

Promoting Responsible Scientific Research



A report from the
**AMERICAN
ACADEMY OF
MICROBIOLOGY**
RECOGNIZING SCIENTIFIC EXCELLENCE



Promoting Responsible Scientific Research

Written by Erika W. Davies and Diane D. Edwards

Report on an American Academy of Microbiology Colloquium held in Washington, DC, from 14 to 15 October 2015

The American Academy of Microbiology (Academy) is the honorific branch of the American Society for Microbiology (ASM), a nonprofit scientific society with nearly 48,000 members. Fellows of the Academy have been elected by their peers in recognition of their outstanding contributions to the microbial sciences. Through its colloquia program, the Academy draws on the expertise of these fellows to address critical issues in the microbial sciences.

This report is based on the deliberations of experts who gathered for two full days to discuss a series of questions developed by the steering committee regarding scientific publication standards and how expectations can be upheld globally, procedures that can be implemented in the laboratory to promote

ethical practices, and the appropriate consequences for data mishandling. This report has been reviewed by the majority of participants, and every effort has been made to ensure that the information is accurate and complete. The contents reflect the views of the participants and are not intended to reflect official positions of the Academy or ASM.

The Academy acknowledges Karl F. Hessler for his assistance in editing this report.

Contents of the report may be distributed further so long as the authorship of the Academy is acknowledged and this disclaimer is included.

BOARD OF GOVERNORS, AMERICAN ACADEMY OF MICROBIOLOGY

Michele S. Swanson, Ph.D., Chair
University of Michigan

Terence Dermody, M.D.
Vanderbilt University

Margaret McFall-Ngai, Ph.D.
University of Hawaii at Manoa

Martin J. Blaser, M.D.
New York University

Stan Fields, Ph.D.
University of Washington

Mary Ann Moran, Ph.D.
University of Georgia

Donald A. Bryant, Ph.D.
Pennsylvania State University

James M. Hughes, M.D.
Emory University

Graham C. Walker, Ph.D.
Massachusetts Institute of Technology

Nancy Craig, Ph.D.
Johns Hopkins University School of Medicine

Steven Lindow, Ph.D.
University of California, Berkeley

Alice S. Weissfeld, Ph.D.
Microbiology Specialists

COLLOQUIUM STEERING COMMITTEE

Arturo Casadevall, M.D., Ph.D.
Johns Hopkins Bloomberg School of Public Health

Ferric C. Fang, M.D.
University of Washington School of Medicine

Lee M. Ellis, M.D.
University of Texas M.D. Anderson Cancer Center

Margaret McFall-Ngai, Ph.D.
University of Hawaii at Manoa

PARTICIPANTS

Linda Amaral-Zettler, Ph.D.
Marine Biological Laboratory,
Woods Hole, Massachusetts

Samantha Elliott, Ph.D.
St. Mary's College of Maryland

Brian Martinson, Ph.D.
HealthPartners Institute for
Education and Research

C. Glenn Begley, MBBS, Ph.D.
TetraLogic Pharmaceuticals

Leonard P. Freedman, Ph.D.
Global Biological Standards Institute

Alexander McAdam, M.D., Ph.D.
Boston Children's Hospital

Arthur Buchberg, Ph.D.
American Association
for Cancer Research

Jarrad Hampton-Marcell, M.Sc.
Argonne National Library

Louis Rice, M.D.
Warren Albert School of Medicine
Brown University

Michael J. Buchmeier, Ph.D.
University of California, Irvine

Joseph R. Haywood, Ph.D.
Michigan State University

Rozanne Sandri-Goldin, Ph.D.
University of California, Irvine

Harold Drake, Ph.D.
Universität Bayreuth

Igor Jouline, Ph.D.
University of Tennessee

Jizhong (Joe) Zhou, Ph.D.
University of Oklahoma

Simon Liu, Ph.D.
U.S. Department of Agriculture

ACADEMY STAFF

Marina Moses, M.S., Dr.P.H.
Director

Isabel Hinestrosa
Board of Governors Program Assistant

Erin Seglem
Board of Governors Program Assistant
(through March 2016)

Virginia Dolen
Program Manager

Daniel Peniston
Colloquium and Outreach
Program Assistant

ASM LEADERSHIP

Timothy Donohue, Ph.D.
ASM Past President

University of Madison-Wisconsin

Introduction

Concerns around the topic of reproducibility, or rather lack thereof, have exploded onto the public and scientific scenes in recent years. A quick search on “reproducibility” in PubMed reveals that numerous editorials have been published in the scientific literature on the subject, especially within the past five years. Growing concern over a perceived crisis in reproducibility led to a 2014 gathering of scientific and publishing experts at a workshop convened by the National Institutes of Health (NIH), Science, and the Nature Publishing Group. The resultant *Principles and Guidelines in Reporting Preclinical Research* were widely disseminated and provided recommendations to journals with the aim of increasing transparency and reproducibility in scientific publishing. In addition to endorsing these principles and guidelines, the American Society for Microbiology (ASM) examined its role in addressing these issues within the sphere of the microbial sciences and turned to the American Academy of Microbiology (Academy) for expertise and leadership on the topic of reproducibility in microbial research.

The Academy represents members fully invested in the highest standards of rigor and reproducibility in experimental design, interpretation, and reporting. A steering committee appointed prior to the event developed the questions that guided discussion during the colloquium, hosted on 14 and 15 October 2015 at the Washington, DC, headquarters of the ASM. Invited participants included biomedical researchers, university science faculty, editors of several peer-reviewed scientific journals, and ASM personnel involved in the society’s extensive publishing efforts. In December 2014, ASM’s then President, Tim Donohue, and the society’s editors in chief had suggested the colloquium topic while attending the annual ASM Journals Editors in Chief meeting.

Colloquium participants considered issues related to reproducibility, the ethical conduct of scientific research, and good practices. The ASM and the Academy sought suggestions on, for example, how publication standards can be upheld globally, procedures that can be implemented in the laboratory to promote ethical practices, and the appropriate consequences for data mishandling. Representing various scientific disciplines, attendees divided into working groups to discuss, periodically reconvening for plenary sessions to report and review the group comments. This report summarizes the plenary discussions during the two-day colloquium.

Colloquium participants considered issues related to reproducibility, the ethical conduct of scientific research, and good practices.

Scope of the Problem

Case study I

William Summerlin

In 1974, William Summerlin was conducting research in transplantation immunology in the laboratory of Robert Good at Memorial Sloan-Kettering Cancer Center in New York City. The research dermatologist had reported successfully transplanting skin between genetically unrelated animals by culturing the skin in a special laboratory medium. The research had major implications for the field of tissue transplantation, but neither he nor other scientists had been able to reproduce his original results.

When Summerlin used a black pen to alter a patch of black mouse skin transplanted onto a white mouse, animal care technicians quickly discovered the fraud. Subsequent investigation by Memorial Sloan-Kettering revealed poorly conducted experiments and misleading statements about research results made by Summerlin in reports and to colleagues. All of his work was discredited, but the 1973 article in *Transplantation Proceedings* was never retracted.

Reproducing prior experimental work, by both original and outside investigators, has long been a foundational strength of the scientific process and is often touted as a “cornerstone” of the scientific endeavor. It is possible that problems with reproducibility were less apparent in the past, when science as an enterprise was considerably smaller. Fewer papers were published by a much smaller community, and those papers had very limited distribution. The current problem with reproducibility may not be a new phenomenon but rather a problem that has been exposed because of burgeoning growth and accessibility. Scientific publishing in the United States and worldwide has exploded in recent years. In 2012, there were an estimated 28,000 scientific journals, over 1 million new papers added to PubMed, and 1.8 to 1.9 million new papers published in science, technology, engineering, and mathematics (STEM; Ware and Mabe 2012). Shifts in scientific publishing—online journals, open access, and digital data in articles—necessitate reexamination of how to ensure reproducibility and transparency in the scientific process.

