Scientific risk-taking and grant funding: A "risky research" agenda for NIH

Pierre Azoulay,[1] MIT Sloan School of Management and NBER

May 2023

***Note:** This piece was prepared for the Building a Better NIH project, a partnership between the Brookings Institution, the Institute for Progress, and the Good Science Project. For more information about the project, please visit the [Building a Better NIH project page](#).*

Few topics elicit as much consensus as the idea that peer review punishes risk taking. Reports abound, ranging from the popular press to U.S. Congressional testimony, from tweets to the pages of *Nature*, and are authored by NIH-funded investigators and leading science policy makers (Kent, 2018; Kolata, 2009; Kornberg, 2007; Nielsen, 2022; Sutter, 2022; Woolston, 2014). So pervasive and prominent are these concerns that it leads scientists to write articles satirizing how risk-adverse grant funders are and muse if even former Nobel Prize winners could receive funding for their seminal discoveries in the modern funding climate (Fields, 2014; Petsko, 2012).

Yet, despite the abundance of anecdotes, there is surprisingly little evidence to support claims that peer review punishes risk taking. Azoulay et al. (2011) compare two grant funding mechanisms—the Howard Hughes Medical Investigator competitions and the R01 grants awarded by NIH—and show the former encourages more risk taking. Beyond just theoretical interest, grant funders potentially punishing risk taking has very direct implications for what research is undertaken. In a recent survey of investigators supported by Fast Grants, 78% stated they would make significant changes to their research program if they were not constrained by the demands of grant funding agencies, including pursuing more ambitious research programs, pivoting to new topics, and testing hypothesis others see as unlikely to succeed (Collison et al., 2021).

More generally, the choice of what scientists choose to study is fundamental to determining the rate and direction of innovation. Not all proposed scientific projects, however, have an equal likelihood of extending the scientific frontier. High-risk, high-reward projects may be more likely to lead to breakthrough innovation (Foster et al., 2015; Uzzi et al., 2013; Veugelers and Wang, 2019; Wang et al., 2017). The costs of undersupplying risky research may be especially salient for early-stage fundamental research, because it is often the type of research providing the broadest shoulders for follow-on innovators to stand on (Aghion et al., 2008; Azoulay et al., 2019; Fleming et al., 2019).

---

[1] The author did not receive financial support from any firm or person for this article or from any firm or person with a financial or political interest in this article. The author is not currently an officer, director, or board member of any organization with a financial or political interest in this article.

This short document has two objectives. First, it reports on new results on scientific risk-taking in the context of competitive renewal decisions for R01 grants at NIH. I believe this evidence is useful both because most research on scientific risk-taking has focused on other settings, and because most research on scientific risk-taking typically studies novelty rather than risk-taking *per se.*

Second, I propose a three-pronged research agenda for NIH as it considers how to encourage risk-taking. First, administrative data (which contains information about applications, not just funded grants) could be used to replicate the results, in particular to study whether the “risk-taking penalty” is present at the stage of a grant’s initial cycle. Second, the NIH could help validate the metrics of scientific risk-taking proposed below. Third, I propose four ideas for randomized controlled interventions that could produce a solid empirical foundation for efforts at reforming the NIH peer review process to alleviate (or even reverse) the risk-taking penalty.

1. New evidence on risk taking and NIH grant renewal

My collaborators and I study how risk taking is penalized or rewarded in 103,164 NIH-funded R01-equivalent grant cycles, 1980-2015 (Greenblatt et al., 2022). After controlling for a robust set of covariates for investigator demographics and talent, institution quality, and grant characteristics, we observe the relationship between risk taking in publications that acknowledge funding from the grant and whether the grant was competitively renewed. Risk taking, much like creativity, is a nebulous concept. There is no consensus on what constitutes risk in science nor how it should be measured (Althaus, 2005; Franzoni and Stephan, 2021). As there are different aspects of risk taking, each of which may have distinct dynamics, it is unlikely a single measure of risk taking can fully represent the phenomenon. We develop four separate measures of risk taking, each of which captures a different way investigators may take risks (or how NIH study sections may evaluate risk). First, risk taking may be associated with more extreme tail outcomes, with greater variance in outcomes and both more exceptionally good and exceptionally bad results (Azoulay et al., 2011). Second, risky research may be more disruptive, seeking to overturn the status quo, rather than consolidative, reinforcing earlier results (Funk and Owen-Smith, 2017). Third, researchers may pivot from what they have done in the past to follow new scientific opportunities, capturing the idea that risk is not only inherent to the project itself but depends also on who is undertaking the project. Finally, researchers may stand out from the crowd, selecting projects that are intellectually distant from those of other investigators (Uzzi et al., 2013). Across each of these measures, we take a non-parametric approach in exploring effects across the full distribution of different levels of risk taking. This lets the data speak for itself in whether any observed effect is driven primarily by differences in conventional levels of risk taking, or alternatively, by extreme tail risk taking – arguably those projects with the most potential to lead to breakthrough innovation.

