Michael Rosenblatt is Executive Vice President and Chief Medical Officer at Merck & Co., Kenilworth, NJ 07033, USA. Email: michael.rosenblatt@merck.com

Citation:
M. Rosenblatt, An incentive-based approach for improving data reproducibility. *Sci. Transl. Med.* **8**, 336ed5 (2016).

10.1126/scitranslmed.aaf5003

RESEARCH TRANSLATION

An incentive-based approach for improving data reproducibility

OVER THE PAST SEVERAL YEARS, A SERIES OF PUBLICATIONS AND POLICY STATEMENTS have generated increasing awareness in the scientific community of the scale and implications of the problem of irreproducible data—or at least lack of robust results—particularly in the realm of basic and translational research (*1*–*3*).

Recent studies have shown that the key findings in 50% or more of published reports in certain fields cannot be reproduced (*1*, *2*, *4*). As the public, government, and private funders of research comprehend the extent of the problem, trust in the scientific enterprise erodes, and confidence in the ability of the scientific community to address this problem wanes. In addition, there is considerable potential for reputational damage to scientists, universities, and entire fields (for example, cancer biology, genomics, and psychology).

The irreproducible data problem has a major impact on one subset of biomedical science: translational research, which sits poised at the interface of academic science and industry research and development (R&D). Research conducted in academia provides new insights into human pathophysiological mechanisms and identifies new targets for drug discovery. The public looks to collaborations between academia and industry, and then ultimately industry, to translate this research into innovations that improve health, namely new medicines and vaccines, as rapidly as possible. This is a vexing issue that threatens the entire biomedical research enterprise with increasing public scrutiny, including from the U.S. Congress.

The diversion of much of a critical component of the translational effort into avenues that have little chance of success wastes not only a portion of the public's investment in academic research and industry's subsequent investment but also valuable time. There is an opportunity cost; not working on potentially successful projects leads to delays in delivering health benefits to the population. Like all scientific endeavor, key early experiments on a disease or biological process build the foundation for further investigation. For drug discovery and development, the time, effort, and expenditure increase as a project moves from laboratory findings to animal efficacy and safety studies and then, ultimately, to clinical trials. The consequences of building on a shaky platform are so great that most biopharmaceutical companies and venture capitalists expend considerable resources in attempts to replicate initial data before committing to further research. On average, it takes approximately two to six scientific personnel 1 to 2 years of work in an industry laboratory to try to reproduce the original investigations; cost estimates average $500,000 to $2 million (*4*).

I have spent my career in both academia and industry. On the basis of this perspective, I propose a potential approach to diminish the data irreproducibility problem at the academia-industry interface. The problem of data that cannot be replicated has many potential origins: pressures to publish papers or secure grants, criteria for career advancement, deficiencies in training, and nonrigorous reviews and journal practices, with fraud an infrequent cause (*1*–*3*, *5*). So it is not surprising that an array of remedies have been proposed to address each of the problems; these include more rigorous methods, standards, statistical analysis, training, and changes in evaluation for promotion and tenure (*1*–*3*, *5*). These approaches prescribe specific treatments and, in general, are recommendations without an "effector" mechanism—they are "process-oriented." Some have the potential to be implemented along with penalties or financial disincentives. But even taken collectively and given that these approaches will involve cultural change, it is doubtful that they will be timely or sufficient. Some have the disadvantage of new regulation, which might stifle creativity and other solutions.

I propose an approach to diminish irreproducible data as a roadblock to effective academia-industry collaborations. The proposal is incentive-based and nonprescriptive (that is, without set guidelines for implementation). It provides financial resources to seek replication of results, with the freedom for academic institutions to select solutions best suited to their context.

Here is the essence of the proposal: What if universities stand behind the research data that lead to collaborative agreements with industry, and what if industry provides a financial incentive for data that can be replicated? Currently, industry expends and universities collect funding, even when the original data cannot be reproduced. In the instance of failure, collaborations dissolve, with resulting opportunity loss for both academia and industry. But what if universities offered some form of full or partial money-back guarantee? With such assurance, companies could proceed with a project more rapidly and more frequently. They would also be likely to pay a premium over current rates for data backed by such assurance over "nonguaranteed" data, even from the same university. This approach places the incentive squarely with the investigator (including his or her laboratory) and the institution—precisely the leverage points for change. The premium would provide universities with the financial wherewithal to cover the cost of affirming their data if they choose to replicate it before entering into a collaboration.

It is the concept of assurance with financial backing and further monetary incentive for the delivery of reproducible data that I am proposing. There are many possible ways to implement such an approach. Because the proposal does not prescribe specific solutions, each institution would be free to select its own means, on the basis of the putative sources of the problem at that institution and its favored remedies. For example, if an institution has high confidence in the data, they might not replicate before partnering. Alternatively, replication could be funded at the originator laboratory, at another laboratory at the same institution, at a laboratory at another university, in collaboration with the industry partner, or at a contract laboratory. Independent commercial laboratories might arise for the express purpose of replicating data.

Some collateral benefits might accrue: Academic investigators might be more careful to bring forward only data in which they had a high degree of confidence (in order to avoid retraction and financial loss for the university). Similarly, "early adopter" research institutions might become preferred partners for industry, stimulating other universities to follow. Last, the results can be measured; reproducibility going forward can be compared with the historical record.

This proposal will likely raise issues and objections and have unintended consequences. Although academia-industry collaborations represent only a minor portion of the biomedical research enterprise, they occupy a crucially important and visible position along the translational pathway toward new therapies. Hence, change in this arena might foster cultural changes more broadly. It is my hope that this proposal will encourage a pilot trial by a partnering university and biopharmaceutical company. It is not meant as a panacea for this complex problem. However, if both parties ultimately spend more time and resources, both internally and collaboratively, on projects that have a higher probability of success, then patients and society would be beneficiaries.

I welcome reactions to this proposal.

–Michael Rosenblatt

REFERENCES AND NOTES

1. Data replication and reproducibility. *Science* Special Issue (2 December 2011). www.sciencemag.org/site/special/data-rep.
2. Challenges in irreproducible research. *Nature* Special Issue (8 October 2015). www.nature.com/news/reproducibility-1.17552.
3. F. S. Collins, L. A. Tabak, Policy: NIH plans to enhance reproducibility. *Nature* **505**, 612–613 (2014).
4. L. P. Freedman, I. M. Cockburn, T. S. Simcoe, The economics of reproducibility in preclinical research. *PLOS Biol.* **13**, e1002165 (2015).
5. C. G. Begley, J. P. Ioannidis, Reproducibility in science: Improving the standard for basic and preclinical research. *Circ. Res.* **116**, 116–126 (2015).

Acknowledgments: W. Chin from PhRMA provided helpful comments on a draft of the manuscript. C. Lines from Merck & Co. assisted with editing the manuscript. **Competing interests:** M.R. is an employee of and owns stock and stock options in Merck & Co. The opinions expressed in this article are those of the author and do not necessarily correspond to the views of the author's employer.

Science Translational Medicine

An incentive-based approach for improving data reproducibility

Michael Rosenblatt

Sci Transl Med **8**, 336ed5336ed5.
DOI: 10.1126/scitranslmed.aaf5003

Science Translational Medicine (ISSN 1946-6242) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Translational Medicine* is a registered trademark of AAAS.