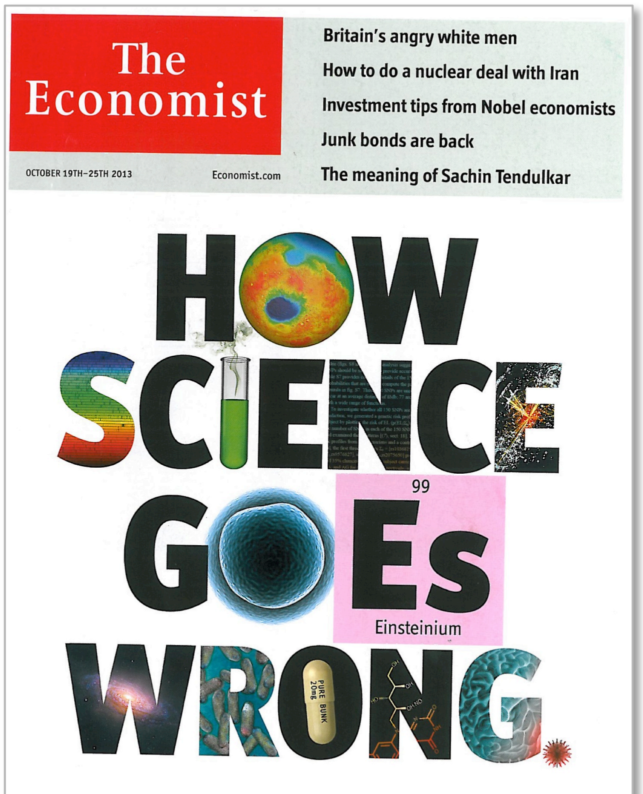
The challenge for all of us

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

PloS Medicine 2005 doi: 10.1371/journal.pmed.0020124

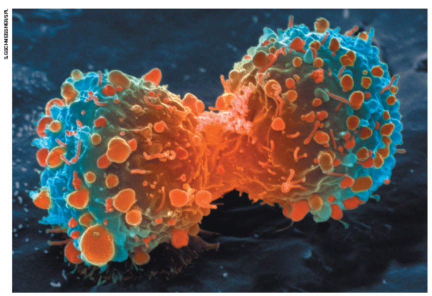


HEALTH

The Human Cost of a Misleading Drug-Safety Study

A reexamination of old data for Paxil found that the antidepressant is more dangerous than the authors let on. How much harm has been done in the 14 years since it was published?

David Dobbs, The Atlantic, 18th SEPT 2015



Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Raise standards for preclinical cancer research

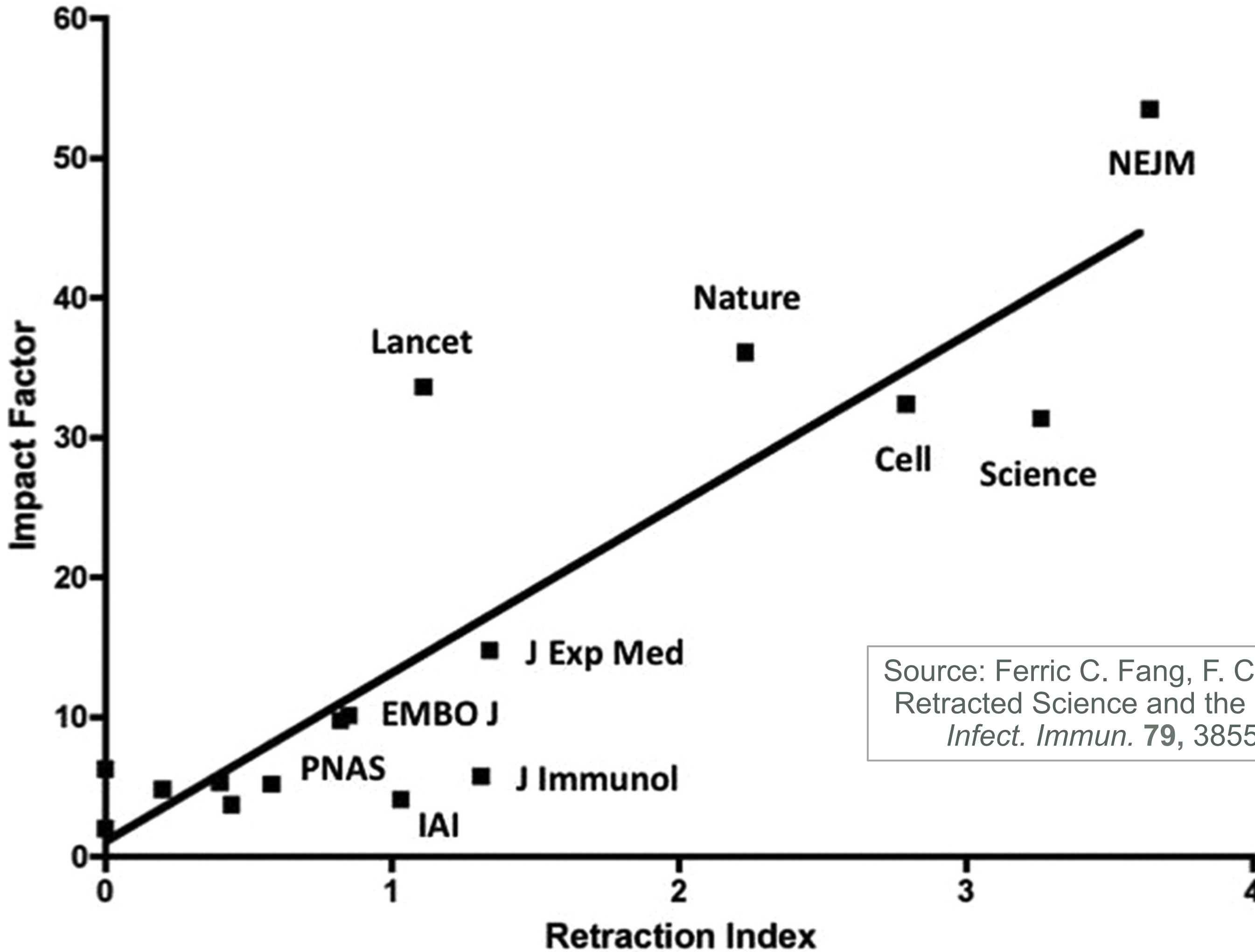
C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit.