Regardless of whether the lack of reproducibility is a new problem or just a newly recognized issue, it is a topic of great concern among researchers, funding institutions, publishers, and society at large. Academy colloquium participants discussed various systemic problems in the current scientific enterprise, as well as the manifestations of these problems in the scientific literature.

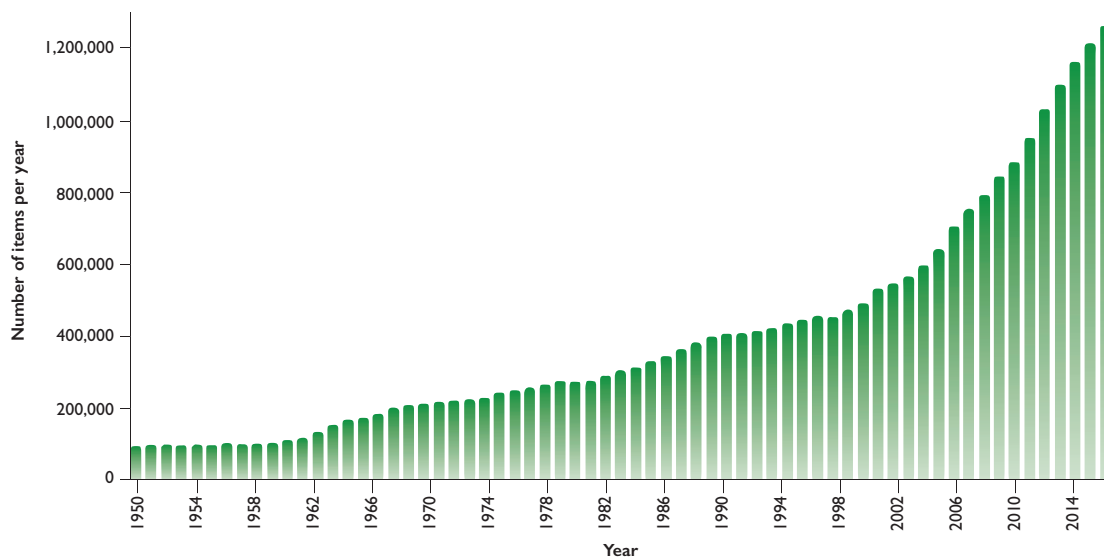
Systemic Problems Drive Data Reproducibility Issues

Colloquium participants agreed that systemic problems within the scientific enterprise contributed to the current environment of suboptimal data reproducibility. Competition for scarce resources likely lies at the heart of these issues.

An examination of these systemic issues can begin by focusing on the key players in this “scientific enterprise.” The primary investigator (PI) is the pivotal scientist in the enterprise, reliant upon academic, government, or private institutions for job placement, promotion, and/or salary. The PI employs laboratory personnel, students, and postdocs to perform experiments. These trainees, in turn, are dependent upon the PI for training, mentorship, graduation, and career advancement. Employing institutions, especially in the academic arena, rarely provide full financial support for the research or for the researchers. Instead, PIs are further dependent upon funding agencies to finance their research and their salaries, with the government funding the majority of life science research. The PI also relies upon the publishing community to communicate results and to maintain the scientific record. Under the current system of “science,” funding, hiring, and promotional decisions are largely based on the PI’s publishing record.

Competition plays a large part in driving this complex system forward. Scientists compete for

Publications Indexed in PubMed



Publications indexed in PubMed. Fang PowerPoint slide, Academy colloquium. Data obtained 9 May 2016.

jobs and for funding by showing their competence and the worth of their scientific ideas. Competition can drive great innovation. But what happens when the delicate balance of this system is altered? Government funding, by its very nature, is dependent upon external factors, such as the economy, the perceived need, and the political environment. A sudden decrease in funding availability upsets the balance, and competition no longer differentiates good science from poor science. Instead, excellent researchers find themselves competing for scarce resources to sustain their research, laboratories, and livelihoods. Is such competition beneficial or detrimental to science (Fang and Casadevall 2015)? Colloquium participants acknowledged that this double-edged sword could persist as a complication in resolving problems with reproducibility. Positive forms of competition incentivize the quest for new knowledge or the creation of a particular product. Negative competition—where competition dictates job security or the ability to continue practicing science—can impair creativity and spawn undesirable research practices.

Concurrent with the pressures induced by decreased funding availability is the rapidly expanding size of today's scientific enterprise—in both numbers of projects and the personnel and resources required—that creates huge pressures on institutions to hit regular home runs with research results. Colloquium participants expressed concern that the focus on “high-impact” science might distort the course of science, such that some important questions are no longer pursued (Casadevall and Fang, 2015b). Publication in high-impact journals has disproportionate rewards for those who succeed. Coupled with the expectation of these journals to publish innovative, flashy, and newsworthy science, these academic and financial

rewards might tempt scientists to decrease rigor, artificially tidy up results, and inflate import in an attempt to submit the “perfect story,” which is, frankly, rare in biology.

At the center of colloquium attendees' concerns regarding heightened competition is a seemingly innocuous metric called the impact factor (IF). The IF is used as a proxy to measure the relative importance of a journal within a scientific field and is calculated by dividing the number of citations a journal receives in a given year by the total number of “citable items” published in that journal during the previous two years (<http://wokinfo.com/essays/impact-factor/>). Unfortunately, the IF metric frequently is misinterpreted by the scientific community as a proxy for the quality or importance of the individual papers published in a journal. Thus, the importance and quality of an innovative, but poorly conducted, study published in a high-impact journal might be misconstrued, overshadowing the rigorous, but less flashy, science published in a lower-impact journal. This distortion reaches throughout the scientific enterprise, as professional and funding decisions are often tied to this misinterpretation of the IF. Consequently, colloquium members recommended a shift away from “IF mania” toward alternative metrics for high-quality research (Casadevall and Fang 2014).

In addition to the systemic problems linked to the current economic incentive structure are problems created by the rapid advance of technology that has led to the acquisition of vast amounts of data, while the knowledge and infrastructure for interpreting, analyzing, and storing said data are lagging behind. Present-day research methods can generate massive data sets and/or necessitate complex statistical analyses. When the amount

Case study 2

Haruko Obokata

In January 2014, two *Nature* articles by Haruko Obokata and her coauthors attracted considerable attention with a simple way to turn ordinary body cells into pluripotent cells. A short soak in a weak citric acid solution could create cells able to grow into any type of cell, a highly desirable achievement in stem cell research. Then excitement over the “stimulus-triggered acquisition of pluripotency” (STAP) cells quickly faded. Days after publication, comments on social media like blogs and Twitter questioned whether article images had been altered and noted that text had been copied from other papers. Most notably, despite the described soak's simplicity, no one else could replicate the original results.

Obokata's employer, the RIKEN Center for Developmental Biology (CDB) in Kobe, Japan, investigated and concluded scientific misconduct. The STAP cells actually were embryonic stem cells. The *Nature* articles were retracted by the journal. Fallout was rapid and harsh. RIKEN completely reorganized its CDB, severely cutting funding and closing or reassigning half of CDB's labs. Obokata's supervisor committed suicide. Other CDB researchers with no ties to the STAP research were negatively impacted. By the end of 2014, after failing to repeat her published results, Obokata resigned. She later was stripped of her doctorate by Waseda University. The severe consequences and surrounding media frenzy initiated additional controversies and soul searching within the global research community, raising questions about appropriate responses to fraud.



Case study 3

Andrew Wakefield

A 1998 article published (and later retracted) by *The Lancet* that linked the onset of autism with childhood vaccination is arguably the most influential retracted scientific paper to date, at least in terms of its impact on public health. At London's Royal Free Hospital, physician Andrew Wakefield conducted research on gastroenterology and autism. In the article, he and a dozen coauthors described the case histories of 12 children, including incidences of chronic enterocolitis, inflammatory bowel disease, regressive developmental disorder, and immunization with the MMR vaccine. At a prepublication news conference, Wakefield expressed his autism-related concerns with using the combination MMR vaccine rather than single-disease vaccines.

