

## Making the scientific literature fail-safe

Arturo Casadevall, Ferric C. Fang

*J Clin Invest.* 2018;128(10):4243-4244. <https://doi.org/10.1172/JCI123884>.

### Viewpoint

If a critical system fails, another should always be there to do its work. —John Downer (1)  
Critical systems are designed to be fail-safe. This does not mean that failures cannot occur, but rather that redundant and compensatory mechanisms are engineered into the system to detect and mitigate failures when they occur. The scientific literature is the critical system by which scientific findings are communicated and archived for subsequent reference and analysis. Hence, the reliability of the scientific literature is of the utmost importance to society. However, in recent years, rising numbers of retracted articles, reproducibility problems, and inappropriately duplicated images have increased concern that the scientific literature is unreliable (2–5). Contributing factors may include competition, sloppiness, prioritization of impact over rigor, poor experimental design, inappropriate statistical analysis, and lax ethical standards (6, 7). Although the number of questionable publications represents a very small percentage of the total literature, even a few problematic publications can reduce the credibility of science. Hence, it is important to redouble efforts to improve the reliability of scientific publications. We suggest a seven-point approach to reengineering the scientific literature so that it is better able to prevent and correct its failures. i. Improving graduate and postgraduate training. Training is the foundation of all scientific endeavors. Contemporary graduate scientific training is designed to prepare trainees to [...]

**Find the latest version:**

<http://jci.me/123884/pdf>



# Making the scientific literature fail-safe

Arturo Casadevall<sup>1</sup> and Ferric C. Fang<sup>2</sup>

<sup>1</sup>Department of Molecular Microbiology and Immunology, Johns Hopkins School of Public Health, Baltimore, Maryland, USA. <sup>2</sup>Departments of Laboratory Medicine and Microbiology, University of Washington, Seattle, Washington, USA.

***If a critical system fails, another should always be there to do its work.***

—John Downer (1)

Critical systems are designed to be fail-safe. This does not mean that failures cannot occur, but rather that redundant and compensatory mechanisms are engineered into the system to detect and mitigate failures when they occur. The scientific literature is the critical system by which scientific findings are communicated and archived for subsequent reference and analysis. Hence, the reliability of the scientific literature is of the utmost importance to society. However, in recent years, rising numbers of retracted articles, reproducibility problems, and inappropriately duplicated images have increased concern that the scientific literature is unreliable (2–5). Contributing factors may include competition, sloppiness, prioritization of impact over rigor, poor experimental design, inappropriate statistical analysis, and lax ethical standards (6, 7). Although the number of questionable publications represents a very small percentage of the total literature, even a few problematic publications can reduce the credibility of science. Hence, it is important to redouble efforts to improve the reliability of scientific publications. We suggest a seven-point approach to reengineering the scientific literature so that it is better able to prevent and correct its failures.

*i. Improving graduate and postgraduate training.* Training is the foundation of all scientific endeavors. Contemporary graduate scientific training is designed to prepare trainees to perform deep investigation into a highly specialized area (8) but does not necessarily provide students with a broad scientific background. Students are taught in a guild-like environment by a mentor who may or may not

have been trained in good scientific practices. Consequently, there is no guarantee that programs are consistently producing scientists who are adequately prepared to do good science. Improving postgraduate training to ensure that trainees are well versed in scientific rigor, statistical analysis, experimental design, and ethics can improve the quality of the scientific literature by improving the quality of the research itself.

*ii. Reducing errors in manuscript preparation.* Roughly one of every 25 articles in the biomedical literature contains an inappropriately duplicated image (4). The majority of inappropriate image duplication results from simple errors in figure assembly (9). However, a minority of these represents intentional efforts to mislead the reader, which constitutes scientific misconduct. Involving multiple individuals in figure preparation prior to manuscript submission may reduce the likelihood of error and also discourage intentional deception.

