Advancing Gold Standard Science to Address America's Chronic Disease Crisis

Council of Councils
May 14, 2026

Jay Bhattacharya, M.D., Ph.D.
Director, National Institutes of Health

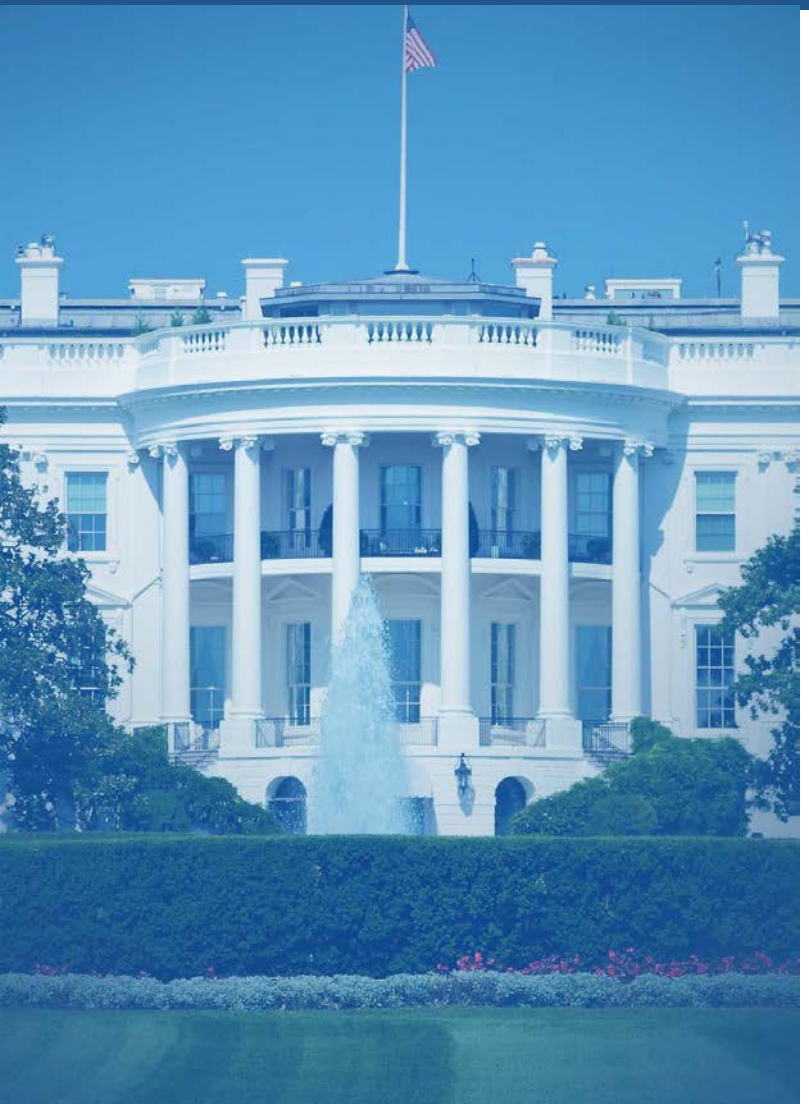
No Improvement in American Longevity for 15 years



US life expectancy flatlined between 2010 and 2019 and collapsed during the pandemic

Source: Mortality Watch
https://www.mortality.watch/explorer?c=USA&t=le&ct=yearly&df=2001&bf=2011&sb=0&ce=1&m=1&pi=0

Make America Healthy Again



Chronic diseases such as heart disease, cancer, and diabetes are:

- Leading causes of death and disability in the U.S.
- Leading drivers of the nation's $4.5 trillion in annual health care costs

NIH's role: Find better ways to prevent, treat, and cure chronic diseases.

Source: CDC

To make America healthy again, we must rethink how we do science

Why Aren’t We Making More Progress?