The *Lancet* paper quickly attracted major public attention, despite its relatively few research subjects. Vaccination rates fell sharply in the United Kingdom. The study helped launch the antivaccination movement in the United States. It also initiated a long series of subsequent publications from other researchers, as well as institutions like the U.S. Institute of Medicine, rejecting causal relationships between the MMR vaccine and autism. In Britain, the General Medical Council revoked Wakefield's medical license, citing, among other findings, that he had not disclosed funding from lawyers suing vaccine manufacturers.

of data exceeds laboratory and/or institutional capabilities, researchers might be forced to parse results in ways that introduce biases or inaccuracies. Cutting-edge technologies and research fields might be too new to have established generally accepted standards in data production and evaluation. Additionally, high-throughput tools may themselves introduce "noise," which can complicate evaluation or reexamination efforts. Advanced technology has created paradigm shifts in scientific publishing. Online journals, open access, and digital data have fundamentally changed how scientists receive and communicate information, necessitating reexamination of how to ensure reproducibility and transparency in the scientific process.

During the colloquium, it was expressed that the real threat to ethical conduct in science lies in the current tension between the existing reward systems and the normative standards of the scientific process. In other words, the normative standards of science that emphasize rigor, reproducibility, and responsibility are out of sync with the current reward system, which emphasizes large numbers of publications and publication in high-impact journals. This intensely competitive, "winner-take-all" model, as perpetuated by research and funding institutions and scientific publishers, is perverting a centuries-old system of objectively seeking new truths, warned colloquium members. Scientists are pressured to show greater productivity at the expense of scientific rigor. Younger scientists might feel coerced to report dubious or doctored data to publish and to obtain funding and facilitate promotion within academia. In this environment of hypercompetition and fixation on "high-impact science," too little time and resources are devoted to developing the infrastructure, analytical tools, and redundancies necessary to maximize reproducibility.

Systemic Problems Manifest as Irreproducibility

A quick look through recent retractions sheds light on some of the factors contributing to the problem of reproducibility in science (Casadevall et al. 2014). Many retractions occur due to mistakes and/or sloppiness. Erroneous duplication of images can occur when researchers accidentally insert the wrong image. Failure to include the appropriate controls can invalidate an experiment. Misuse of statistical analyses might lead to misinterpretation of the significance of a finding. Failure to validate a cell line (or other reagents) may introduce confounding effects due to contamination. Another factor leading to irreproducible data is one of bias. Scientists may pick and choose among experimental outcomes, ignoring contradictory results in favor of those that support their hypothesis. Outliers may be specifically removed from analyses. At the furthest extreme, scientific misconduct can and does occur, often

in an attempt to cover up mistakes such as those described above. At the farthest end of the spectrum are scientists who intentionally commit fraud in an attempt to further their own careers, at the expense of the scientific community and ultimately at the expense of the public.

Attendees of the Academy colloquium addressed the causes and effects of three major contributors to lack of reproducibility: (i) sloppy science, (ii) bias in interpreting and reporting results, and (iii) misconduct. While there might be debate over the relative importance of these factors, there was agreement at the Academy colloquium that intentional misconduct, though the most headline grabbing, probably accounts for relatively few cases.

Sloppy science, however, is widespread and includes published research with uncontrolled, or improperly controlled, experimental variables, contaminated reagents, and/or inappropriate use of statistics. A colloquium presenter shared that a significant percentage of recently surveyed academic faculty admitted to instances of sloppy science, including inadequate recordkeeping, insufficient monitoring of research projects, cutting corners, and circumventing or ignoring certain material-handling protocols (Martinson et al. 2005). Advanced technologies likely add to mistakes in the reporting of data as well, as databases full of similar digital images can lead to improper cataloging of images (or other data) and the cutting and pasting of incorrect or duplicate images for publication. Further, image-editing tools offer new opportunities for the introduction of mistakes into published figures. It is likely that more rigorous training and oversight in science practice and reporting could have a large impact on improving reproducibility.

Bias is likely a more difficult factor to combat, since it relies upon knowledge of a scientist's published and unpublished work. It is a well-known and frequently aired complaint that negative results are unlikely to be published, especially in "top-tier" journals. Multiple problems stem from this deficiency in the scientific literature. First, the lack of communication over what does not work likely leads to unnecessary repetition of experimental conditions across laboratories. Second, selection of only positive results likely skews the literature in a nonreproducible fashion. For example, a group of scientists performs an experiment 20 times and obtains the expected results in only three of the experiments. These three experiments may be included in a paper, but they do not represent the reality of the situation, in which it is more likely NOT to get those results. Similarly, scientists might exclude outliers from their published data set, which artificially decreases variability and likely increases the apparent significance of their reported finding.

Given the systemic problems in the current scientific enterprise, as discussed above, it is not surprising that people manipulate the system in attempts to get their work published in the pursuit of funding and career advancement. The NIH follows Public Health Service policies on research misconduct, Title 42 Code of Federal Regulations part 93, and defines research misconduct as “fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results.” While misconduct is generally considered to play a small role in the current reproducibility crisis, a 2012 study using the PubMed database found that about two-thirds of article retractions were linked to misconduct (Fang et al. 2012). Regardless of the extent to which misconduct contributes to the current problems surrounding reproducibility, colloquium participants agreed that vigilance is required on the part of institutions, publishers, and funding agencies to monitor and respond to acts of research misconduct. Improved and consistent training regarding the responsible conduct of research is recommended for scientists at all levels of their careers and across academic, federal, and private institutions.

Consequences of Poor Scientific Practices

Everyone loses when the scientific process in biological sciences fails to meet standards of rigor and responsibility—scientists in all stages of their careers, research and academic institutions, biomedicine-related industry, and any stakeholder in the public trust in science. Flawed research squanders scarce funding and good reputations. One study estimated that papers supported by federal funds that were retracted in 2012 for misconduct each represented an average waste of \$393,000 in research dollars, totaling \$58.5 million that year (Stern et al. 2014). It is not possible to adequately quantify the oppor-

tunity costs associated with irreproducible studies. The costs are, however, undoubtedly sizeable and important, whether in lost time, money, and careers; negative publicity; or patient costs and lives. Colloquium participants agreed that, going forward, the focus should be *preventing* these opportunity costs, regardless of amounts.

Journals also have a stake in the publication of research that cannot be reproduced. Retraction of problematic research papers may result in negative publicity for journals, though colloquium consensus suggests otherwise. In general, respected journals have protocols and author guidelines in place and should be credited for efforts to retract faulty papers or, in less extreme cases, correct mistakes.

Poor science can also pose a threat to public health and patient lives; there are publicized examples of rippling negative impacts long after misconduct or neglect is exposed. For example, the erroneous notion that measles, mumps, and rubella (MMR) vaccination is associated with autism persists in the population, leading some to refuse immunization despite the facts that the original paper was retracted due to misconduct (Editors of *The Lancet* 2010) and that numerous subsequent studies have shown no association.