*iii. Presubmission criticism.* Although peer review is intended to detect and correct errors prior to publication, the process involves only a small number of reviewers and is well known to be imperfect (10). Critical input from a broader range of colleagues may lead to identification of weaknesses in a manuscript and allow authors to improve the quality of their published work. Presubmission criticism may be informally obtained by asking others to read a manuscript before submission or by posting the manuscript on a preprint server and alerting colleagues in the field that the data are available in prepublication form. Both authors of this commentary have received presubmission criticism of manuscripts posted as preprints that led to improvements. A more longstanding mechanism for obtaining presubmission

criticism is to present unpublished data at meetings and seminars.

*iv. Robust review and editorial procedures.* After a manuscript is submitted for publication, the peer review and editorial processes are major checkpoints for quality improvement. Reviewers can identify errors, and training may improve their ability to detect problematic data. Journals can use software to identify plagiarism, image manipulation, or data anomalies (11–13). Dedicated statistical editors and reviewers can help to ensure that complex data sets are appropriately analyzed. The *JCI* requires authors to submit copies of the original Western blot images used for figure construction. Having access to the original data may discourage manipulation and can be used to address questions that may arise with regard to figure presentation.

*v. Postpublication criticism.* The development of sites such as *PubPeer* allows readers to anonymously post critical feedback after a manuscript has been published (14). Such comments can alert the scientific community to potential problems concerning a published manuscript and allow the authors to respond. Some concerns may be easily addressed, while others may require correction or even retraction of an article. As it may be difficult to fully evaluate published results without access to the primary data, journals have a responsibility to respond to readers' concerns and to work with authors to resolve them. Historically, both journals and institutions have sometimes failed to live up to their obligations in addressing problematic articles and allegations of research misconduct (15, 16). Although postpublication review occurs relatively late in the process, it provides an important safeguard that allows even published findings to be corrected.

*vi. Increasing journal-based research.* Journals tend to focus more on publishing scientific information than on analyzing their own performance and are often secretive about their publication practices.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Reference information:** *J Clin Invest*. 2018;128(10):4243–4244. <https://doi.org/10.1172/JCI123884>.

However, much of what we have learned about problems with the scientific literature has come from editors and journals willing to analyze and share their experiences. As one example, the journal *Molecular and Cellular Biology* recently published their experience in dealing with inappropriate image duplications (9). Their experience suggests that the majority of inappropriate image duplications are the result of simple errors, although approximately 10% of these led to retractions, and that prepublication image screening may be more efficient than waiting to deal with problems after publication. Going forward, it will be important for more journals to examine their own experiences and share them with the scientific community in order to establish best practices and improve the entire publishing enterprise.

vii. *Fostering a culture of rigor.* In recent decades, many life science researchers have learned to accept a culture of impact, which stresses publication in high-impact journals, flashy claims, and packaging of results into tidy stories. Today, a scientist who publishes incorrect articles in high-impact journals is more likely to enjoy a successful career than one who publishes careful and rigorous studies in lower-impact journals, provided that the publications of the former are not retracted. This misplaced value system creates perverse incentives for scientists to participate in a “tragedy of the commons” that is detrimental to science (17). The culture of impact must be replaced by a culture of rigor that emphasizes quality over quantity. A focus on experimental redundancy, error analysis, logic, appropriate use of statistics, and intellectual honesty can help make research more rigorous and likely to be true (18). The publication of confirmatory or contradictory findings must also be encouraged to allow the scientific literature to provide a more accurate and comprehensive reflection of the body of scientific evidence (19).

There is no single simple remedy for improving the reliability of the scientific literature. The scientific enterprise is a complex system with multiple interacting components, each of which has a critical role to play in ensuring the integrity of the whole (20). Reengineering the system to incorporate multiple fail-safe features from data acquisition to postpublication review will better prevent, detect, and correct failures and result in a more reliable scientific literature. Science is a human endeavor and, as such, will never be perfect. Nevertheless, remarkable achievements have been made in science and technology, which remain humanity’s greatest hope for the many challenges it currently faces. To obtain the full benefit of science, its literature must be reliable. For too long, science has relied on the mantra that it is self-correcting. Science can be self-correcting, but only through the concerted efforts of all scientists working at multiple levels. The steps outlined here provide a blueprint to begin this process.