Too much **safe, incremental science**

Research doesn’t **replicate**

Funding concentrated in few institutions

Solution: Democratizing Science

Too much **safe, incremental science**

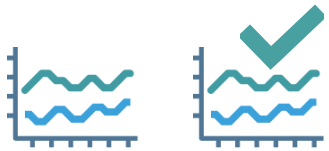




Unified Funding Strategy

Research doesn't **replicate**





Replication initiative

Funding concentrated in few institutions





Competitive infrastructure support

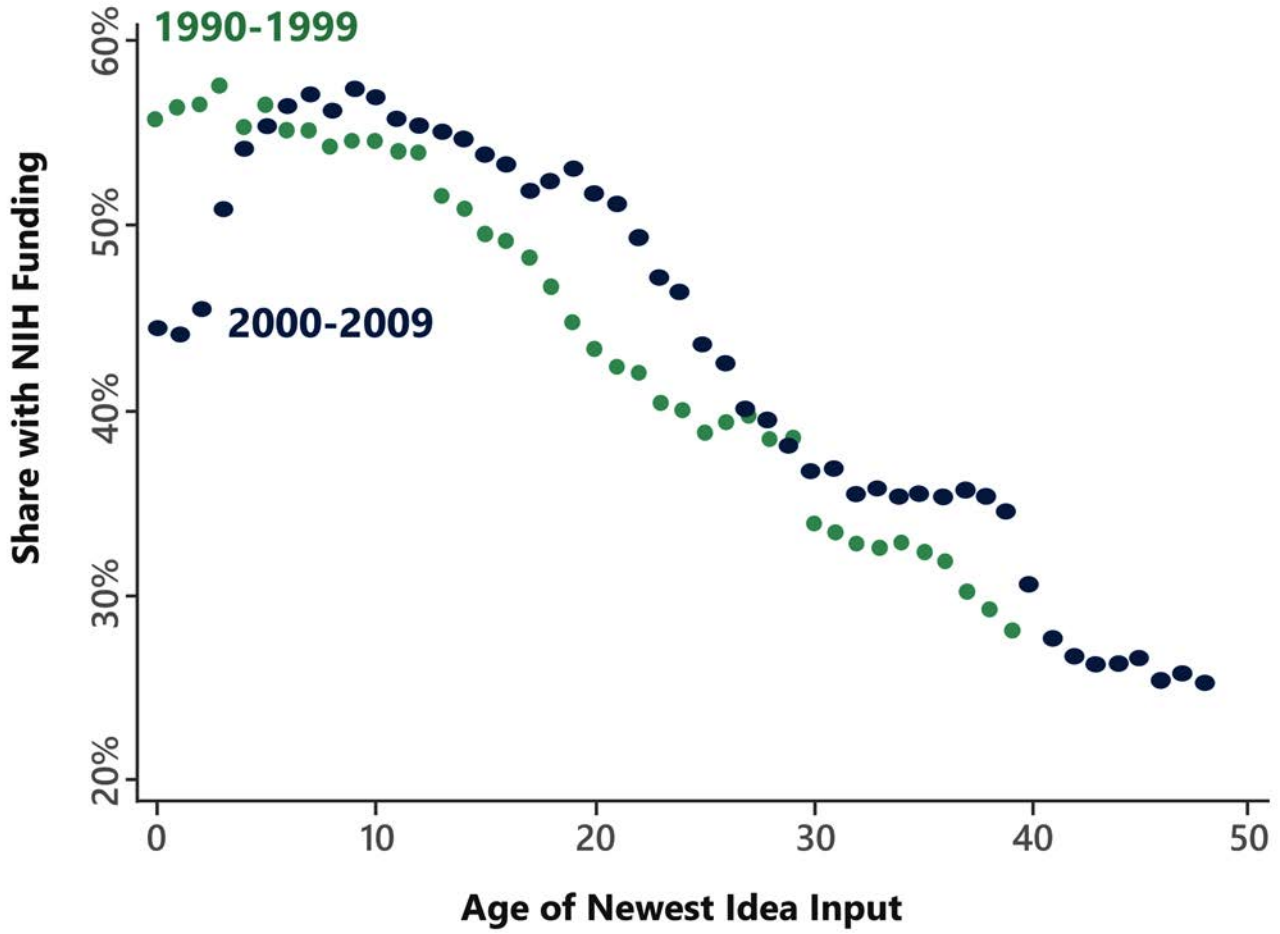
Funding Bold Ideas

Unified Funding Strategy

- End pay-lines for selecting projects to fund
- Peer review important + other key factors