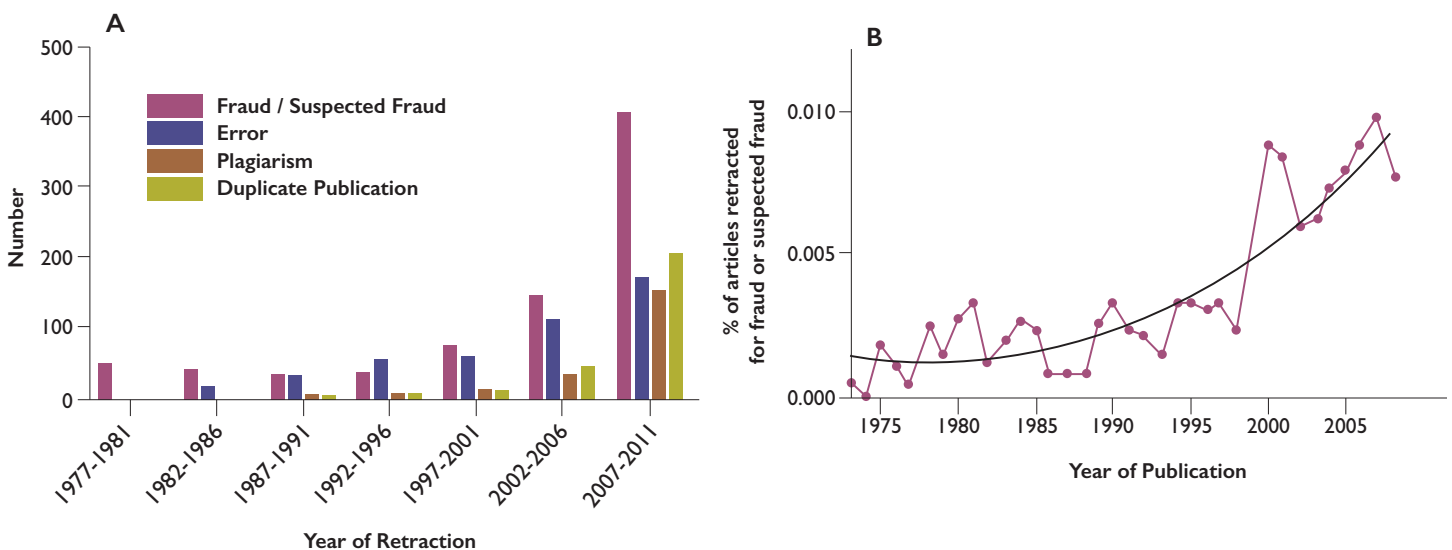
Unethical practices in scientific research affect everyone, from the general public to students to colleagues in the science community. Poor research practices can provoke negative public opinion about and distrust of science, raising further the already high stakes for changes within the scientific enterprise. We all are stakeholders in how science is conducted, in particular the patients, taxpayers, trainees, and other scientists who might be misled by specific results.

Case study 4

Naoki Mori

Image manipulation led to dozens of journal retractions for virologist Naoki Mori. A researcher at the University of the Ryukyus in Nishihara, Okinawa, he had altered images depicting protein or nucleic acid gels in his articles. He received a 10-year publishing ban from the ASM, whose journals have retracted eight Mori papers (*Infection and Immunity* and *Journal of Virology*). There have been more than 30 additional retractions of Mori papers by non-ASM journals.

Mori's misconduct was exposed when a journal reviewer noticed the duplication and misrepresentation of a previously published figure in a submitted article. The journal notified the University of the Ryukyus, which conducted an internal investigation and subsequently notified *Infection and Immunity* and other journals about problems with images in Mori papers. The university fired Mori in 2010, reinstating him in 2013.



The role of suspected fraud and other systemic problems in retractions of scientific publications. (Fang et al. 2012, PNAS).

Academy Recommendations

ACADEMY RECOMMENDATIONS FOR DATA REPRODUCIBILITY

Design rigorous and comprehensive evaluation criteria to recognize and reward high-quality scientific research.

Require universal training in good scientific practices, appropriate statistical usage, and responsible research practices for scientists at all levels, with training content regularly updated and presented by qualified scientists.

Establish open data as the standard operating procedure throughout the scientific enterprise.

Encourage scientific journals to publish negative data that meet standards of quality.

Agree upon common criteria among scientific journals for retraction of published papers, to provide consistency and transparency.

Strengthen research integrity oversight and training.

Participants in the Academy colloquium agreed that there is a current problem with reproducibility in the scientific endeavor. The incremental advance of science requires a solid foundation of rigorous and reproducible information to drive basic science toward translatable solutions for our world. Sloppy science, reporting bias, and misconduct contribute to the publication of erroneous and nonreproducible information, which obstructs the forward movement of potentially beneficial lines of scientific inquiry. Colloquium attendees considered cultural changes aimed at fixing the systemic problems discussed above. Further, they encouraged the development of standard and transparent processes for responding to deliberate and accidental errors in scientific literature. By encouraging cultural changes and consistent identification and responses to erroneous

information, colloquium participants hope to increase the reproducibility of published data.

The multiple recommendations articulated during the Academy colloquium collectively target promoting best practices in biological research. They address a diverse range of problem areas identified by the attendees as contributing to the current problems with lack of reproducibility in published research. It is crucial that everyone with a direct stake in science collaborates to achieve uniform, top-quality standards. The collective scientific community must respond consistently to substandard research practices. All stakeholders, including funding agencies, research and academic institutions, journals, professional associations, individual investigators, and research groups, have additional responsibilities specific to their respective roles in scientific practice.

Changing the Culture

Revamp the Research Funding Structure and Develop Appropriate Metrics To Assess Scientists and Their Work

The present state of research funding is a central concern discussed at the colloquium. In the words of one participant, "scarcity of funding is the primary driver of sloppy and dishonest science." The current reliance upon soft (grant) money for scientific research, coupled with the training of too many scientists for too few positions, is a pivotal systemic problem and requires redress. In

the absence of a large and sustained increase in research funding through the usual mechanisms, colloquium participants called for the creation of more sustainable research funding and the balancing of the scientific workforce supply and demand, possibly through the training of fewer scientists and/or the creation of more staff scientist positions.

Colloquium participants further voiced concern over the continued misuse of IFs in the evaluation of scientists for tenure, promotion, and funding. As previously discussed, IF is a measure of a journal's citations and should not be used as a proxy to measure the importance of individual papers or authors. Consideration of the totality of a scientist's contributions, including teaching, mentoring, and engagement with the scientific community through such activities as peer review and editing, would likely be more fair and balanced. The participants strongly favored the development of alternative metrics for use in combination with other evaluations for a comprehensive, holistic assessment of a scientist's contribution and value.

Specific suggestions by colloquium attendees included the development of alternative criteria to specifically address research quality and impact; a role for journals in identifying and recognizing current or previously published papers of importance to the community, such as the production of summary journal issues that provide updates on the impacts of previously published

papers; and standardized criteria for the ranking of journals, independently of IF (e.g., *U.S. News & World Report* university rankings).

RECOMMENDATION

Design rigorous and comprehensive evaluation criteria to recognize and reward high-quality scientific research.

- Defuse the unwarranted importance of IFs and simple metrics in ranking research, scientists, and journals.
- Support efforts by universities and funding agencies to minimize or eliminate the current use and misuse of IFs and simple metrics as criteria for academic advancement.
- Evaluate researchers being considered for promotion or funding based on their entire output and longer-term contributions (including teaching, administration, mentoring, and patient care), with reviews conducted by relevant peers.

Improve Training in Good Scientific Practice

A top priority for colloquium participants is improved training in the execution and responsible conduct of scientific research. Scientific training is, by its very nature, individualized to particular fields of study, institutions, and laboratories. Such training typically focuses on STEM students and postdocs, though there are data showing that the problem of misconduct spans generations and includes senior scientists (Fang et al. 2013). Therefore, colloquium participants called for intensified, expanded efforts to educate scientists on fundamental best practices

in both conducting and reporting research at all research institutions and across the career spectrum. Colloquium participants expressed hope that institutions and senior scientists will focus more on training the next generation of scientists in rigor and responsibility, rather than viewing them simply as a means for the generation of large amounts of data.

Colloquium participants suggested innovative ideas that could be implemented by institutions to improve rigor and reproducibility through training.



