INFECTION AND IMMUNITY, Mar. 2009, p. 929–932 Vol. 77, No. 3
0019-9567/09/$08.00+0 doi:10.1128/IAI.01567-08

EDITORIAL

NIH Peer Review Reform—Change We Need, or Lipstick on a Pig?[▽]

"Individual investigators have sustained and driven our scientific enterprise for centuries, and the scientific enterprise, in turn, needs to sustain their productivity."
—Leo T. Furcht, Past President, FASEB (8)

A Crisis in NIH Funding

Find two scientists together, and chances are they are complaining about grants. The American research community is presently in its sixth year of a funding crisis exacerbated by an earlier period of growth that created new funding commitments and recruited additional scientists to the workforce. Resources in the system are insufficient to support current demands for research funds, and scientists are devoting unprecedented time and effort into competing over the dwindling funds available. Robert Siliciano, a prominent virologist, testified to a Congressional committee that 60% of his time is dedicated to seeking research funding (25). There are simply not enough resources for the number of scientists.

Would a successful business organize its research and development department so that employees spend more than half their hours writing detailed 5-year plans and then provide resources for only a 10th of them, leaving the rest to languish? Of course not. Yet, that is essentially the status of the nation's scientific enterprise in 2009. Over the past 5 years, the NIH budget has declined 13% after correcting for inflation (1). A greater emphasis on centrally defined research priorities in an era of declining budgets has had a particularly harsh impact on individual investigator-initiated research, the traditional engine of scientific progress. A precipitous 46% decline in R01 grants awarded between 2000 and 2007 underscores this trend (19).

We must face the fact that the ongoing funding imbalance is causing lasting harm to the nation's scientific enterprise, undermining both productivity and innovation. The crisis comes at a particularly inopportune time, as biomedical research will have a very important role to play in the world's economic recovery (3). For some scientists, their very jobs are at stake. This is because salary support for many American scientists is more dependent on grant revenues than in other countries (4). Additional casualties of the funding crisis are more difficult to measure but nevertheless real: deteriorating morale and a perceptible decline in scientific collegiality and cooperation. As David Sarnoff once observed, "Competition brings out the best in products and the worst in people" (26).

Peer Review Reform

In the midst of the funding crisis comes an NIH initiative to reform its peer review process. Called "Enhancing Peer Review," the program was initiated in July 2007, modified in response to feedback from numerous scientists, and is now scheduled to be phased in between January 2009 and January 2010. Reform of the current peer review process was motivated by a concern about the enormous administrative burden of the review process (5). Other justifications included complaints about review quality, very low funding rates among new investigators, and the declining NIH budget. The fact that none of these concerns is new (7, 27) suggests that solutions will not come easily.

The stated goals of the current NIH peer review reforms are to reduce the administrative burden associated with the grant process, enhance review quality, and increase support for new or early-stage investigators. The central elements of the program include (i) new funding targets for early-stage investigators (within 10 years of completing a terminal research degree or medical residency), (ii) shortened applications (a reduction from 25 to 12 pages), (iii) a new one- to nine-point scoring system with separate scores for individual criteria (impact, investigators, innovation, feasibility, and environment), (iv) limitation of grant applications to two submissions, and (v) incentives for long-term reviewers. Further details can be found at http://enhancing-peer-review.nih.gov/index.html.

Impact of the Changes

We commend the NIH for their efforts, as some of the changes should improve the existing system. For example, abbreviating the length of grant applications will reduce the workload for applicants and reviewers as well as lessen the emphasis on experimental minutiae. In fact, compelling arguments can be made for even shorter applications, particularly for established productive investigators (23), and many non-NIH grant applications are already less than 12 pages. The renewed emphasis on new and early-stage investigators is welcome news for scientists early in their careers, who have watched the average age of R01 recipients rise steadily. These individuals represent the future of science and warrant special consideration.

The benefits of other proposed changes are less clear. A nine-point scale will decompress the current scoring system. However, the scoring itself will remain a subjective process, and the new scale will not mitigate the problems inherent in selecting the most meritorious projects when resources are so limited (24). The elimination of third resubmissions of revised proposals, referred to as "two strikes and you're out" (12), is concerning, given the limitations of peer review (see below). If a proposal has been rejected twice, the new criteria suggest that it should be regarded as nonviable and never seen again. Yet science evolves, and new supporting data or context may cast previously rejected ideas in a more favorable light. It is particularly curious that the present success rate of third (A2) submissions is substantially higher than that of first or second (A1) submissions (12), yet these are the very applications that will be eliminated under the new system. One suspects that the new rule will drive applicants to be more devious in disguising and repackaging revised proposals, causing additional distractions for both applicants and reviewers. An even greater con-

▽ Published ahead of print on 21 January 2009.

cern is that some unsuccessful applicants will simply abandon science.