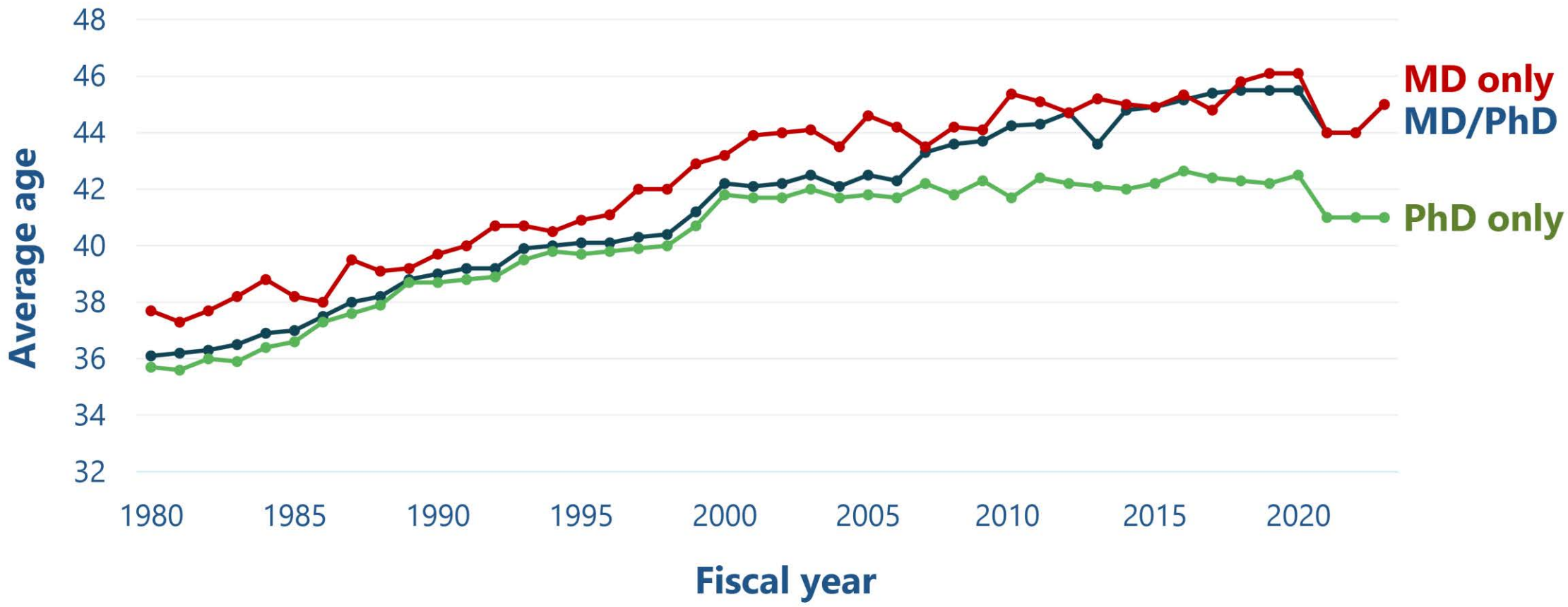
- **Goals:**
  - Gives NIH Institutes and Centers flexibility to fund high-risk, high-reward research
  - Align funding with institute missions
  - Prioritize early-career scientists and novel ideas

Supporting Edge Science



Packalen and Bhattacharya (2020) NIH Funding and the Pursuit of Edge Science. PNAS

Average age of first-time R01 (1980-2023)

