# Reproducibility of Research
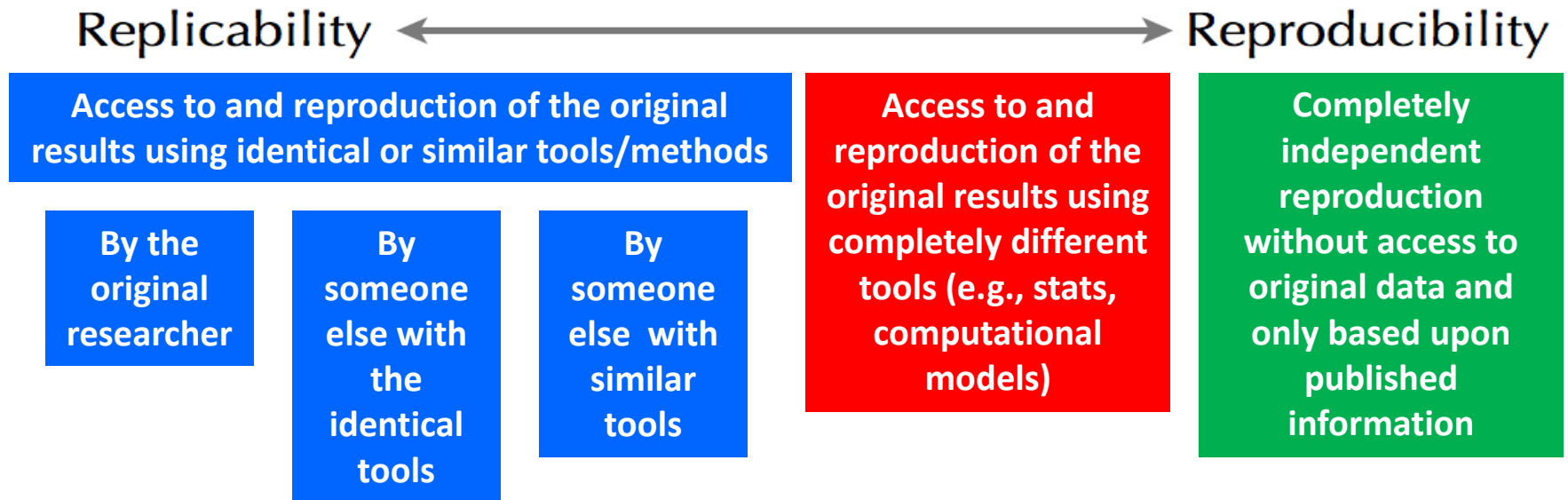


APLU Council on Research Summer Meeting
June 29, Park City, UT

- Reproducibility, replicability, repeatability, rigor, transparency, independent verification...are all foundational to the **Scientific Method**

- These terms are not synonymous but constitute a **continuum** (which varies by discipline)

Replicability ←——————————————→ Reproducibility

| Access to and reproduction of the original results using identical or similar tools/methods | Access to and reproduction of the original results using completely different tools (e.g., stats, computational models) | Completely independent reproduction without access to original data and only based upon published information |
|---|---|---|
| By the original researcher / By someone else with the identical tools / By someone else with similar tools | | |

# Why Now?

- The reproducibility of research has become a topic of great interest in the past few years, driven in part by
  - Complexity of today's research **tools and problems** being studied across all disciplines
  - **Inability to reproduce results** of some studies
  - Concerns about **rushing to publication** with poor experiment design or misuse of statistics
  - Greatly increased emphasis by **Congress** on transparency, accountability, and return on investment
  - New **open data** requirements that will make available vast new quantities of data/outcomes for verification
  - Increasing use of highly complex **computational models** that contain inherent elements of chaos/uncertainty

# Points of View

- Although everyone agrees that reproducibility is foundational to the Scientific Method, the **motivation for change** and understanding of the issues varies
  - Researchers
  - Professional associations
  - Journals (print and online)
  - Provosts
  - Senior Research Officers
  - Lawmakers
  - Funding agencies and other research sponsors
  - Inspectors General

# Some of the Issues/Questions...

- Reproducibility does not imply **correctness**

- Lack of reproducibility does not imply **incorrectness**

- Only some disciplines have **formal methods courses** or teach about the Scientific Method/experiment design.