"Fifty-three papers were deemed 'landmark' studies. ... Nevertheless, scientific findings were confirmed in only 6 (11%) cases."

Nature **483**, 531–533 (2012) doi:10.1038/483531a

The challenge for journals

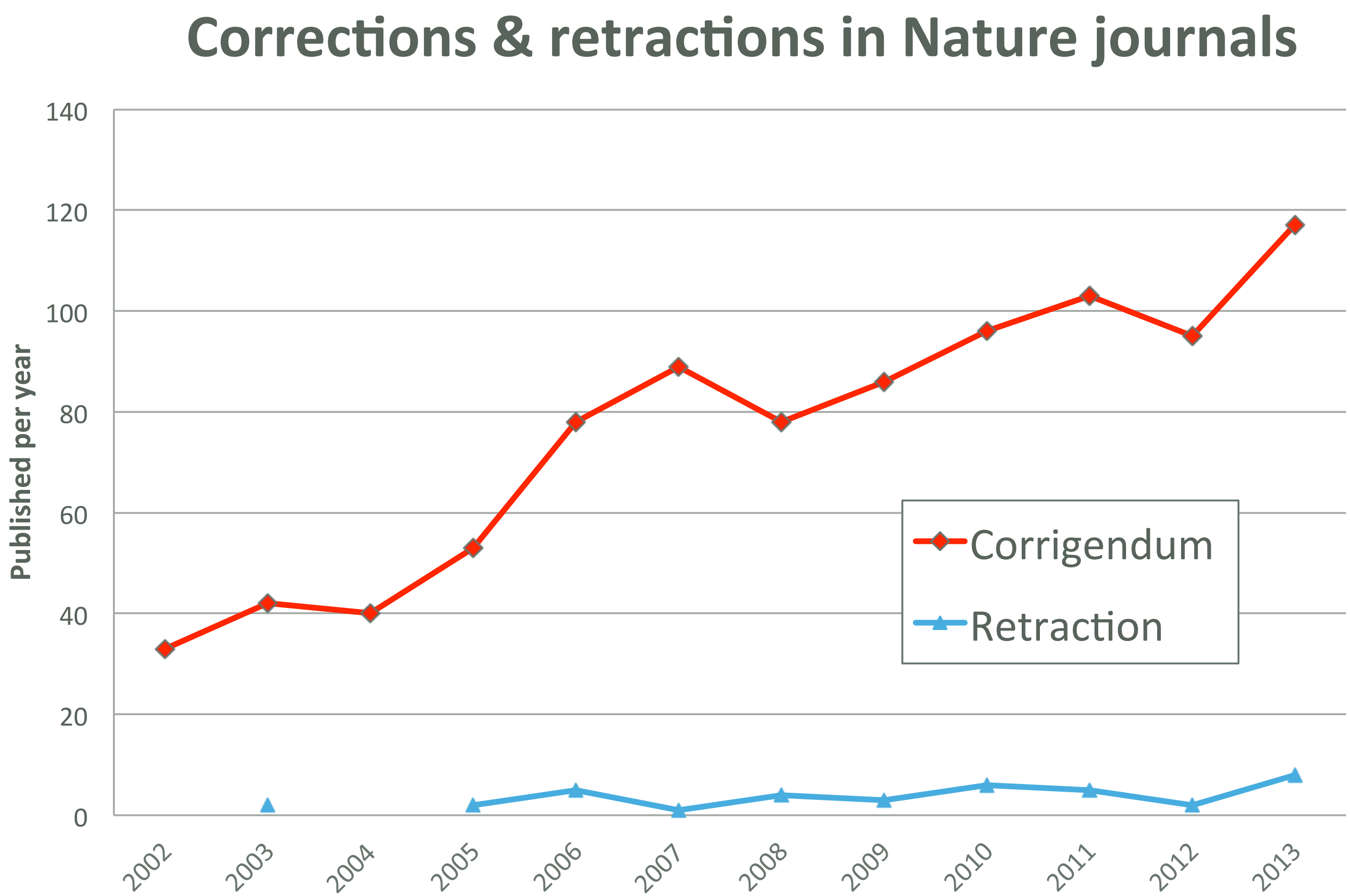
Retractions are ever more frequent and scale with impact factor



Source: Ferric C. Fang, F. C. & Casadevall, A. Retracted Science and the Retraction Index. *Infect. Immun.* **79**, 3855-3859 (2011).