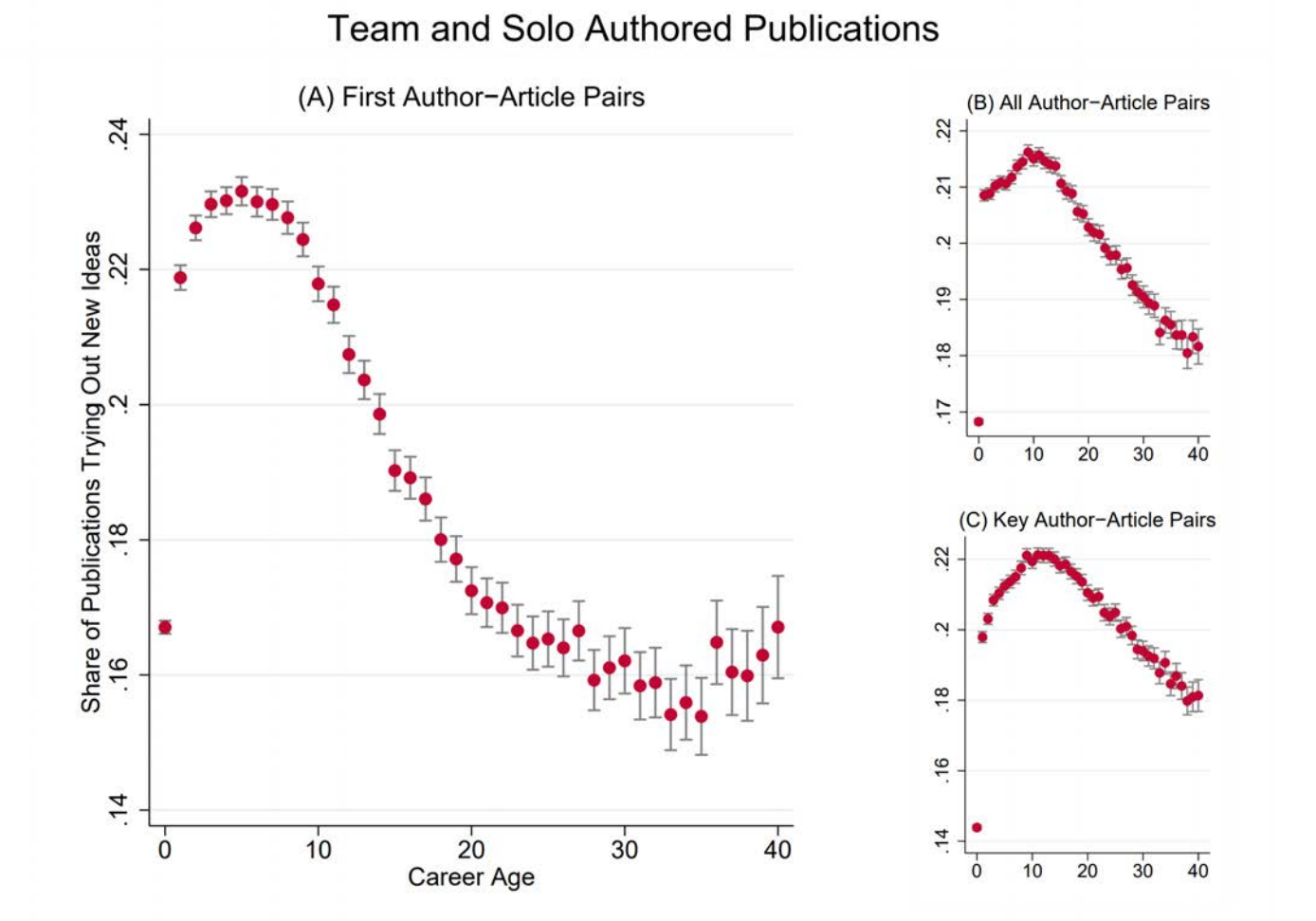


https://grants.nih.gov/policy-and-compliance/policy-topics/early-stage-investigators/related-data

Garrison HH, Ley TJ. Physician-scientists in the United States at 2020: Trends and concerns. FASEB J. 2022;36(5):e22253. doi:10.1096/fj.202200327

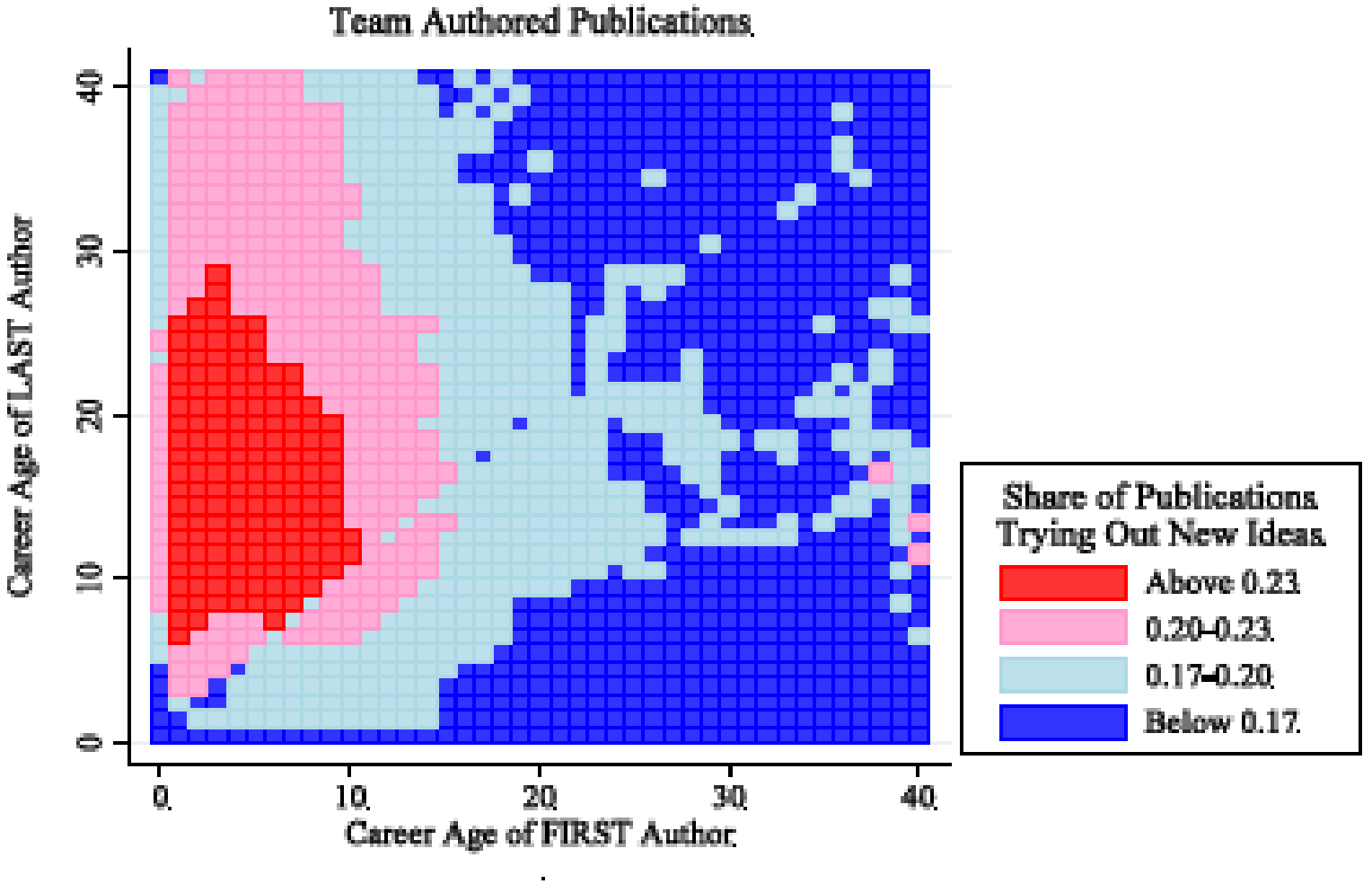
https://web.archive.org/web/20250404233158/https:/nexus.od.nih.gov/all/2024/07/03/continued-support-for-early-stage-investigators-in-fy-2023/