What Problems Have Not Been Fixed?

(i) A persistent imbalance between resources and applicants. Problems with peer review are more visible during times of scarce research funding. There is no escape from the relentless math of an unprecedented number of applications (now approaching 80,000 annually) chasing a shrinking number of grant dollars. If the amount of available funds does not increase dramatically, diversion of funds to early-stage investigators could conceivably make things worse by heightening competition among senior investigators and reducing available resources for scientists during what should otherwise be their period of peak productivity (6). In addition to their direct scientific contributions, senior investigators represent an invaluable source of wisdom, guidance, and inspiration for younger scientists. Furthermore, as new investigators become established (a transition that occurs rapidly in science), the shift in funding priorities that allowed their early success will soon disappear and leave them to compete with senior investigators for a shrinking pool of resources. This raises the uncomfortable question of whether we should be training more scientists when there is already a shortage of support for those already trained. Clearly, there is a need for recruiting new investigators to maintain the pool of scientists and to provide fresh outlooks on scientific problems. However, the strategy for ensuring the success of new investigators needs to be considered carefully. Short-term increases in support for trainees and new investigators will only exacerbate the current dearth of grant support and job opportunities as these trainees progress in their careers, unless the total level of support for research is dramatically increased.

(ii) The imperfect science of grant reviewing. For a system that determines the fate of scientific proposals, peer review is remarkably unscientific. Analyses have revealed that the NIH peer review system is statistically weak, imprecise, and prone to bias (11, 13). At its extremes, the error and variability in the review process become almost laughable. One of our colleagues recently witnessed an application to receive a perfect score of 1.0 when part of a program project application, but the identical application was unscored as a stand-alone R01. Almost no scientific investigation has been performed to examine the predictive accuracy of study section peer review. With more than a half-century of study section assessments on record, it would be interesting to know the frequency with which major scientific discoveries were recognized and anticipated by study sections. For example, what fraction of applications scored above or below the 10th percentile has been associated with major recognized scientific discoveries during the past 50 years? Similarly, what percentage of important scientific discoveries that were initially reviewed as proposals was rejected? Putting a stronger scientific foundation and accountability into peer review would enhance confidence in the system and facilitate evidence-driven improvements (10).

(iii) Disincentives for novelty. Reviewer biases favor topics well understood and appreciated by the study section and disfavor less conventional ideas or understudied topics. This leads to greater homogeneity in science. Applicants learn to write conservative proposals to avoid creating targets for reviewers. Nobel laureate Roger Kornberg recently observed, "In the present climate especially, the funding decisions are ultraconservative. If the work that you propose to do isn't virtually certain of success, then it won't be funded. And of course, the kind of work that we would most like to see take place, which is groundbreaking and innovative, lies at the other extreme" (16). The playful curiosity (15) and open-ended thinking that characterize the best science have become increasingly rare as scientists are driven by funding anxiety to propose safe, conservative, short-term projects. (Of course, conservatism is no guarantee of success in the peer review arena; reviewers may then characterize such a proposal as "unambitious" instead of "risky.") There is a basic distinction between the peer review of manuscripts, in which the work has been done, and the peer review of grants, in which the work is being proposed. The difference is as fundamental as reviewing a movie versus forecasting the weather. Scientists have a limited ability to predict a priori which experimental paths will be most fruitful. Therefore, grant reviews at best involve probability and uncertainty. If only projects that are certain to succeed are given support, then it becomes a virtual certainty that many worthwhile projects will fail to receive support. One could argue that most projects in which one can predict success with certainty are, by their very nature, unlikely to be the type of highly innovative science that leads to major breakthroughs, since certainty in prediction means that one is operating within existing boundaries of knowledge. Study sections normally base their decisions on consensus and thus function within the sphere of what Kuhn called "normal science" (14), which discourages innovative deviations from established paradigms. The drive toward conformity ignores the essential role of serendipity in science (21)—unexpected results are often the most exciting and fruitful ones. The pressure to eschew innovation will not be alleviated by a few Transformative, Eureka, New Innovator, or Pioneer awards. What is actually needed is not separate funding earmarks for high-risk research but rather sufficient breathing room to accommodate a few high-risk ideas as appropriate components of any research project.