The challenge for journals

Corrections are on the rise, too



The biggest problem isn't fraud

Although the lion share of retractions arise from misconduct, poor reproducibility does not

We are not talking about results that were:

- **Falsified** — misconduct is a problem but it's still in the minority
- **Wrong** — legitimate observations but subsequent work disproves the hypothesis — that's how science is meant to work!

We are talking about results that are

- **Poorly described preventing verification** — independent experimenter cannot observe the same results under similar conditions.
- **Overstated** — failure to consider alternative explanations
- **Misrepresented** — data claimed to be more robust than they actually are, CHERRY PICKING!
- **Sloppy** — failure to account for (or even consider) sources of error, poor use of statistics, poor controls

So, what are the solutions?

Part of the solution is education

We mustn't take it for granted that everyone knows what good research is

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

EDITORIAL

Rigorous Science: a How-To Guide

Arturo Casadevall,[a] Founding Editor in Chief, *mBio*, Ferric C. Fang,[b] Editor in Chief, *Infection and Immunity*

Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA[a]; Departments of Laboratory Medicine and Microbiology, University of Washington School of Medicine, Seattle, Washington, USA[b]

ABSTRACT **Proposals to improve the reproducibility of biomedical research have emphasized scientific rigor. Although the word "rigor" is widely used, there has been little specific discussion as to what it means and how it can be achieved. We suggest that scientific rigor combines elements of mathematics, logic, philosophy, and ethics. We propose a framework for rigor that includes redundant experimental design, sound statistical analysis, recognition of error, avoidance of logical fallacies, and intellectual honesty. These elements lead to five actionable recommendations for research education.**

Rigor is a prized quality in scientific work. Although the term is widely used in both scientific and lay parlance, it has not been precisely defined (1). Rigor has gained new prominence amid concerns about a lack of reproducibility in important studies (2, 3), an epidemic of retractions due to misconduct (4), and the discovery that the published literature is riddled with problematic images (5). Insufficient rigor may be slowing the translation of basic discoveries into tangible benefits (6, 7). New initiatives aim to understand deficiencies in scientific rigor and to make research more rigorous (8–10). Here, we consider the meaning of rigorous science and how it can be achieved.

The word rigor is derived from an old French word, "*rigueur*," meaning strength and hardness (11). In scientific vernacular, the underlying concept of strength resonates in the expressions "hard data" and "solid work" used to convey a sense of reliable and trustworthy information. In common usage, the word "rigor" has evolved to mean the quality of being exact, careful, or strict (12). Although the words "exact" and "careful" also apply to science, additional definition is needed since practicing rigorous science means more than mere exactness and care in experimental design. An experiment in which all components were exact in their proportions and the procedures carefully executed would still not be considered rigorous in the absence of appropriate controls. Hence, the definition of scientific rigor requires a deeper exploration than can be provided by simple perusal of the dictionary.

The scientific literature adds surprisingly little to our understanding of rigor, with the term almost always used without definition, as if its meaning is self-evident. The NIH has recently defined scientific rigor as "the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results" including "full transparency in reporting experimental details so that others may reproduce and extend the findings" (13). While we credit the NIH for providing a starting point for discussion, we find the NIH definition of rigor to be both excessively wordy and disconcertingly vague, as well as complicated by an insistence on transparency and reproducibility, which may be desirable but are arguably separate from rigor.

A WORKING DEFINITION OF SCIENTIFIC RIGOR

We suggest that rigorous science may be defined as theoretical or experimental approaches undertaken in a way that enhances confidence in the veracity of their findings, with veracity defined as truth or accuracy. Rigorous science could be entirely theoretical, as exemplified by a thought experiment used to illustrate a principle, such as Schrödinger's cat or Maxwell's demon in physics, or entirely experimental, as illustrated by Cavendish's measurement of the gravitational constant at the end of the 18th century. However, in the biomedical sciences, most research has both theoretical and experimental aspects.

A PENTATEUCH FOR SCIENTIFIC RIGOR

Different fields vary in the level of uncertainty that they are willing to accept with regard to conclusions. Certainty in science is often couched in terms of the probability that the null hypothesis may be rejected, which in turn depends on the methodologies employed. For example, the Higgs boson was announced when physicists were certain to "five sigma" or a P value of 3×10^{-7} (14). In contrast, many biological and medical studies accept a P value of 0.05, although more stringent criteria have been advocated (15). Does this make physics more rigorous than biology? Not necessarily—differences in the complexity of physical and biological phenomena as well as limitations in methodology determine the level of certainty that is practically achievable in these disciplines. Hence, a definition of rigorous science cannot rely on strict and arbitrary levels of certainty.