- It is normal for some published results to be **refuted**; what role does reproducibility play?  Does a gap exist between scientific **values** and scientific **practices**?

- At **what stage** should reproducibility be addressed?
  - Proposal, peer review, experimentation, journal submission, peer review?

- Journal articles usually don't provide **sufficient information** to allow for reproducibility

- Who bears the **cost** of refuting claims that results are not reproducible?

# 2015 America COMPETES Act (NSF Part)

## SEC. 117. RESEARCH REPRODUCIBILITY AND REPLICATION.

(a) SENSE OF CONGRESS.—It is the sense of Congress that—

    (1) the gold standard of good science is the ability of a researcher or research lab to reproduce a published method and finding;

    (2) there is growing concern that some published research findings cannot be reproduced or replicated, which can negatively affect the public's trust in science;

    (3) there are a complex set of factors affecting reproducibility and replication; and

    (4) the increasing interdisciplinary nature and complexity of scientific research may be a contributing factor to issues with research reproducibility and replication.

(b) REPORT.—The Director shall—

    (1) not later than 45 days after the date of enactment of this Act, enter into an agreement with the National Research Council to provide, within 18 months after the date of enactment of this Act, a report to assess research and data reproducibility and replicability issues in interdisciplinary research and to make recommendations on how to improve rigor and transparency in scientific research; and

    (2) not later than 60 days after receiving the results of the assessment under paragraph (1), submit a report to the Committee on Science, Space, and Technology of the House of Representatives and the Committee on Commerce, Science, and Transportation of the Senate on the findings of the assessment, together with the agreement or disagreement of the Director and Board with each of its findings and recommendations.

# National Science Foundation

- NSF is beginning a **wide-ranging discussion** among stakeholders of the quality and utility of the results of the research it funds
  - Emphasis is on **"reliability"** as encompassing reproducibility, replicability, robustness, etc
- A continuing theme for NSF is creating the most **transparent** processes possible, and this effort is part of that theme
- Directorates will be initiating, either individually or collectively, **workshops** and other activities that can stimulate an energetic and informed discussion
  - One workshop will take place in DC on **September 10**

# National Institutes of Health



NIH | National Institutes of Health
Turning Discovery Into Health

Search

Health Information | Grants & Funding | News & Events | Research & Training | Institutes at NIH | About NIH

NIH Home > Research & Training

## RIGOR AND REPRODUCIBILITY

Rigor and Reproducibility

- Principles and Guidelines
- Publications
- Training
- Meetings and Workshops
- Expanded Guidelines
- Application Instructions

### Rigor and Reproducibility

Two of the cornerstones of science advancement are rigor in designing and performing scientific research and the ability to reproduce biomedical research findings. The application of rigor ensures robust and unbiased experimental design, methodology, analysis, interpretation, and reporting of results. When a result can be reproduced by multiple scientists, it validates the original results and readiness to progress to the next phase of research. This is especially important for clinical trials in humans, which are built on studies that have demonstrated a particular effect or outcome.

In recent years, however, there has been a growing awareness of the need for rigorously designed published preclinical studies, to ensure that such studies can be reproduced. This webpage provides information about the efforts underway by NIH to enhance rigor and reproducibility in scientific research.

Johns Hopkins University students in a laboratory. (Johns Hopkins University Photo)
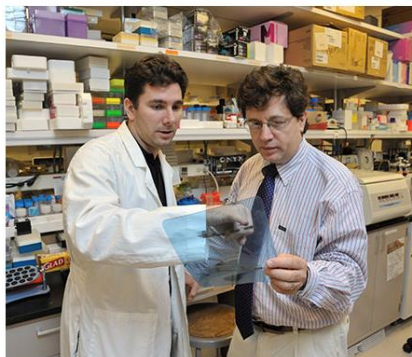
Richard Siegel, M.D., Ph.D., (right) NIAMS Clinical Director and Chief of the Autoimmunity Branch, discusses research results with postdoctoral fellow Martin Pelletier, Ph.D. (Bill Branson, photographer)

#### Email Updates

To sign up for updates please enter your e-mail address.