Of particular concern to colloquium members is improved training in statistical analytical methodology, since misapplication of statistics is common in the scientific literature (Ioannidis 2005; Strasak et al. 2007; Worthy 2015; Weissgerber et al. 2016; Chavalarias et al. 2016). Participants also agreed that the topic of data sharing and compliance with data sharing policies be included in the training of all scientists. They discussed the utility of review clubs for researchers, from graduate students to PIs, to vet experimental techniques and the resultant data. Journal clubs were mentioned as valuable

teaching venues for examination of exemplary (and nonexemplary) published papers. Importantly, participants agreed that consideration of data sharing be included in the training of all scientists.

RECOMMENDATION

Require universal training in good scientific practices, appropriate statistical usage, and responsible research practices for scientists at all levels, with training content regularly updated and presented by qualified scientists.

Increase Reproducibility and Transparency through Data Sharing

Colloquium participants recognized the need for greater transparency in the reporting of scientific data in any and all efforts to improve reproducibility. Without access to data, it is virtually impossible to evaluate reproducibility and possible causes or explanations for any lack of reproducibility. It was recently reported that only 13% of authors with scientific articles published in 2014 made their data freely accessible to others (Womack 2015). Colloquium participants agreed that data generated by experiments should be accessible outside the research group conducting the experiments.

Opinions varied as to how open data could or should be made available. Colloquium participants agreed that specific requirements for data availability would likely vary by discipline. Therefore, the scientists, institutions, funding agencies, and/or scientific societies encompassing specific fields should develop best practices and guidelines to address the following: the types or levels of data to be shared (e.g., raw vs. processed), designating responsibility for the storing and sharing of data, and the appropriate centralization of data.

Examples of well-established standards for data sharing include those developed by the Genomic Standards Consortium and various repositories of “omics data.”

RECOMMENDATION

Establish open data as the standard operating procedure throughout the scientific enterprise.

- Develop common standards for data sets and laboratory notebooks (i.e., “high-value” data) and for access to those data.
- Make the commitment to share data a prerequisite to publication.
- Archive open data from published papers in a predetermined location and a useable format for optimal accessibility.
- Standards and procedures should be informed by partners at national and international levels, including federal agencies, funding institutions, research and academic institutions, journal publishers, and scientific societies.

Publish Negative Data

Rigorously executed experiments that address specific hypotheses often result in “negative” data—results that do not support the hypothesis. Such results often go unpublished, since they are viewed as lacking innovation. Participants in the Academy colloquium argued that both positive and negative data have value for the scientific community and that reporting of negative data should be encouraged in respected venues. As with positive data, the value of publishing a set of negative data is proportional to the value of the original research question.

Colloquium participants suggested that there should be a formal mechanism for the publication

of negative data, likely through an alternative to the mechanism for the publication of standard journal manuscripts, such as a separate online forum or dedicated section in the journal. Participants suggested that the data and detailed experimental methods could be published without extensive background and discussion sections. Regardless of the format, the participants suggested that peer review be included to ensure that negative studies conform to the same standards of quality and rigor as standard journal content.

RECOMMENDATION

Encourage scientific journals to publish negative data that meet standards of quality.

Responding to the Problem

Develop Common Processes for Correcting the Scientific Literature

In addition to suggesting cultural changes aimed at alleviating the reproducibility crisis in published science, Academy colloquium participants considered appropriate responsibilities and responses to the current situation. Instructions to authors (ITAs) are used by many journals to inform authors of their responsibilities, including compliance with data, material, and software sharing policies, in addition to other ethical considerations. Colloquium participants expressed concern that while most journals use ITAs, the policies are not consistently followed or enforced. They encouraged the development of tools, such as checklists, to help authors adhere to publishing ethics policies. While reproducibility problems typically become apparent following publication, publishers are increasingly using approaches to detect problems prior to publication, as through the use of forensic image analysis software to screen images and plagiarism detection software to detect text similarities. However, colloquium participants expressed concern that the current tools for

image analysis are inadequate for the detection of some types of image manipulation and duplication. Nonetheless, participants encouraged attempts to correct or prevent erroneous, nonreproducible, data from being published.

Journals and authors have a responsibility to correct the scientific literature if and when problems are identified following publication. There are multiple mechanisms available for publishers, including author corrections, expressions of concern, and retractions. However, implementation varies among journals and even from editor to editor, creating confusion instead of clarity. Colloquium participants expressed interest in standard practices for the implementation and communication of corrective actions.

RECOMMENDATION

Agree upon common criteria among scientific journals for retraction of published papers, to provide consistency and transparency.



Expand Training and Improve Responses to Breaches in Responsible Conduct of Research

The burden of investigation into issues of misconduct often is beyond the purview and capabilities of publishers. Instead, such investigations must be conducted by the employing institutions, which have access to detailed data in the form of laboratory notebooks and computer hard drives, in addition to the testimony of laboratory personnel and coworkers. Funding agencies typically play an oversight role in this process, with responsibility for ensuring that appropriate training and policies are in place at institutions receiving federal funds. The Office of Research Integrity (ORI) provides such oversight for Health and Human Services-funded research, as does the Office of the Inspector General for National Science Foundation-funded research. Colloquium participants expressed concern that the current system of institutional investigation and training with government oversight is insufficient due to a lack of consistency and transparency in the process. The funding institutions lack the funding and resources necessary to effectively monitor responsible conduct-of-research education and investigative efforts across the multitude of institutions that receive government funds. Further, privately or industry-funded research operates outside any government requirements and oversight.

Colloquium participants expressed concern over the primary role of institutions in the investigative process. There are several obvious disincentives for institutions to conduct investigations into allegations of possible misconduct, such as the possibility of losing institutional funding associated with a researcher's federal grant and the negative publicity for faculty or staff participating in wrongdoing. Most research institutions have mechanisms for

the anonymous reporting of problematic behavior, but the reporting mechanism and/or the chain of command can fail some complainants. Colloquium participants collectively urged greater care in safeguarding whistleblowers' careers and suggested that the investigative process and responsibility should be removed from the department of concern. Such responsibility might be better assumed by an institution's dean, provost, or president or an independent body outside the institution.

The Academy colloquium participants also discussed the upsurge in anonymous online platforms like PubPeer for reporting and commenting on suspected cases of poor practice. Some participants prefer to have confidential allegations come directly to the journal. However, several argued that there is value in an anonymous platform for evaluation of published data, especially in cases where there is a fear of retaliation. Colloquium participants recognized that the growing use of PubPeer and similar anonymous platforms likely stems from frustration over perceived inaction and inconsistency by journals, institutions, and funding agencies.

RECOMMENDATION

Strengthen research integrity oversight and training.

- Establish a comprehensive, consistent, and transparent system to detect and report problems to both research institutions and the ORI and to enforce research integrity at all levels.
- Develop standards for institutional ethics training and refresher courses that will include training for scientists at all career levels.



Conclusion

The Academy colloquium's focus on improving the reproducibility of published data paralleled the growing call to action within the U.S. scientific community to resolve problems with reproducibility. The NIH has launched several initiatives to enhance reproducibility and transparency. In 2013, it created an online forum (PubMed Commons) for discourse about published articles. NIH guidelines for reporting preclinical research now include more rigorous statistical analysis, data and material sharing, and expanded methods sections in published papers. Also encouraged is the use of checklists for authors of submitted papers to assess the rigor of experimental design. Nature journals now require the submission of a "[Reporting Checklist for Life Science Articles](#)," to provide details on experimental design and statistics, biological reagent validation, and data sharing. The NIH guidelines also encourage the development of best-practice guidelines for digital data and validation of biological reagents. Many journals and societies have endorsed the NIH guidelines, which should lead to continued adaptation of journal policy to NIH guidelines. Effective January 2016, NIH grant applications must satisfy new policies and instructions focused on the strength of the scientific premise underlying the proposed research, rigorous experimental design, close attention to biological variables, and authentication of key biological and/or chemical resources. Finally, the NIH recently launched a new web portal to provide guidance and training on rigor and reproducibility: <http://www.nih.gov/research-training/rigor-reproducibility>.