Traditional Chinese philosophy, Hinduism, Islam, and Judaism are each founded on five elements, pillars, or sacred texts. In Judaism, the first five books of the Hebrew bible are collectively referred to as the Pentateuch. Here, we humbly propose a Pentateuch for scientific rigor (Fig. 1).

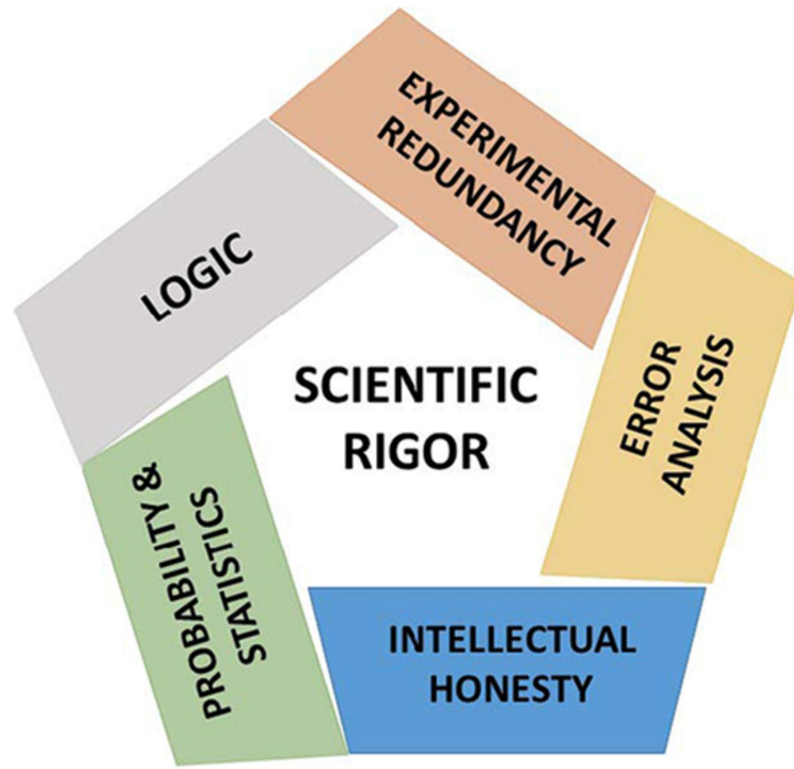
(i) Redundancy in experimental design. Good laboratory practices include proper controls, dose-response studies, determination of time courses, performance of sufficient replicates, and corroboration of major findings using independent experimental approaches and methods. It is important to establish whether a finding is generalizable, using a variety of cell types or species. New findings should lead to new predictions, which can in turn be experimentally tested. Experimental confirmation of predictions provides added assurance that the original findings are valid. Like rigor, redundancy is a multidimensional quality composed of many elements (Table 1). Redundancy in experimental design can enhance confidence in experimental results.

Published 8 November 2016

Citation Casadevall A, Fang FC. 2016. Rigorous science: a how-to guide. mBio 7(6): e01902-16. doi:10.1128/mBio.01902-16.

Address correspondence to Arturo Casadevall, acasade1@jhu.edu.

A Pentateuch for Scientific Rigour



Casadevall, A. and Fang, F. C. Rigorous Science: a How-To Guide. *mBio* **7**, e01902-16 (2016); doi:10.1128/mBio.01902-16

Encourage community standards

Case study: perovskite solar cells

- Rapid improvement in the efficiency of perovskite solar cells lead to a surge in claims of ever higher performance.
- Extraordinary claims based on poor experimental design lead to demands from referees that all claims be submitted to third party verification before peer review.
- Independent verification before peer review, enforced by journals, is now standard practice.

editorial

Solar cell woes

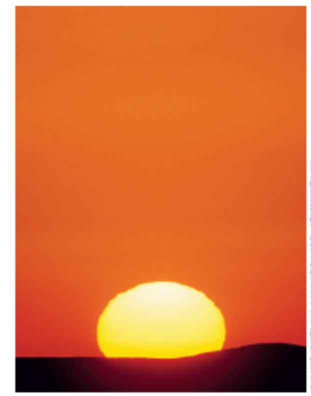
The pressure to publish results claiming organic solar cells with high efficiencies is leading to pervasive problems of false reporting within the community.

Widespread misreporting of power conversion efficiencies — a key metric for judging the performance of solar cells — is damaging the organic photovoltaics field and risks bringing it into disrepute. That's the damning conclusion of a new study performed by scientists from the University of Konstanz in Germany. Writing in a *Nature Photonics* Commentary (page 669), Lukas Schmidt-Mende and colleagues describe how they analysed 375 papers related to organic and organic–inorganic hybrid solar cells that were published in 13 journals during the period 2011–2012[1]. In each case, the team compared the claimed measured short-circuit current data (J_{SC}) of the cell with what could be reasonably expected from the cell's published external quantum efficiency (EQE). Shockingly, they found that in around one-third (37%) of cases a significant (>20%) discrepancy between the two existed, which draws serious doubts over the measurement techniques being used and the validity of efficiency claims that are being reported. In a number of cases, the discrepancies were larger than 100%. Furthermore, some of the papers reporting suspicious data have been highly cited and run the risk of becoming long-standing false benchmarks of performance.