Submit

#### Related Links

Letter from Dr. Stephen I. Katz: An Update on the NIH Initiative to Enhance Research Rigor and Reproducibility

#### Contact Us

Please send email to NIHReprodEfforts@od.nih.gov.

## Principles and Guidelines for Reporting Preclinical Research

NIH held a joint workshop in June 2014 with the Nature Publishing Group and Science on the issue of reproducibility and rigor of research findings, with journal editors representing over 30 basic/preclinical science journals in which NIH-funded investigators have most often published. The workshop focused on identifying the common opportunities in the scientific publishing arena to enhance rigor and further support research that is reproducible, robust, and transparent.

The journal editors came to consensus on a set of principles to facilitate these goals, which a considerable number of journals have agreed to endorse (see Endorsements — Principles and Guidelines for Reporting Preclinical Research ⬇ (PDF - 75KB)). These principles are shown below.

[Expand All]

⊞ Rigorous statistical analysis

⊞ Transparency in reporting

⊞ Data and material sharing

⊞ Consideration of refutations

⊞ Consider establishing best practice guidelines for:

⊞ Endorsements — Principles and Guidelines for Reporting Preclinical Research

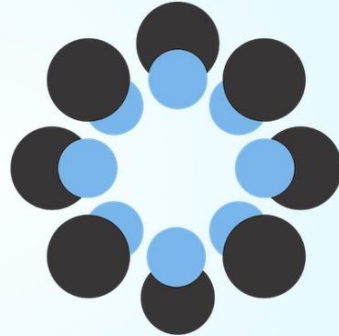# Community-Initiated Non-Profit



http://centerforopenscience.org

# Corporate For-Profit



http://validation.scienceexchange.com/#/

# Courses Now Being Offered



https://www.coursera.org/course/repdata

# Computer Code

# Open science decoded

Tony Hey and Mike C. Payne

Granting access to publications and data may be a step towards open science, but it's not enough to ensure reproducibility. Making computer code available is also necessary — but the emphasis must be on the quality of the programming.

In February 2013, the US Office of Science and Technology Policy, in the Executive Office of the President, issued a memorandum requiring that Federal agencies investing in research develop clear policies to support increased public access to the results of their research (http://go.nature.com/OhZNcR). The memo stipulated that "such results include peer-reviewed publications and digital data", where digital data is defined as material necessary to validate other scientists' research. It was soon followed by similar declarations from the Global Research Council in May (http://go.nature.com/qHIgxU) and from the G8 Science Ministers in June 2013 (http://go.nature.com/TeKjLs).

Clearly, there is now increasing global momentum towards 'open science'. This necessarily requires not only open access to research publications, but also to the metadata and data required to validate and make sense of the results of research. And improving the comprehensibility and reproducibility of computational science is an important step in this endeavour. One approach is through executable papers, which not only provide the text, tables and figures of a conventional paper, but also offer access to the software, data and computer environment used to produce the results. There are now several interesting attempts at providing support for executable papers[1,2]. However, there are still many research challenges in this area and it remains to be seen whether such approaches will become widely adopted.

One might argue that the physics community is leading the charge in response to these obstacles. In the 1980s, forty years after the birth of the electronic computer, Nobel Prize winner Ken Wilson and others argued for the importance of computational science for scientific exploration and discovery. Wilson even went so far as to call computational science the "third paradigm
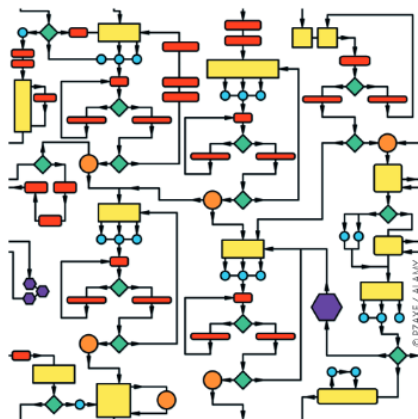
of science, supplementing theory and experimentation"[3]. And indeed, computer simulations are now used effectively to gain understanding of complicated physical systems that are either too complex to be solved analytically or inaccessible to experiment — or both.