Participants in the Academy colloquium will work with ASM journals' editors and staff to take the lead on several of their recommendations. ASM journals have already adopted open data policies,

whereby authors must make data available—including the data, metadata, and methods used to reach the conclusions in the submitted paper and any additional data needed to replicate the study findings. This policy applies for ASM's two new journals mSystems and mSphere, as well as mBio. Additional roles for ASM in the uptake of the Academy colloquium recommendations include (i) expansion of open data policies to other journals and the consistent enforcement of open data policies across ASM journals, (ii) transparent and consistent management for the correction and/or retraction of flawed published material, (iii) development of mechanisms for the publication of negative data, and (iv) development of training for scientists on proper experimental design and responsible conduct of research.

Participants in the Academy colloquium emphasized that active participation is necessary across the scientific enterprise in order to improve the reproducibility of scientific research. Scientists, academic and research institutions, journals, professional societies, funding agencies, and the public must all be engaged. It is not enough simply to make recommendations and guidelines—they must be followed by implementation. Scientists and PIs must educate their students, postdocs, and laboratory personnel on acceptable and unacceptable research and reporting practices. Research institutions should take more responsibility in reviewing employees' research data and in providing training for all scientific staff. Editors of journals should consistently enforce clearly stated standards in publishing. Funding agencies can collaborate to develop common policies and requirements for applicants, who likely seek funding from multiple sources. Professional societies need to develop common standards and facilitate cultural change within their scientific fields.

References

- Achenbach J. 27 January 2015a. The new scientific revolution: reproducibility at last. Washington Post, Washington, DC. http://www.washingtonpost.com/national/health-science/the-new-scientific-revolution-reproducibility-at-last/2015/01/27/ed5f2076-9546-11e4-927a-4fa2638cd1b0_story.html. Accessed 20 November 2015.
- Achenbach J. 27 January 2015b. Many scientific studies can't be replicated. That's a problem. Washington Post, Washington, DC. <http://www.washingtonpost.com/news/speaking-of-science/wp/2015/08/27/trouble-in-science-massive-effort-to-reproduce-100-experimental-results-succeeds-only-36-times/>. Accessed 20 November 2015.
- Alberts B, Kirschner MW, Tilghman S, and Varmus H. 2014. Rescuing US biomedical research from its systemic flaws. Proc Natl Acad Sci U S A 111(16):5773–5777. <http://doi.org/10.1073/pnas.1404402111>.
- American Society for Cell Biology, Task Force on Scientific Reproducibility. 2015. How can scientists enhance rigor in conducting basic research and reporting research results? White Paper. American Society for Cell Biology, Bethesda, MD. <http://www.ascb.org/reproducibilityreport/>
- Bal L. 20 March 2015. Is science broken? The reproducibility crisis. BioMed Central. <https://blogs.biomedcentral.com/on-biology/2015/03/20/is-science-broken-a-reproducibility-crisis/>.
- Bebeau MJ. 2014. An evidence-based guide for ethics instruction. J Microbiol Biol Educ 15(2):124–129. <http://doi.org/10.1128/jmbe.v15i2.872>.
- Begley CG. 2013. Six red flags for suspect work. Nature 497:433–434. <http://doi.org/10.1038/497433a>.
- Begley CG, and Ellis LM. 2012. Raise standards for preclinical cancer research. Nature 483:531–533. <http://doi.org/10.1038/483531a>.
- Begley CG, and Ioannidis JP. 2015. Reproducibility in science: improving the standard for basic and preclinical research. Circ Res 116:116–126. <http://doi.org/10.1161/CIRCRESAHA.114.303819>.
- Bertuzzi S. 9 June 2015. Sensational numbers and the real irreproducibility crisis in preclinical research. American Society for Cell Biology, Bethesda, MD. <http://www.ascb.org/imaginary-numbers-and-the-real-irreproducibility-crisis-in-preclinical-research/>.
- Bowen A, and Casadevall A. 2015. Increasing disparities between resource inputs and outcomes, as measured by certain health deliverables, in biomedical research. Proc Natl Acad Sci U S A 112(36):11335–11340. <http://doi.org/10.1073/pnas.1504955112>. <http://www.pnas.org/content/112/36/11335.abstract>
- Buck S. 2015. Solving reproducibility. Science 348(6242):1403. <http://doi.org/10.1126/science.aac8041>.
- Callahan M, Wears RL, and Weber E. 2002. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. JAMA 287:2847–2850. <http://doi.org/10.1001/jama.287.21.2847>.
- Casadevall A, Bardeen JM, Bond JR, and Kaiser N. 2012. Misconduct accounts for the majority of retracted scientific publications. Proc Natl Acad Sci U S A 109(42):17028–17033. <http://doi.org/10.1073/pnas.1212247109>.
- Casadevall A, Ellis LM, Davies, EW, McFall-Ngai, M, and Fang FC. 2016. A framework for improving the quality of research in the biological sciences. mBio 7(4):e01256-16. <http://doi.org/10.1128/mBio.01256-16>.
- Casadevall A, and Fang FC. 2010. Reproducible science. Infect Immun 78(12):4972–4975. <http://doi.org/10.1128/IAI.00908-10>.
- Casadevall A, and Fang FC. 2014. Causes for the persistence of impact factor mania. mBio 5(2):e00064-14. <http://doi.org/10.1128/mBio.00064-14>.
- Casadevall A, and Fang FC. 25 July 2015a. Is science in crisis? Op-Ed. Baltimore Sun, Baltimore, MD. <http://www.baltimoresun.com/news/opinion/oped/bs-ed-science-crisis-20150725-story.html>.
- Casadevall A, and Fang FC. 2015b. Impacted science: impact is not importance. mBio 6(5):e01593-15. <http://doi.org/10.1128/mBio.01593-15>.
- Casadevall A, Steen RG, and Fang FC. 2014. Sources of error in the retracted scientific literature. FASEB J 28(9):3847–3855. <http://doi.org/10.1096/fj.140256735>.
- Chang AC, and Li P. 2015. Is economics research replicable? Sixty published papers from thirteen journals say “usually not.” Finance and Economics Discussion Series 2015-083. Board of Governors of the Federal Reserve System, Washington, DC. <http://doi.org/10.17016/FEDS.2015.083>.
- Chavalarias D, Wallach JD, Li AH, and Ioannidis JP. 2016. Evolution of reporting *P* values in the biomedical literature, 1990–2015. JAMA 315(11):1141–1148. <http://doi.org/10.1001/jama.2016.1952>.
- Collins FS, and Tabak LA. 2014. NIH plans to enhance reproducibility. Nature 505(7485):612–613. <http://doi.org/10.1038/505612a>.
- Cressy D. 13 October 2015. Poorly designed animal experiments in the spotlight. Nature. <http://doi.org/10.1038/nature.2015.18559>. http://www.nature.com/news/poorly-designed-animal-experiments-in-the-spotlight-1.18559?WT.ec_id=NEWSDAI-LY-20151014&spMailingID=49774566&spUserID=Mzc0OTAwM-jl3NDgSl&spJobID=781883625&spReportId=NzgxODgzNjlIS0. Accessed 15 November 2015.
- Devereaux ML. 2014. Rethinking the meaning of ethics in RCR Education. J Microbiol Biol Educ 15(2):165–168. <http://doi.org/10.1128/jmbe.v15i2.857>.
- Dorsey ER, de Roulet J, Thompson JP, Remnick JJ, Thai A, White-Stellato Z, Beck CA, George BP, and Moses H, III. 2010. Financial anatomy of biomedical research, 2003–2008. JAMA 303(2):137–143. <http://doi.org/10.1001/jama.2009.1987>.
- Editors of The Lancet. 2010. Retraction—ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. Lancet 375(9713):445. [http://doi.org/10.1016/S0140-6736\(10\)60175-4](http://doi.org/10.1016/S0140-6736(10)60175-4).