(iv) Reviewer expertise. Another important difference between peer review of manuscripts and grants is the expertise of the reviewers. Journal editors can consult any scientist in the world about a manuscript, but an NIH grant review is essentially limited to the expertise of the 30 scientists in the room, who may each be outstanding in their fields but cannot hope to encompass the entire range of applications considered. Some of the most active scientists can rarely participate in study sections because they are submitting their own applications virtually every grant cycle. Grants are typically reviewed by at least some individuals with substantially less expertise in a given subject than the applicant, and such reviews tend to focus on critiquing "grantsmanship" instead of the science itself. The grant application, originally intended only as a tool to facilitate the distribution of research funds, has thereby acquired an undeserved status as an object of obsession. When there is insufficient funding available for the number of quality proposals submitted, funding decisions become increasingly capricious. Decisions to deny funding must be justified, so critiques become packed with unhelpful and generic demands for more preliminary data, expected results, anticipated pitfalls, alternative approaches, and timetables, etc. Such criticisms may not be the true reason for denying funding, but they force applicants and reviewers to prolong the dance of revision and review. There has always been a disconnection between the science that is described in grants and that which is performed in reality, but today the chasm is absurdly wide.

(v) Large disparities in funding among investigators. What is the optimal amount of funding for a single investigator? This is a fundamental question for which there is no answer. Clearly, research is very difficult if not impossible in the absence of funding. At the other extreme, one can imagine that efficiency declines as groups become very large and the efforts of a single investigator become diffused. This problem might be studied using available economic tools and the results used to optimize the allocation of funds, but this issue, like peer review, should be the subject of rigorous scientific analysis. Study sections should perhaps consider not only the productivity of an applicant but the productivity relative to dollars awarded, with the caveat that the number of publications alone should not be the sole parameter considered (18). In fact, it might be argued that new tools should be developed to help study sections to objectively gauge the productivity of investigators and the potential impact of proposed research. While the awarding of grants on the basis of productivity seems reasonable in principle, it should be recognized that this creates a positive-feedback loop that can aggravate the already highly inhomogeneous distribution of research funding. *Nature* reported that 200 scientists had six or more grants from NIH in 2007 and that just 19 researchers accounted for 165 research grants totaling $160,000,000 (9). Expensive "big science" technology-driven projects and large clinical trials are important factors contributing to this trend. The success of a fortunate few prominent scientists in the current funding environment cannot conceal the funding woes of the majority in a mirroring of the widened gap between the haves and have-nots in the economy at large. The NIH peer review reform initiative considered a variety of suggestions to address the maldistribution of grants, such as requiring principal investigators to spend at least 20% effort on each grant, but failed to arrive at a consensus.

(vi) Lengthy review process. Presently, investigators must wait 4 months to receive a critique of their application. Such delays are likely to represent purposeful inefficiency, so that revised applications can be submitted only every 8 months, or two grant cycles. This will not change under enhanced peer review.

(vii) Administrative burden. In recent years, the scarcity of research funding has been paralleled by burgeoning administrative burdens. It was estimated that 42% of the total faculty research time is consumed by administrative activities required for compliance with research grants, such as progress reports, accounting, animal protocols, human protocols, and select-agent regulations, etc. (17). Although scientists owe society full accountability on these important issues, the enormous energies spent on paperwork introduce friction into the scientific process that contributes to inefficiency and lower productivity.

(viii) Wasted time and human power. Withholding funding will not make investigators more productive, nor will it be easy for unfunded investigators to come up with the additional preliminary data needed to make an amended proposal more persuasive. Once an investigator struggles through a few years of inadequate funding, it becomes less and less likely that their research program will ever return to a competitive level. The system as it exists is exceedingly wasteful in that we are not allowing many highly trained and competent scientists to work to their full potential, and some are unable to work at all. Although the administrative process of peer review will be somewhat decompressed, the massive time commitment to the grant funding process demanded of applicants will continue unabated under the new NIH reforms. The increased diversion of scientists' efforts into grant procurement rather than research, along with the temptation to pursue funds earmarked by Congress and the NIH for specific programs rather than continuing the natural lines of investigation initiated by the individual investigator, results in reduced productivity. In response to a news feature in *Science* entitled "U.S. output flattens and NSF wonders why" (20), John Moore replied, "The number of papers that are written is diminishing because scientists are able to spend less time writing papers! Instead, we spend ever-more time on. . .writing, rewriting, and re-rewriting grant applications as the NIH's pay line drops to catastrophically low levels" (22). The funding shortage will continue to undermine recruitment of the best and brightest young minds to research. The costs will be difficult to measure—it is hard to quantify the discoveries not made and the great scientists who never were.