**Reproducible research**

Computational science presents a challenge to the traditional notion of scientific reproducibility. Several different programs written by different researchers can seek to explore the physics of the same complex system and these may use different algorithms and/or different numerical approximations. Exact numerical agreement between the results of the two different programs is therefore not to be expected, and in practice, reproducibility involves finding very similar quantitative results for key physical parameters. However, all too often, neither the source code nor enough detail of the precise computing environment is included with the published paper to enable other scientists to reproduce the results. In addition, of course, access to the

particular high-end supercomputer that obtained the results of the simulation may not be possible. These issues form part of the challenge of 'reproducible research' for computational science[4].

One of the first scientists to recognize the need for reproducibility in computational science was the geophysicist Jon Claerbout. As early as 1990, he set a goal of reproducibility for all the (non-open-access) reports coming out of his Stanford Exploration Project, identifying reproducibility as "a way of organizing computational research that allows both the author and the reader of a publication to verify the reported results" (http://go.nature.com/FMw34h).

Twenty years later, in December 2012, the Institute for Computational and Experimental Research in Mathematics at Brown University hosted a workshop on reproducibility in computational and experimental mathematics. Out of this workshop came a list of the five best publication practices (http://go.nature.com/wtqOLn). The list declares that data and code should both be made available and accessible, in the sense that others can reproduce published findings. It underlines the importance of properly citing data that is not generated exclusively for the study in question. And it stipulates guidelines for granting publishers copyright permissions, and for organizing supplementary materials.

Following such best-practice guidelines in principle gives researchers the information needed to reproduce computational results using the same code and data used by the original researcher. However, there is a related but distinct problem of numerical reproducibility[4]. Round-off error can be exacerbated when simulations are scaled up to run on large parallel systems. This can be problematic for the original authors of the code — and even more so for researchers seeking to reproduce results

© PZAXE / ALAMY

367

# Key Questions for Today

- To what extent do your **faculty understand and appreciate the importance of reproducibility**, the strong and increasing emphasis being placed upon it nationally, and the importance of making sure students in relevant disciplines understand the issue?  What role might your office play in building this appreciation?

- To what extent might **research misconduct** be an important contributor to problems of irreproducibility if it occurs on your campus?  What changes might be made to the Research Integrity Office that would lead to clear improvement in the reproducibility of results?

# Key Questions for Today

- To what extent is **poor research design or improper analysis** a potential contributor to irreproducibility on your campus?  Does your campus offer centralized support or services for assisting faculty with research design, statistical analyses, data visualization, etc?

- To what extent is **poor oversight and management** of students, postdocs, and staff a potential contributor to irreproducibility?  How do faculty on your campus acquire this expertise?

# Key Questions for Today

- There is a strong preference, on the part of journals, for researchers to **report "significant" results**.  Because of the strong link between publication and success with federal grants, and because of the strong link between both and career advancement, faculty are at risk of interpreting and reporting on the results of their research in a biased way.  This bias may be overt, but it may also be quite subtle.  Has your university addressed this issue?

- What is the **role of the senior research officer** in ensuring research quality?  What sort of training programs should be offered, and to whom?  Should there be QA/QC beyond or in addition to peer review?

# Key Questions for Today

- Mandated **open access policies** for data, and in many cases computer code, may lead to new challenges in reproducibility, including substantial costs for re-doing analyses or experiments for which funding has ended.  Has your institution given thought to this issue and how such work might be funded?

- Have you experienced situations on your campus where research results have been openly or quietly **challenged** owing to real or perceived irreproducibility?  If so, what was the resolution?

- Congress is weighing in on reproducibility (e.g., the 2015 America COMPETES Act from the House Science, Space and Technology Committee), suggesting that it is rampant across all fields of science and engineering.  **How can we and COR be helpful** in clarifying the issues and working with Congress on positive steps forward?