- Ellis LM. 2015. The erosion of research integrity: the need for culture change. *Lancet Oncol* 16(7):752–754. [http://doi.org/10.1016/S1470-2045\(15\)00085-6](http://doi.org/10.1016/S1470-2045(15)00085-6).
- Enserink M. 13 October 2015. Most animal research studies may not avoid key biases. *Science*. <http://doi.org/10.1126/science.aad4706>. http://news.sciencemag.org/plants-animals/2015/10/most-animal-research-studies-may-not-avoid-key-biases?utm_campaign=email-news-latest&et rid=LELLIS@MDANDERSON.ORG&et_cid=47557. Accessed 15 November 2015.
- Fang FC, Bennett JW, and Casadevall A. 2013. Males are overrepresented among life science researchers committing scientific misconduct. *mBio* 4(1): e00640-12. <http://doi.org/10.1128/mBio.00640-12>
- Fang FC, and Casadevall A. 2015. Competitive science: is competition ruining science? *Infect Immun* 83(4):1229–1233. <http://doi.org/10.1128/IAI.02939-14>.
- Fang FC, Steen RG, and Casadevall A. 2012. Misconduct accounts for the majority of retracted scientific publications. *Proc Natl Acad Sci U S A* 109(42):17028–17033. <http://doi.org/10.1073/pnas.1212247109>
- Grinnell F. 2014. The interrelationship between research integrity, conflict of interest, and the research environment. *J Microbiol Biol Educ* 15(2):162–164. <http://doi.org/10.1128/jmbe.v15i2.851>.
- Haug CJ. 2015. Peer-review fraud—hacking the scientific publication process. *N Engl J Med* 373(25):2393–2395. <http://doi.org/10.1056/NEJMp1512330>.
- Heitman E. 2014. Cross-cultural considerations in U.S. research ethics education. *J Microbiol Biol Educ* 15(2):130–134. <http://doi.org/10.1128/jmbe.v15i2.860>.
- Hotz RL. 10 August 2015. Scientists observe odd phenomenon of multiplying co-authors. *Wall Street Journal*, New York, NY. <http://www.wsj.com/articles/SB22470103492741843612004581156691220105286>
- Ioannidis JP. 2005. Why most published research findings are false. *PLoS Med* 2(8): e124. <http://doi.org/10.1371/journal.pmed.002012>.
- Johnson VE. 2013. Revised standards for statistical evidence. *Proc Natl Acad Sci U S A* 110(48):19313–19317. <http://doi.org/10.1073/pnas.1313476110>.
- Kolata G. 16 October 2015. A new, life-or-death approach to funding heart research. *New York Times*, New York, NY. http://www.nytimes.com/2015/10/20/health/a-new-life-or-death-approach-to-funding-heart-research.html?hp&action=click&pgtype=Homepage&module=first-column-region®ion=top-news&WT.nav=top-news&_r=0. Accessed 15 November 2015.
- Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, Crystal RG, Darnell RB, Ferrante RJ, Fillit H, Finkelstein R, Fisher M, Gendelman HE, Golub RM, Goudreau JL, Gross RA, Gubitza AK, Hesterlee SE, Howells DW, Huguenard J, Kelner K, Koroshetz W, Krainc D, Lasic SE, Levine MS, Macleod MR, McCall JM, Moxley RT, III, Narasimhan K, Noble LJ, Perrin S, Porter JD, Steward O, Unger E, Utz U, and Silberberg SD. 2012. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 490(7419):187–191. <http://doi.org/10.1038/nature11556>.
- Larivière V. 2015. Long-distance interdisciplinarity leads to higher scientific impact. *PLoS One* 10(3): e0122565. <http://doi.org/10.1371/journal.pone.0122565>. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0122565>
- Larivière V, and Gingras Y. 2010. On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology* 61(1):126–131. <http://doi.org/10.1002/asi.21226>. [http://www.chss.uqam.ca/Portals/0/docs/sts8020/\(23\)Larivière.GIngras.pdf](http://www.chss.uqam.ca/Portals/0/docs/sts8020/(23)Larivière.GIngras.pdf). Accessed 12 November 2015.
- Lawrence PA. 2007. The mismeasurement of science. *Curr Biol* 17(15): r583–R585. <http://doi.org/10.1016/j.cub.2007.06.014>.
- Lorsch JR, Collins FS, and Lippincott-Schwartz J. 2014. Fixing problems with cell lines. *Science* 346(6216):1452–1453. <http://doi.org/10.1126/science.1259110>.
- Macleod MR, Lawson McLean, A, Kyriakopoulou, A, Serghiou, S, de Wilde, A, Sherratt, N, Hirst T, Hemblade R, Bahor Z, Nunes-Fonseca C, Potluru A, Thomson A, Baginskaite J, Egan K, Vesterinen H, Currie GL, Churilov L, Howells DW, and Sena ES. 2015. Risk of bias in reports of *in vivo* research: a focus for improvement. *PLoS Biol*. 13(10): e1002273. <http://doi.org/10.1371/journal.pbio.1002273>.
- Marcus A, and Oransky I. 2014. What studies of retractions tell us. *J Microbiol Biol Educ* 15(2):151–154. <http://doi.org/10.1128/jmbe.v15i2.855>.
- Marcus A, and Oransky I. 19 August 2015. Getting the bogus studies out of science. *Wall Street Journal*, New York, NY. <http://www.wsj.com/articles/getting-the-bogus-studies-out-of-science-1440024409>
- Martinson BC. 2011. The academic birth rate. *EMBO Rep* 12(8):758–762. <http://doi.org/10.1038/embor.2011.142>.
- Martinson BC, Anderson MS, and de Vries R. 2005. Scientists behaving badly. *Nature* 435:737–738. <http://doi.org/10.1038/435737a>.
- McDowell GS, Gunsalus KT, MacKellar DC, Mazzilli SA, Pai VP, Goodwin PR, Walsh EM, Robinson-Mosher A, Bowman TA, Kraemer J, Erb ML, Schoenfeld E, Shokri L, Jackson JD, Islam A, Mattozzi MD, Krukenberg KA, and Polka JK. 2014. Shaping the future of research: a perspective from junior scientists. *FI000Res* 3:291. <http://doi.org/10.12688/fi000research.5878.2>.
- McGee R. 2014. Evaluation in RCR training—are you achieving what you hope for? *J Microbiol Biol Educ* 15(2):117–119. <http://doi.org/10.1128/jmbe.v15i2.853>.
- McNutt M. 2014. Journals unite for reproducibility. *Science* 346(6210):679. <http://science.sciencemag.org/content/346/6210/679.long>.
- National Academy of Sciences. 1992. *Responsible science*, vol. 1. Ensuring the integrity of the research process. National Academies Press, Washington, DC. <http://www.nap.edu/read/1864/chapter/2#13>. National Institutes of Health. 2015a. Principles and guidelines for reporting preclinical research. National Institutes of Health, Bethesda, MD. <http://www.nih.gov/research-training/rigor-reproducibility/principles-guidelines-reporting-preclinical-research>. Accessed 17 November 2015.