Early-Career Researchers Contribute a Larger Share of New Ideas



Packalen and Bhattacharya (2019) "Age and the Trying Out of New Ideas." Journal of Human Capital.

Teams of Early Career and Later Career Researchers Most Likely to Try Out New Ideas



Solving the Replication Crisis

Replication as the Basis for Truth in Science



Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

Published research findings are sometimes refuted by subsequent evidence, with ensuing confusion and disappointment. Refutation and controversy is seen across the range of research designs, from clinical trials and traditional epidemiological studies [1–3] to the most modern molecular research [4,5]. There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims [6–8]. However, this should not be surprising. It can be proven that most claimed research findings are false. Here I will examine the key factors that influence this problem and some corollaries thereof.

The Essay section contains opinion pieces on topics of broad interest to a general medical audience.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p-value less than 0.05. Research is not most appropriately represented and summarized by p-values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

should be interpreted based only on p-values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful. "Negative" is actually a misnomer, and the misinterpretation is widespread. However, here we will target relationships that investigators claim exist, rather than null findings.

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance [10,11]. Consider a 2×2 table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let R be the ratio of the number of "true relationships" to "no relationships" among those tested in the field. R is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α. Assuming that c relationships are being probed in the field, the expected values of the 2×2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the 2×2 table, one gets $PPV = (1 - \beta)R/(R - \beta R + \alpha)$. A research finding is thus

Citation: Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2(8): e124.