Address correspondence to: Arturo Casadevall, Department of Molecular Microbiology and Immunology, Johns Hopkins School of Public Health, 615 N. Wolfe Street, Room E5132, Baltimore, Maryland 21205, USA. Phone: 410.955.3457; Email: [acasadevall@jhu.edu](mailto:acasadevall@jhu.edu).

1. Downer J. *When Failure Is an Option: Redundancy, Reliability and Regulation in Complex Technical Systems*. London, United Kingdom: ESRC Centre for Analysis of Risk and Regulation (CARR), London School of Economics and Political Science; 2009.
2. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. *Nature*. 2012;483(7391):531–533.
3. Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. *Proc Natl Acad Sci U S A*. 2012;109(42):17028–17033.
4. Bik EM, Casadevall A, Fang FC. The prevalence of inappropriate image duplication in biomedical research publications. *MBio*.

- 2016;7(3):e00809-16.
5. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.
6. Casadevall A, Fang FC. Reforming science: methodological and cultural reforms. *Infect Immun*. 2012;80(3):891–896.
7. Fang FC, Casadevall A. Reforming science: structural reforms. *Infect Immun*. 2012;80(3):897–901.
8. Bosch G, Casadevall A. Graduate biomedical science education needs a new philosophy. *MBio*. 2017;8(6):e01539-17.
9. Bik EM, Fang FC, Kullas AL, Davis RJ, Casadevall A. Analysis and correction of inappropriate image duplication: The Molecular and Cellular Biology experience [published online ahead of print July 23, 2018]. *Mol Cell Biol*. <https://www.doi.org/10.1128/MCB.00309-18>.
10. Walker R, Rocha da Silva P. Emerging trends in peer review—a survey. *Front Neurosci*. 2015;9:169.
11. Higgins JR, Lin FC, Evans JP. Plagiarism in submitted manuscripts: incidence, characteristics and optimization of screening—case study in a major specialty medical journal. *Res Integr Peer Rev*. 2016;1:13.
12. Powell K. Your own desktop crime lab. *Nat Med*. 2006;12(5):493.
13. Nuijten MB, Hartgerink CH, van Assen MA, Epskamp S, Wicherts JM. The prevalence of statistical reporting errors in psychology (1985–2013). *Behav Res Methods*. 2016;48(4):1205–1226.
14. Didier E, Guaspere-Cartron C. The new watchdogs’ vision of science: A roundtable with Ivan Oransky (Retraction Watch) and Brandon Stell (PubPeer). *Soc Stud Sci*. 2018;48(1):165–167.
15. Resnik DB, Dinse GE. Scientific retractions and corrections related to misconduct findings. *J Med Ethics*. 2013;39(1):46–50.
16. Gunsalus CK, Marcus AR, Oransky I. Institutional research misconduct reports need more credibility. *JAMA*. 2018;319(13):1315–1316.
17. Casadevall A, Fang FC. Causes for the persistence of impact factor mania. *MBio*. 2014;5(2):e00064–e00014.
18. Casadevall A, Fang FC. Rigorous science: a how-to guide. *MBio*. 2016;7(6):e01902-16.
19. van Assen MA, van Aert RC, Nuijten MB, Wicherts JM. Why publishing everything is more effective than selective publishing of statistically significant results. *PLoS One*. 2014;9(1):e84896.
20. Broccardo C, et al. RePAIR Consensus guidelines: Responsibilities of publishers, agencies, institutions, and researchers in protecting the integrity of the research record. CODEX. <https://publicationethics.org/files/RePAIR%20Consensus%20Guidelines.pdf>. Accessed July 31, 2018.