications; get 50 of them out of here.' We understand that you might have a good pool of applications. Maybe out of 100 only 35" would be rejected at the start. But he is pushing reviewers to trim the ranks early. Once the initial cut has been made, the NIH staffers who assist the panels will quickly send out notices of rejection. And instead of receiving a long summary sheet, reviewers will get a short statement that does little more than convey notes written by the primary and secondary reviewers.

This truncated process, Varmus says, is intended to give reviewers more time to spend on the top-ranked proposals and to make reviewing a more satisfying experience. Varmus is worried about a "self-perpetuating disenchantment" in which study group members get turned off by the tedium of the process, avoid service, and allow it to decline further. He would like to see "the highest quality people" serving on panels.

Varmus also argues that quick reviews will benefit those who are rejected by giving them a "very clear signal that this is not an application that can be moved into the fundable category simply by responding to a series of complaints" listed in a pink sheet. Institute staffers, he claims, will gladly help with advice on "retooling," taking a new tack, or considering another topic.

It's too early to tell how the experiment is working, but Demsey reports that the first panel to meet was able to isolate only 20% of the low-ranking grants. He notes, however, that this panel had already done some of its reviews and wasn't able to make the adjustment to triage quickly enough. The second panel did achieve its goal of 50%. Others are still in progress. Next, NIH will survey reviewers and reviewees for their reactions, and then, says Demsey, NIH will either drop the triage idea, apply it to all study sections, or mount another trial run.



Trying again. Resubmissions make up a growing percentage of the total number of RO1 grant proposals sent to NIH.

More changes ahead

If triage works, it may be just the first step in a tune-up of the peer-review system. NIH has already regrouped its 100 study sections into 19 broad "review groups"—a move that should make it easier to realign study groups as areas of science wax and wane. (NIH made the change in part to get around a White House mandate last spring ordering all agencies to lower costs by cutting the number of advisory panels by at least 33%.) Although the change won't reduce paperwork or save money, Baldwin says it will "give us more latitude" to shift members from one panel to another as needed to cope with technical questions. It may also make it easier to change assignments and possibly to alter what molecular biologist Keith Yamamoto of the University of California, San Francisco, refers to as "anachronisms" that have been "locked into the system."

Other changes will be considered at a 2-day "brainstorming" meeting to be held this spring, organized by Yamamoto with encouragement from Varmus. Yamamoto wants NIH staffers and reviewers in the areas of cell and molecular biology to consider whether study sections reflect the best science being done in laboratories and whether "we are overcovering or undercovering any areas," says Yamamoto. The goal, according to Yamamoto, is to help the NIH review system focus on the most exciting research.

Because the peer-review system is at the heart of NIH's operations, researchers are likely to be skittish about even the most modest changes. So NIH officials are already bracing for complaints from applicants who feel they haven't been given a fair shake under the new system. But Yamamoto thinks that triage is already an unqualified success in one small way: It shows researchers that someone is paying attention to their concerns, and that's bound to improve morale.

–Eliot Marshall