This study is not the first to raise concerns over the integrity of organic solar cell characterization, but worryingly

reliable data (*Nature Mater.* 13, 837; 2014). The difficulties and challenges involved in characterizing nanostructured solar cells are also discussed in the September issue of *Nature Nanotechnology* (*Nature Nanotech.* 9, 657; 2014).

So what needs to be done to address the situation? Fortunately, it seems that the adoption of a small number of simple steps could be highly effective in stamping out the publication of sloppy and dubious data. Several groups have now published guidelines on how organic cells should be characterized, describing the common pitfalls that need to be avoided[5,6]. As suggested by Lukas Schmidt-Mende and co-workers in their Commentary, it's most important that solar cell papers provide adequate description of the methodology used for characterizing the cells as well as all experimental details. This not only brings the benefit of much-needed transparency for readers but will also serve to increase the reproducibility of published work. In particular, confirmation that a mask was used for measurements and information about the mask's size, the illumination source, the number of cells measured and the variation in their performance is especially important. These are details that *Nature Photonics* will now be requesting from authors and checking to ensure is present prior to the publication of all solar cell papers that feature power conversion



© MIKE WATSON / MOODBOARD / THINKSTOCK

Widespread problems with the characterization of organic solar cells threaten to bring the field into disrepute.

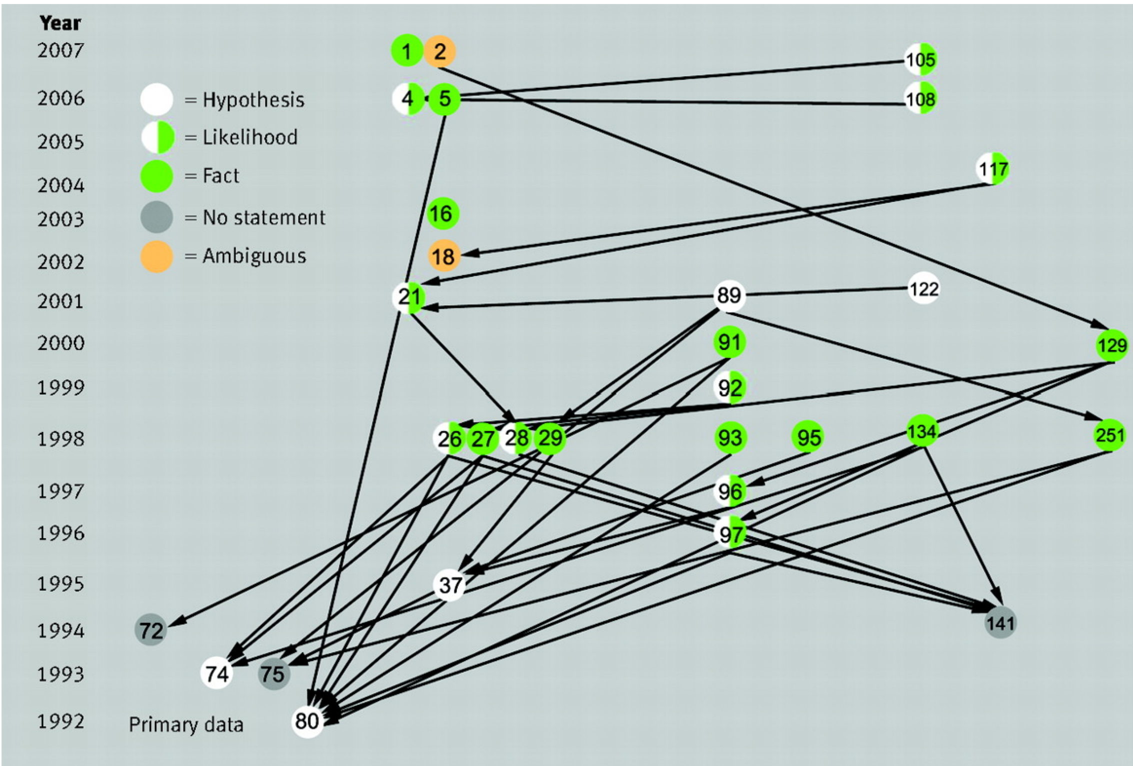
which is important to gain the trust of the community, but also means that they will be eligible for entry into official charts and tables that track the performance of solar cells[7,8].

We would, however, like to take the opportunity to make it explicitly clear that

Nature Photonics **8,** 665 (2014); doi:10.1038/nphoton.2014.212

Don’t cite anything you haven’t read!