- National Institutes of Health. 2015b. Updated application instructions to enhance rigor and reproducibility. <http://www.nih.gov/research-training/rigor-reproducibility/updated-application-instructions-enhance-rigor-reproducibility>. National Institutes of Health, Bethesda, MD. Accessed 17 November 2015.
- National Institutes of Health. 2015c. Journals' endorsement of NIH guidelines. National Institutes of Health, Bethesda, MD. <http://www.nih.gov/sites/default/files/research-training/initiatives/reproducibility/rigor-reproducibility-endorsements.pdf>.
- Open Science Collaboration. 2015. Psychology. Estimating the reproducibility of psychological science. *Science* 349(6251): aac4716. <http://doi.org/10.1126/science.aac4716>.
- Pickett CL, Corb BW, Matthews CR, Sundquist WI, and Berg JM. 2015. Toward a sustainable biomedical research enterprise: finding consensus and implementing recommendations. *Proc Natl Acad Sci U S A* 112(35):10832–10836. <http://doi.org/10.1073/pnas.1509901112>.
- Reardon S. 2015. US vaccine researcher sentenced to prison for fraud. *Nature* 523(7559):138–139. <http://doi.org/10.1038/nature.2015>.
- Stemwedel JD. 2014. Life after misconduct: promoting rehabilitation while minimizing damage. *J Microbiol Biol Educ* 15(2):177–180. <http://doi.org/10.1128/jmbe.v15i2.827>.
- Stern AM, Casadevall A, Steen RG, and Fang FC. 2014. Financial costs and personal consequences of research misconduct resulting in retracted publications. *eLife* 3:e02956. <http://doi.org/10.7554/eLife.02956>.
- Strasak AM, Zaman Q, Pfeiffer KP, Göbel G, and Ulmer H. 2007. Statistical errors in medical research—a review of common pitfalls. *Swiss Med Wkly* 137(3/4):44. <http://www.smw.ch/for-readers/archive/backlinks?url=/docs/archive200x/2007/03/smw-11587.html>.
- Van Noorden R. 3 June 2015. Sluggish data sharing hampers reproducibility effort. *Nature*. <http://doi.org/10.1038/nature.2015.17694>. <http://www.nature.com/news/sluggish-data-sharing-hampers-reproducibility-effort-1.17694>. Accessed 20 November 2015.
- Wager E. December 2014. How should journal editors respond to cases of suspected misconduct? *J Microbiol Biol Educ* 15(12):146–150. <http://doi.org/10.1128/jmbe.v15i2.829>.
- Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, Berelowitz M, Dhillon AP, Thomson MA, Harvey P, Valentine A, Davies SE, and Walker-Smith JA. 1998. Retracted: ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet* 351(9103):637–641.
- Ware M, and Mabe M. 2012. The STM report: an overview of scientific and scholarly journal publishing, 3rd ed. International Association of Scientific, Technical and Medical Publishers, The Hague, Netherlands. http://www.stm-assoc.org/2012_12_11_STM_Report_2012.pdf.
- Weissgerber TL, Garovic VD, Milin-Lazovic JS, Winham SJ, Obradovic Z, Trzeciakowski JP, and Milic NM. 2016. Reinventing biostatistics education for basic scientists. *PLoS Biol* 14(4): e1002430. <http://doi.org/10.1371/journal.pbio.1002430>.
- Winey M. 15 July 2015. How Can Scientists Enhance Rigor in Conducting Basic Research and Reporting Research Results? American Society for Cell Biology, Bethesda, MD. <http://www.ascb.org/wp-content/uploads/2015/11/How-can-scientist-enhance-rigor.pdf>.
- Womack RP. 2015. Research data in core journals in biology, chemistry, mathematics and physics. *PLoS One* 10(12): e0143460. <http://doi.org/10.1371/journal.pone.0143460>.
- Worthy G. 2015. Statistical analysis and reporting: common errors found during peer review and how to avoid them. *Swiss Med Wkly* 145: w14076. <http://doi.org/10.4414/smw.2015.14076>.
- Yong E. 27 August 2015. How reliable are psychology studies? Atlantic, Washington, DC. <http://www.theatlantic.com/science/archive/2015/08/psychology-studies-reliability-reproducibility-no-sek/402466/>.

Appendix A

Definitions

The U.S. Department of Health and Human Services defines scientific misconduct as “fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results.”

Reproducibility—for any experimental research, an independent investigator should be able to replicate the original experiment, under the same conditions, and achieve similar or comparable results.

Definitions, American Society for Cell Biology Task Force

- **Analytic replication**—attempts to reproduce the results from the same original data via reanalysis
- **Direct replication**—attempts to reproduce the same results using the same conditions and methods as the original experiment
- **Systematic replication**—aims at obtaining the same finding of a given publication, but under different conditions, for example, with a different cell line, mouse strain, etc.
- **Conceptual replication**—aims to demonstrate the validity of a concept or a finding using a different paradigm

Appendix B

Statement of Task

The American Academy of Microbiology will host a colloquium to discuss scientific data irreproducibility and how to prevent these practices from occurring among the biological sciences. Any alteration made to original data is considered scientific misconduct and can result in strong action, possibly leading to the demise of one's

career and significant financial costs. The U.S. government defines scientific misconduct as “fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results.” Publishing novel findings is a difficult task, and with fierce competition among researchers to be the first group to relay the data, modifying data, unfortunately, can become appealing.

With the increased use of computational and imaging methods in biology, the manipulation of scientific data that appear in the literature has become a significant issue. For example, with imaging software, it is now very simple and can be tempting to make adjustments to digital image files rather than repeating the experiment to produce a better image. Data irreproducibility continues to be a problem for a multitude of reasons that this colloquium will address. The intense competition to obtain funding opportunities, cultural differences in laboratory practices, and general naivety of students, laboratory technicians, postdocs, and principal investigators can result in incorrect scientific reporting.

Appendix C

Colloquium Discussion Questions

1. Why has the ability to reproduce biological scientific data become a significant issue?
 - 1a. How do institutions contribute to the problem of data irreproducibility?
 - 1b. What current research practices and journal policies contribute to irreproducibility?
 - 1c. Do cultural differences and the meaning of ethical behavior influence scientific practices?
 - 1d. Does the current culture of science favor high impact over carefully performed science?



2. What constitutes data reproducibility?
 - 2a. What percentage of irreproducibility is caused by (i) uncontrolled variables, (ii) sloppy science (i.e., poor stats, cell line contamination), (iii) bias, and (iv) scientific misconduct?
 - 2b. Are there different accepted levels of data reproducibility among the biological sciences (e.g., cancer research)?
3. How can data irreproducibility be addressed at researcher and institutional levels before the problem escalates?
 - 3a. What can be done to enhance transparency in reporting scientific data?
 - 3b. Is competition beneficial or detrimental to science? Can competition be eliminated from science?
 - 3c. How can high-quality work be rewarded in a scientific culture suffering from IF mania?
4. Do poor research practices affect the public's and research community's views of scientific research?
 - 4a. Are there problems with the incentive structure of science?
 - 4b. Are there viable alternative economic systems that would enhance cooperation and creativity and reduce the debits of competition?
 - 4c. How do funding agencies "reward" solid, well-performed studies rather than flashy science?
 - 4d. How can they hold the researchers that they fund responsible?
5. What are the ramifications of unethical practices?
 - 5a. Who is affected by scientific misreporting?
 - 5b. What is the opportunity cost associated with irreproducible studies?
 - 5c. How do retractions impact the reputation of a journal?
6. Difficulty in publishing negative results may prompt selective reporting of data. How can the publication of negative results be facilitated?
7. How can institutions and journals enhance the transparency and robustness of scientific reporting?
 - 7a. Should training courses in scientific integrity and good practices be mandatory for both trainees and experienced scientists?
 - 7b. Should journals use software to monitor image manipulation? What are the costs of instituting these measures?
 - 7c. How can whistleblowers report scientific misconduct at the institutional level?
 - 7d. Should there be a direct line to the ORI?
 - 7e. How do we address this in other nations lacking an ORI?
 - 7f. Can PubPeer and PubMed Commons improve scientific integrity?
8. How can journals ensure complete compliance with data set disclosure and software sharing?
9. How is data reproducibility presently addressed?
 - 9a. What will be the impact of the initiative implemented by the NIH on data misrepresentation?
 - 9b. Would tougher penalties for research misconduct be helpful?
 - 9c. How can data standards be raised?

American Academy of Microbiology
1752 N Street, NW
Washington, DC 20036

academy.asm.org