Conclusions

Peer review remains a cornerstone of science and a true leveler of the playing field on which applicants compete. We recognize that it is imperative for the quality and fairness of peer review to be optimized during times of resource scarcity. However, the problem of inadequate resources cannot be compensated for by changes in the mechanism by which available funds are allocated. In fact, peer review cannot work effectively when funding is so limited (24). Thus, the NIH Enhanced Peer Review initiative is an essentially cosmetic reform that might reduce the administrative burden associated with grant applications but will do little to improve the current plight of most individual investigators or to ensure that meritorious proposals receive support. The reinvigoration of American science will require much bolder action that considers the needs of the entire research workforce and not only the scientific elite. The foremost priority is to restore a balance between funding and applicants that returns pay lines to reasonable levels. As the outgoing NIH Director Elias Zerhouni has acknowledged, "Peer review doesn't need to be as stringently quality-focused when there is a lot of money" (2). The current system for awarding research grants is error-prone and wasteful and discourages innovation. America must put its scientists back to work making discoveries instead of endlessly writing grants. Now *that* would be a change we can believe in.

We are grateful to Liise-anne Pirofski for her insightful comments and suggestions.

REFERENCES

1. **AAAS.** 2008. AAAS analysis of R&D in the FY 2009 budget. AAAS, Washington, DC.
2. **Anonymous.** 2008. A metareview at the NIH. Nat. Med. **14:**351.
3. **Anonymous.** 2008. Research and recovery. Nat. Med. **14:**1129.
4. **Benderly, B. L.** 7 May 2008. Taken for granted: lost in space. Sci. Careers. doi:10.1126/science.caredit.a0800064.
5. **Bonetta, L.** 2008. Enhancing NIH grant peer review: a broader perspective. Cell **135:**201–204.

6. **Brumfiel, G.** 2008. Older scientists publish more papers. Nature **455:**1161.
7. **Finn, R.** 1995. NIH study section members acknowledge major flaws in the reviewing system. Scientist **9:**7.
8. **Furcht, L. T.** 2007. NIH funding: what does the future look like? Science **316:**198–200.
9. **Hand, E.** 2008. 222 NIH grants: 22 researchers. Nature **452:**258–259.
10. **Hannun, Y. A.** 2008. NIH: grants revamp needs grounding in evidence. Nature **452:**811.
11. **Johnson, V. E.** 2008. Statistical analysis of the National Institutes of Health peer review system. Proc. Natl. Acad. Sci. USA **105:**11076–11080.
12. **Kaiser, J.** 2008. National Institutes of Health. Two strikes and you're out, grant applicants learn. Science **322:**358.
13. **Kaplan, D., N. Lacetera, and C. Kaplan.** 2008. Sample size and precision in NIH peer review. PLoS ONE **3:**e2761.
14. **Kuhn, T. S.** 1962. The structure of scientific revolutions. University of Chicago, Chicago, IL.
15. **Larsson, U.** 2005. Cultures of creativity: birth of a 21st century museum. Science History Publications, Sagamore Beach, MA.
16. **Lee, C.** 28 May 2007. Slump in NIH funding is taking toll on research, p. A06. Washington Post, Washington, DC.
17. **Leshner, A. I.** 2008. Reduce administrative burden. Science **322:**1609.
18. **Lewison, G., J. Anderson, and J. Jack.** 1995. Assessing track records. Nature **377:**671.
19. **Mandel, H. G., and E. S. Vesell.** 2008. Declines in NIH R01 research grant funding. Science **322:**189.
20. **Mervis, J.** 2007. Scientific publishing. U.S. output flattens, and NSF wonders why. Science **317:**582.
21. **Meyers, M. A.** 1995. Science, creativity and serendipity. Am. J. Roentgenol. **165:**755–764.
22. **Moore, J. P.** 2007. Speaking out about U.S. science output. Science **318:**913.
23. **Pagano, M.** 2006. American idol and NIH grant review. Cell **126:**637–638.
24. **Petsko, G. A.** 2006. The system is broken. Genome Biol. **7:**105.
25. **Siliciano, R.** 19 March 2007. Testimony before the Senate Committee on Appropriations: Subcommittee on Labor, Health and Human Services, Education, and Related Agencies. U.S. Senate Committee on Appropriations, Washington, DC.
26. **Van Wart, M.** 2005. Dynamics of leadership in public service: theory and practice. M. E. Sharpe, Armonk, NY.
27. **Wessely, S.** 1998. Peer review of grant applications: what do we know? Lancet **352:**301–305.

Ferric C. Fang
Editor in Chief, Infection and Immunity

Arturo Casadevall
Editor, Infection and Immunity

The views expressed in this Editorial do not necessarily reflect the views of the journal or of ASM.