Greenberg, S. A. How citation distortions create unfounded authority: analysis of a citation network. *BMJ* **339,** b2680 (2009); doi: 10.1136/bmj.b2680



“Fig. 5 — Conversion of hypothesis to fact through citation alone. Citations on statement that accumulation of β amyloid “precedes” other abnormalities in inclusion body myositis muscle. Statement as fact is supported through citation to papers that only state it as hypothesis (for example, references 5 to 80, 91 to 80, 134 to 74) or sometimes supported by citation to papers that contain no statements addressing it (for example, references 91 to 72, 251 to 75; dead end citations). This phenomenon might be called citation transmutation.”

Better methods sections, longer methods sections

Eliminate length limits for methods sections

extended methods online

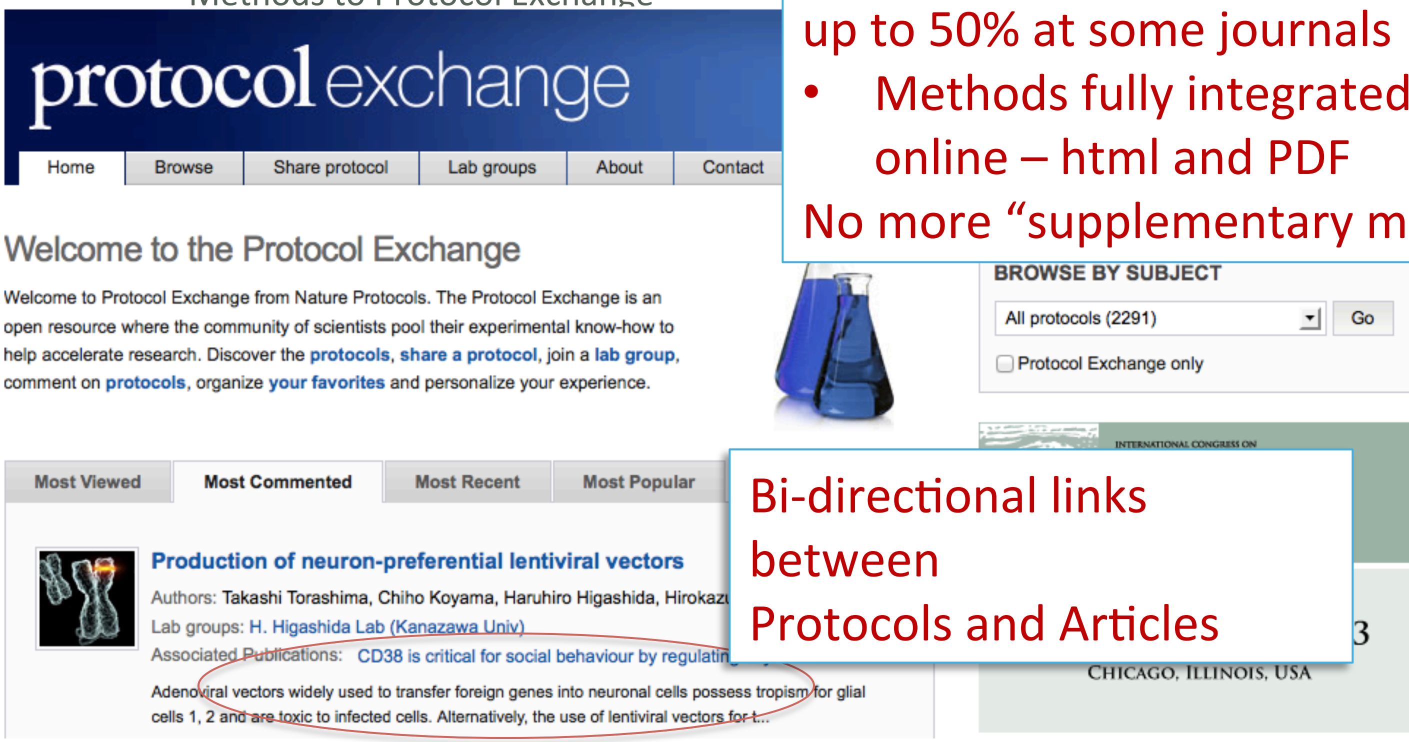
Encouraged to submit

Methods to Protocol Exchange

Net effect:

- Length increased by average 25% up to 50% at some journals
- Methods fully integrated with article online – html and PDF

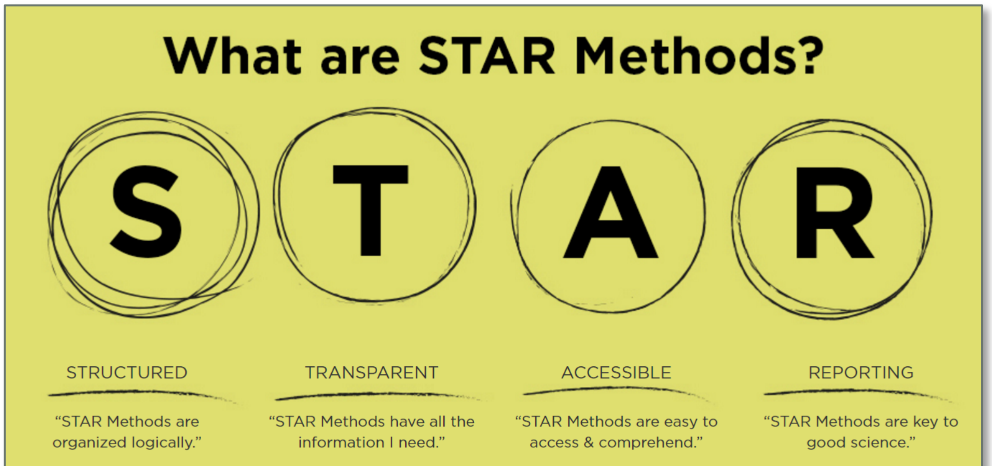
No more “supplementary methods”