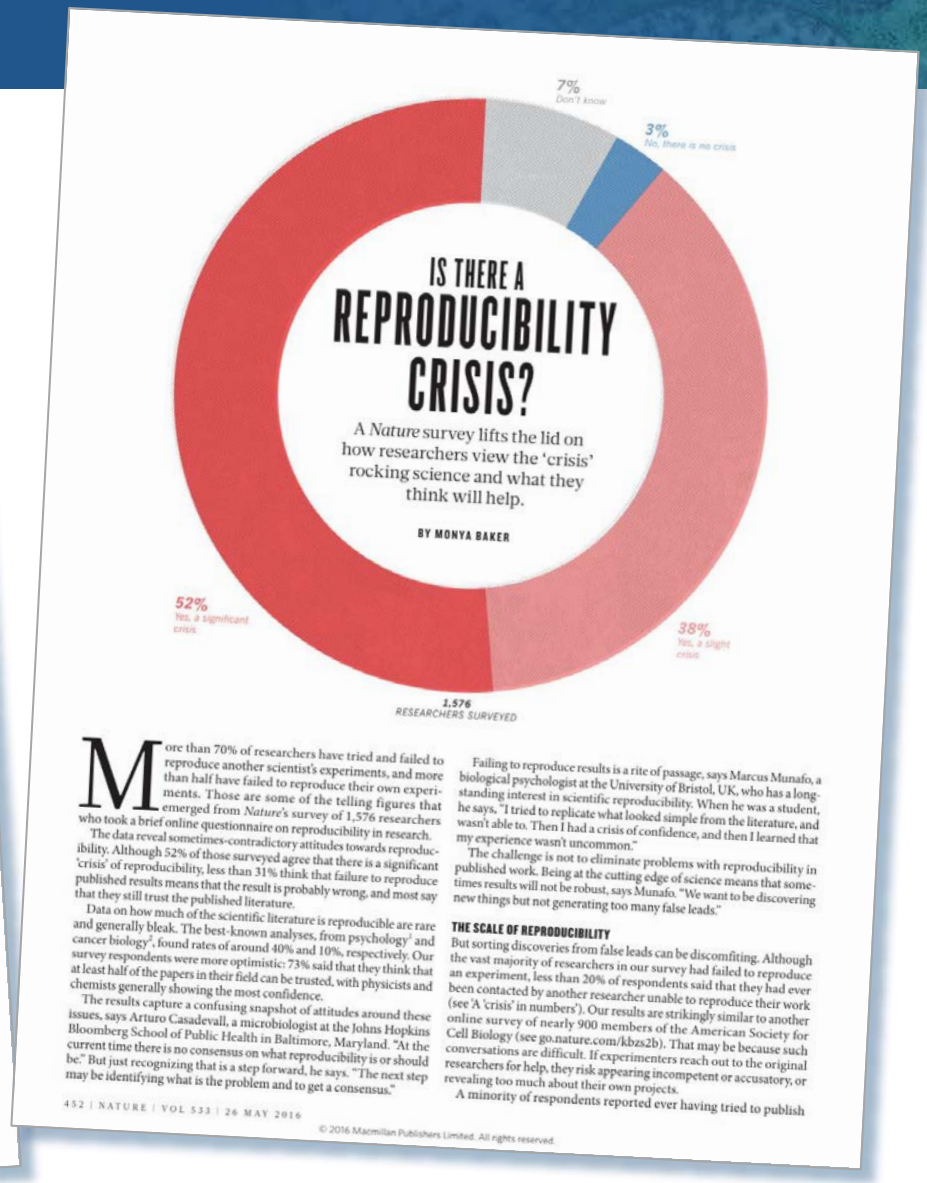
Copyright: © 2005 John P. A. Ioannidis. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abbreviation: PPV, positive predictive value

John P. A. Ioannidis is in the Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, and Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: jioannid@cc.uoi.gr

Competing Interests: The author has declared that no competing interests exist.

DOI: 10.1371/journal.pmed.0020124



IS THERE A REPRODUCIBILITY CRISIS?

A *Nature* survey lifts the lid on how researchers view the 'crisis' rocking science and what they think will help.

BY MONYA BAKER

More than 70% of researchers have tried and failed to reproduce another scientist's experiments, and more than half have failed to reproduce their own experiments. Those are some of the telling figures that emerged from *Nature*'s survey of 1,576 researchers who took a brief online questionnaire on reproducibility in research.

The data reveal sometimes-contradictory attitudes towards reproducibility. Although 52% of those surveyed agree that there is a significant 'crisis' of reproducibility, less than 31% think that failure to reproduce published results means that the result is probably wrong, and most say that they still trust the published literature.

Data on how much of the scientific literature is reproducible are rare and generally bleak. The best-known analyses, from psychology[2] and cancer biology[3], found rates of around 40% and 10%, respectively. Our survey respondents were more optimistic: 73% said that they think that at least half of the papers in their field can be trusted, with physicists and chemists generally showing the most confidence.

The results capture a confusing snapshot of attitudes around these issues, says Arturo Casadevall, a microbiologist at the Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland. "At the current time there is no consensus on what reproducibility is or should be." But just recognizing that is a step forward, he says. "The next step may be identifying what is the problem and to get a consensus."

Failing to reproduce results is a rite of passage, says Marcus Munafo, a biological psychologist at the University of Bristol, UK, who has a long-standing interest in scientific reproducibility. When he was a student, he says, "I tried to replicate what looked simple from the literature, and wasn't able to. Then I had a crisis of confidence, and then I learned that my experience wasn't uncommon."

The challenge is not to eliminate problems with reproducibility in published work. Being at the cutting edge of science means that sometimes results will not be robust, says Munafo. "We want to be discovering new things but not generating too many false leads."

THE SCALE OF REPRODUCIBILITY

But sorting discoveries from false leads can be discomfiting. Although the vast majority of researchers in our survey had failed to reproduce an experiment, less than 20% of respondents said that they had ever been contacted by another researcher unable to reproduce their work (see 'A 'crisis' in numbers'). Our results are strikingly similar to another online survey of nearly 900 members of the American Society for Cell Biology (see go.nature.com/kbzs2b). That may be because such conversations are difficult. If experimenters reach out to the original researchers for help, they risk appearing incompetent or accusatory, or revealing too much about their own projects.