Across each of these measures, we find that risk taking is penalized. When comparing grant cycles in the top and bottom decile of risk taking, grants with greater risk taking have a 9.5% lower renewal rate (20.5% decline) when measuring risk taking using extreme tail outcomes, an 11.1% lower renewal rate (24.4% decline) when measuring risk taking by its disruptiveness, a 7.7% lower renewal rate (16.9% decline) when measuring risk taking by an investigator pivoting from her prior research, and a 6.1% lower renewal rate (12.4% decline) when measuring risk taking by standing out from what other investigators are studying.

The obdurate nature and magnitude of this risk penalty is shocking in light of the often-expressed desire by NIH officials to encourage and support “high-risk, high-reward” projects. Of course, two caveats are in order when interpreting these results. First, these four metrics have not been validated using external sources of data. Second, the data at our disposal is limited to information about funded grants. In particular, this implies that we cannot distinguish grants whose principal investigators unsuccessfully applied for a renewal from grants that simply lapsed because the scientists shifted their interest or left academic science. Our team would love the opportunity to make progress on both fronts, as explained below. The rest of this document will proceed on the assumption that the risk penalty we estimated corresponds to a robust, pervasive, and salient phenomenon within the NIH peer review system.

2. A “risky research” agenda

I propose a three-pronged agenda to advance the state of knowledge in this area: (1) a retrospective analysis using NIH administrative data; (2) the validation of risk-taking metrics using external sources; and (3) randomized, controlled interventions.

2.1 Observational data

The results mentioned above demonstrate that insights can be extracted from publicly available data. However, access to administrative data—especially information about unfunded applications—would provide a firmer foundation for empirical work in this area. First, one could examine whether the results extend to grants in their initial cycle. Second, one could distinguish between grants that the NIH declined to renew and those for which the grantee failed to submit a competitive renewal application.

2.2 Metrics

The research mentioned above proposes four metrics of risk-taking that draw on bibliometric data (publications and citations) on the one hand, and the structure of the review process on the other hand (e.g., other grants funded by the same study section and institute). While these metrics have a certain amount of face validity, they need much more thorough characterization and validation. How do they correlate with expert judgement of risk taking? How do they correlate with reviewers scores in the peer review process? With access to full-text information for the grant application, is it possible to use modern word embedding techniques (such as SciBERT) to score applications for risk taking?

2.3 Randomized controlled interventions

I sketch below four ideas for interventions.

- Experiment #1. Randomly select a subset of R01 grantees and announce to them that if/when they choose to apply for a competitive renewal, their application should not contain any preliminary data. The hypothesis to be tested is that this early knowledge would allow the new investigator to choose a bolder agenda, relative to a situation where she will face the burden of early proof.
- Experiment #2. Randomly select a subset of R01 grantees and announce to them (at the time they receive notice that their project has been funded) that the length

of their grant is going to be extended by n years, where $1 \leq n \leq 5$. The treated investigators would be secure in the knowledge that they face a longer time horizon than expected to plan their investigations. Nothing else would change (Epstein, 2011).[2]

- Experiment #3. For a random subset of special emphasis panels, implement a version of quadratic voting so that the intensity of evaluator sentiment, as well as the variance in sentiment, can be taken into account to generate a proposal score. Do the applications selected for funding under quadratic voting take on more scientific risk than those selected for funding under the traditional system?
- Experiment #4. For a random subset of R01 grantees (in the first year of the grant cycle), announce that they will not have to submit yearly progress reports, and in fact are encouraged to follow the logic of their investigations, wherever it leads them. Do the treated applicants take on more risk? Are they more or less likely to submit a competitive renewal?

When assessing risk and reward for each of these interventions, I can imagine having three dimensions in mind: (i) cost, (ii) time to generate results, and (iii) organizational overhead needed.

	Cost	**Time**	**Organizational overhead**
Experiment #1	Low	One cycle	Significant
Experiment #2	High	One cycle	Low
Experiment #3	Low	Immediate	Intermediate
Experiment #4	Low	One cycle	Low-ish

As can be seen, these experiments vary in their profile of costs and benefits to NIH. To various extents, however, they can be conducted in the regular course of peer review business. The biggest challenge is therefore likely to be the agency's own appetite for risk.