Bi-directional links between Protocols and Articles

Better methods sections, longer methods sections

Cell Press
http://www.cell.com/star-methods



The antiseptic power of sunlight

I regularly tell young researchers that one way to decide whether something it ethical or not is to ask themselves,

"If this behaviour or practice were to become public knowledge, would I have any reason to feel uncomfortable or ashamed by it?"

And if the answer is 'yes' or even 'maybe', **don't do it!**

The main point here is that poor practice can only thrive in the dark. And **good practice comes by doing everything in the open**.

More transparent reporting

http://www.nature.com/authors/policies/checklist.pdf

Corresponding Author Name: ______________________

Manuscript Number: ______________________

nature publishing group | reporting checklist for life sciences articles

Reporting Checklist For Life Sciences Articles

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. For more information, please read Reporting Life Sciences Research.

Figure legends

Each figure legend should contain, for each panel where they are relevant:

- the **exact sample size (*n*)** for each experimental group/condition, given as a number, not a range;
- a **description of the sample collection** allowing the reader to understand whether the samples represent **technical or biological replicates** (including how many animals, litters, cultures, etc.);
- a **statement of how many times the experiment shown was replicated in the laboratory**;
- **definitions of statistical methods and measures**:
  - very common tests, such as *t*-test, simple χ^2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - **statistical test results**, e.g., ***P* values**;
  - definition of '**center values**' as **median or average**;
  - definition of **error bars as s.d. or s.e.m.**

Any descriptions too long for the figure legend should be included in the methods section.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents and animal models. Below, provide the page number(s) or figure legend(s) where the information can be located.

Statistics and general methods

Encourage more sharing of data

SCIENTIFIC DATA

BioMed Central
The Open Access Publisher
biomedcentral.com

(GIGA)^n SCIENCE
Editor-in-Chief:
Laurie Goodman (USA)

Publish a Data Note of your research in *GigaScience* and get credit for your data

figshare
credit for **all** your research

biosharing

Altmetric

DataCite
Helping you to find, access, and reuse data

DRYAD
Data available in Dryad

Sharing platforms needn't just be about sharing (or at all)

If your data isn't saved in a secure, persistent repository it may as well not exist

figshare
credit for **all** your research

A **cloud based** research data management system where you can:

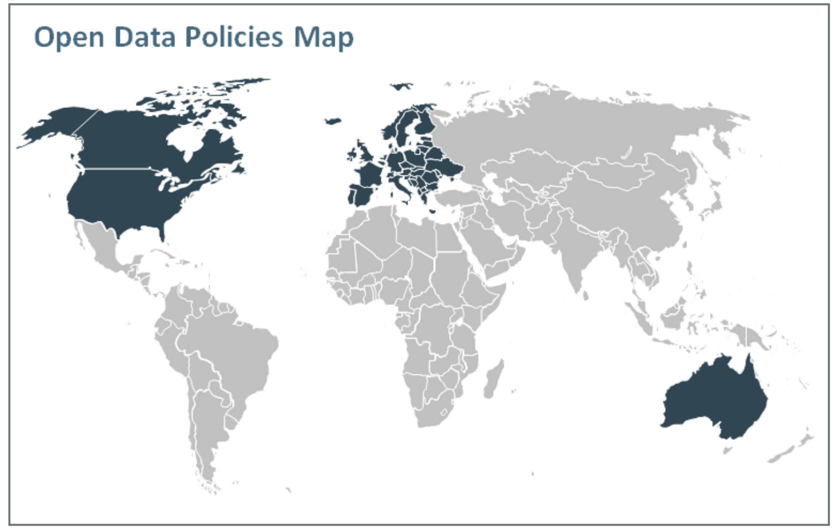
Manage your research outputs privately and securely

Make your research outputs citable & discoverable

Why share data?

Because funders increasingly require it

- At least 28 research funders globally have policies or mandates requiring archiving of data as a condition of grants, including:
  - National Science Foundation (NSF)
  - National Institutes of Health (NIH)
  - Wellcome Trust
  - Bill and Melinda Gates Foundation
- Some of these require data to be linked to publications including:
  - Research Councils UK (as part of open access policy)
  - Engineering and Physical Sciences Research Council (EPSRC)