A minority of respondents reported ever having tried to publish

© 2016 Macmillan Publishers Limited. All rights reserved.

Sources:
Baker, M. ***1,500 scientists lift the lid on reproducibility.*** *Nature 533, 452–454 (2016). DOI: 10.1038/533452a*
Ioannidis JPA (2005) ***Why most published research findings are false.*** *PLoS Med 2(8): e124.*

Replication Crisis: Empirical Basis

nature

Explore content ~ About the journal ~ Publish with us ~ Subscribe

nature > news > article

NEWS | 25 April 2025

Huge reproducibility project fails to validate dozens of biomedical studies

Unique reproducibility effort in Brazil focuses on common methods rather than a field — and prompts call for reform.

By Rodrigo de Oliveira Andrade

Falling short

The authors judged a paper's replicability by five criteria, including whether at least half of the replication attempts had statistically significant results in the same direction as the original paper. Only 21% of the experiments were replicable using at least half of the applicable criteria.

The authors also found that the effect size – the magnitude of the observed impact in the experiments – was, on average, 60% larger in the original papers than in the experimental follow-ups, indicating that published results tend to overestimate the effects of the interventions tested.

Invest in the Science of Replication



Fund replication research



Establish a replication platform

PubMed

Enhanced visibility through PubMed



New metrics for scientific contribution


Funding Scientific
Excellence Everywhere

More Equitable Distribution of Funds

Current system concentrates resources at a few institutions

Research project support

Research grants



Research infrastructure support

Infrastructure funding



Thank you



National Institutes of Health

Turning Discovery Into Health

What’s Next

Closing the Gap Between Health Data and Health Answers

U.S. health data distributed over too many systems

- Fragmented and difficult to use quickly, consistently
- Slow answers to national health questions
- Inconsistent evidence
- High cost of data collection

NIH Real-World Data Network

Accelerating discovery



Gold Standard Science: A Vision for Transforming Biomedical Research

Purpose: Modernizing how NIH funds science to drive innovation, rigor, and fairness.