[2] This may strike the reader as being very reminiscent of the MERIT R37 program. This is correct, but (1) to our knowledge, the R37 mechanism was never the target of a rigorous evaluation; and (2) the R37 was only awarded to applicants who score in the top 2 percentiles of the score distribution. But one may wonder if the investigators very skilled at impressing study section members are necessarily those whose taste for exploration are constrained by the current system.

References

Aghion, P., Dewatripont, M., & Stein, J.C. (2008). Academic freedom, private-sector focus, and the process of innovation. *RAND Journal of Economics*, 39(3):617-635. https://scholar.harvard.edu/files/stein/files/academic-freedom-rand-2008-final_0.pdf

Althaus, C. E. (2005). A disciplinary perspective on the epistemological status of risk. *Risk Analysis*,25(3):567-588. https://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2005.00625.x

Azoulay, P., Graff Zivin, J.S., & Manso, G. (2011). Incentives and creativity: evidence from the academic life sciences. *RAND Journal of Economics*, 42(3):527-554. https://onlinelibrary.wiley.com/doi/10.1111/j.1756-2171.2011.00140.x

Azoulay, P., Graff Zivin, J.S., Li, D., & Bhaven Sampat. (2019). Public R&D and Private-Sector Patenting: Evidence from NIH Funding Rules. *The Review of Economic Studies,* 86(1):117-152. https://academic.oup.com/restud/article/86/1/117/5038510?login=true

Collison, P., Cowen, T., & Hsu, P. 2021. *What We Learned Doing Fast Grants*. Future. https://future.com/what-we-learned-doing-fast-grants/

Epstein, J. A. (2011). Enhancing discovery and saving money with MERIT. *Journal of Clinical Investigation*, *121*(4), 1226–1226. https://doi.org/10.1172/JCI57708

Fields, S. (2014). Would Fred Sanger get funded today? *Genetics*, 197(2):435-439. https://academic.oup.com/genetics/article/197/2/435/6074290

Fleming, L., Greene, H., Li, G., Marx, M., & Yao, D. (2019). Government-funded research increasingly fuels innovation. *Science*, 364(6446):1139-1141. https://www.science.org/doi/10.1126/science.aaw2373

Foster, J.G., Rzhetsky, A., & Evans, J.A. (2015). Tradition and Innovation in Scientists' Research Strategies. *American Sociological Review,* 80(5):875-908. https://journals.sagepub.com/doi/10.1177/0003122415601618

Franzoni, C., & Stephan, P. 2021. *Uncertainty and Risk-Taking in Science: Meaning, Measurement and Management*. National Bureau of Economic Research Working Paper #28562. https://www.nber.org/papers/w28562

Funk, R.J., & Owen-Smith, J. (2017). A Dynamic Network Measure of Technological Change. *Management Science*, 63(3):791-817. https://doi.org/10.1287/mnsc.2015.2366

Greenblatt, W., Maity, S.K., Roger P. Levy, and Pierre Azoulay. 2022. "Does Grant Peer Review Penalize Scientific Risk Taking? Evidence from the NIH." Working Paper, MIT.

Kent, D. 2018. *Scientists can't take risks until they are in their 40s*. University Affairs. https://www.universityaffairs.ca/opinion/the-black-hole/scientists-cant-take-risks-until-they-are-in-their-40s/

Kolata, G. (2009). *Grant System Leads Cancer Researchers to Play it Safe.* New York Times. https://www.nytimes.com/2009/06/28/health/research/28cancer.html

Kornberg, R.D. (2007). *2006 Nobel Laureates*. U.S. Senate Committee on Commerce, Science, and Transportation. https://www.commerce.senate.gov/2007/5/2006-nobel-laureates

Nielsen, M. (@michael_nielsen). (2022). *(A brief afternoon rant.)* Twitter. https://twitter.com/michael_nielsen/status/1559318292120027137

Petsko, G.A. (2012). Goodbye, Columbus. *Genome Biology*, 13(5):155. https://genomebiology.biomedcentral.com/articles/10.1186/gb-2012-13-5-155

Sutter, P. M. (2022). *Risk Aversion Is Ruining Science*. Undark. https://undark.org/2022/04/27/risk-aversion-is-ruining-science/

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical Combinations and Scientific Impact. *Science*, 342(6157):468-472. https://www.science.org/doi/10.1126/science.1240474

Veugelers, R. & Wang, J. 2019. Scientific novelty and technological impact. *Research Policy*, 48:1362-1372. https://www.sciencedirect.com/science/article/abs/pii/S0048733319300459?via%3Dihub

Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy,* 46:1416-1436. https://www.sciencedirect.com/science/article/abs/pii/S0048733317301038?via%3Dihub

Woolston, C. (2014). Sanger's legacy stirs up digital chatter. *Nature*, 511(9). https://www.nature.com/articles/511009f