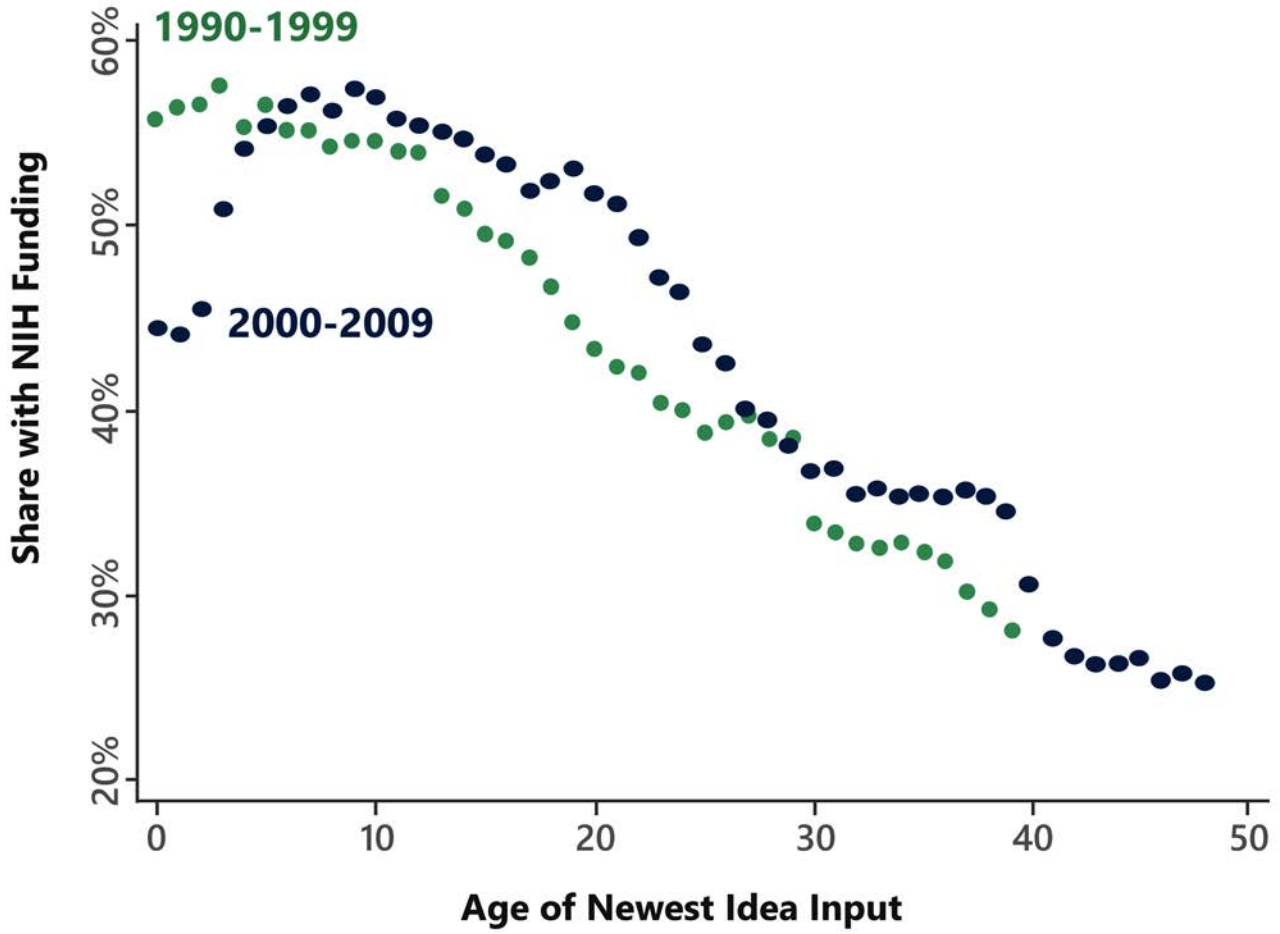


Funding Bold Ideas: Unified Funding Strategy

- End pay-lines for selecting projects to fund
- Peer review important + other key factors

- **Goals:**
  - Gives NIH Institutes and Centers flexibility to fund high-risk, high-reward research
  - Align funding with institute missions
  - Prioritize early-career scientists and novel ideas

Change Over Time: 1990-2009



Packalen and Bhattacharya (2020) NIH Funding and the Pursuit of Edge Science. PNAS

Core Principles of Research Reliability

Reproducibility

Can the same result be produced using the same data and methods?

Example: Another team runs the same code on the same dataset and gets the same findings.

Replicability

Does the result hold up when the study is done again?

Example: A different lab repeats the experiment (new data, similar methods) and finds similar results.

Generalizability

Does the finding apply beyond the original study?

Example: A treatment works not just in one group, but across ages, regions, or environments.



Supporting Early Career Investigators

Early Career Researchers are a Primary Source for New Ideas in Biomedicine

THE UNIVERSITY OF CHICAGO PRESS JOURNALS

This journal | Search | Advanced Search | Log In | Register

Access provided by National Institutes of Health - United Stat

Subscribe/renew | Browse issues | Forthcoming | Contributors | About

Journal of Human Capital > Volume 13, Number 2

PREVIOUS ARTICLE

Age and the Trying Out of New Ideas

Mikko Packalen and Jay Bhattacharya

PDF | PDF PLUS | Abstract | Full Text | Supplemental Material

Details | Figures | References | Cited by



Journal of Human Capital
Volume 13, Number 2
Summer 2019

Abstract

The aging of the scientific workforce and graying of grant recipients are central concerns in science policy. These trends are potentially important because older scientists are often seen as less open to new ideas than younger scientists. Here we put this hypothesis to an empirical test. A text analysis of 20 million biomedical research articles shows that papers published by younger researchers are more likely to build on new ideas. Collaboration with an experienced researcher also matters. Papers with a young first author and a more experienced last author are more likely to try out newer ideas than papers published by other team configurations. Our results buttress the case for funding young researchers but also provide a caution against unconditional idolatry of youth over experience.

Supporting Early-Career Investigators to Spur Innovation

- We need a **culture shift** that benefits early-stage investigators
- We are considering innovative approaches, such as:
  - Use the number of postdocs moving onto independence as a metric of success for large grants
  - Expand institutional K awards: allow universities to compete for K award slots so they have an existing position and can quickly hire an investigator to avoid delays

Many of Today's Treatments are Built on NIH-Supported Research

NIH-supported research contributes to the scientific foundation behind nearly all FDA-approved drugs, especially through basic research on drug targets.



~100% of FDA-approved drugs (2010–2019)

Source: NIH funding analyses of FDA-approved drugs, 2010–2019 (e.g., Cleary et al., PNAS 2018; follow-on studies): Comparison of Research Spending on New Drug Approvals by the National Institutes of Health vs the Pharmaceutical Industry, 2010-2019 | Health Policy | JAMA Health Forum | JAMA Network