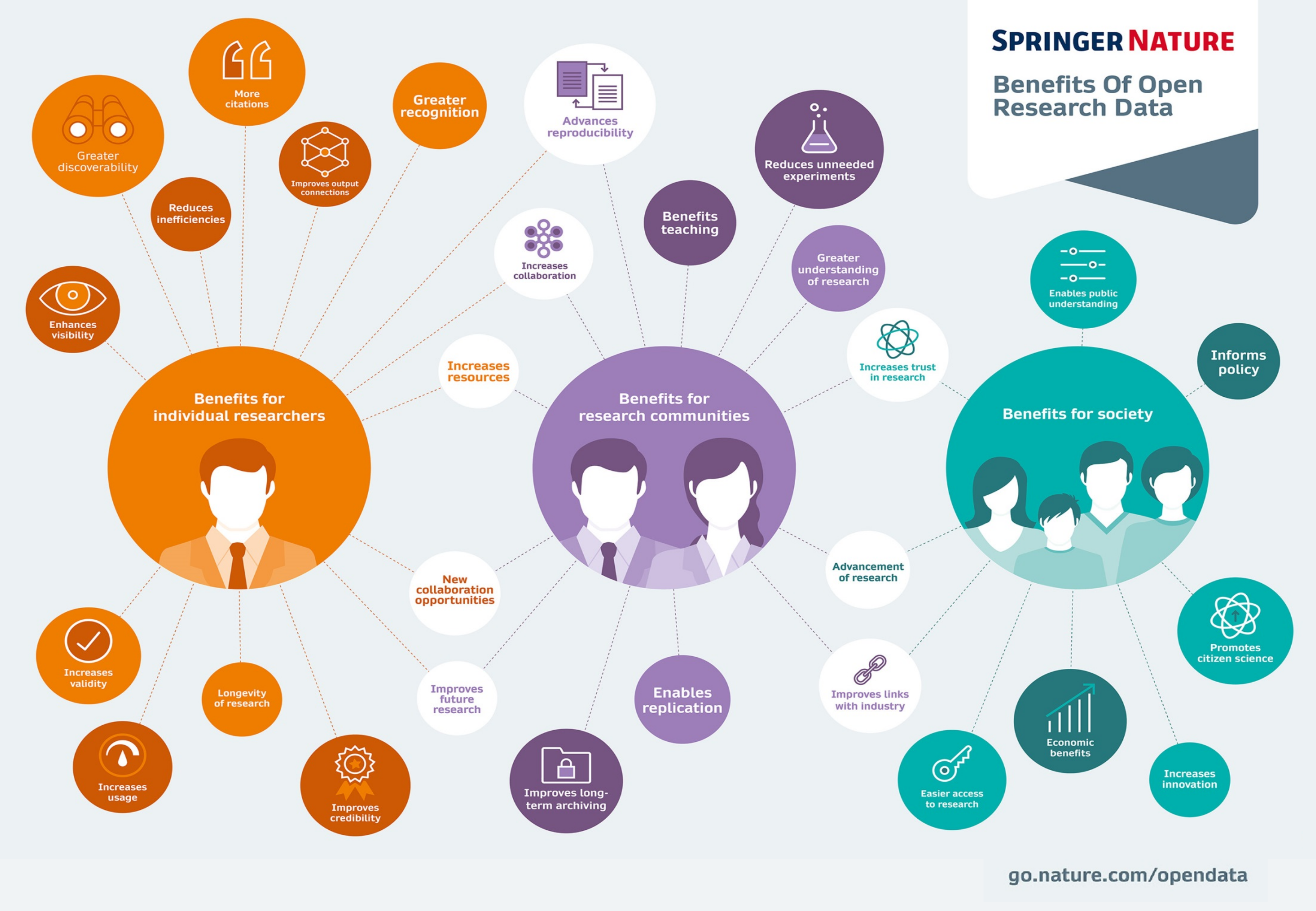
Source: Hahnel, M: Global funders who require data archiving as a condition of grants. *Figshare*. https://dx.doi.org/10.6084/m9.figshare.1281141.v1 (2015)

Why share data?

Because that’s what science should be about

“Publishing research without data is advertising, not science!”

— Cameron Neylon, Graham Steel, and others.


SPRINGER NATURE
Benefits Of Open Research Data
Benefits for individual researchers
Greater discoverability
More citations
Improves output connections
Greater recognition
Reduces inefficiencies
Enhances visibility
Increases validity
Increases usage
Longevity of research
Improves credibility
Benefits for research communities
Advances reproducibility
Increases collaboration
Increases resources
New collaboration opportunities
Improves future research
Improves long-term archiving
Enables replication
Benefits teaching
Reduces unneeded experiments
Greater understanding of research
Benefits for society
Increases trust in research
Enables public understanding
Informs policy
Advancement of research
Improves links with industry
Easier access to research
Economic benefits
Increases innovation
Promotes citizen science
go.nature.com/opendata

Open data in China

Case study: China’s open corrosion data platform

Corrosion costs around US$4 trillion a year globally.

Share corrosion data

To prevent disasters, **Xiaogang Li** and colleagues call for open data infrastructures to collate information on materials failures.

In November 2013, an oil pipeline in the Chinese city of Qingdao exploded, killing 62 people and wounding 136. Eight months later, a similar explosion in Kaohsiung caused 32 deaths and 321 injuries. The pipelines were made of steel of the

A lack of knowledge hinders our ability to prevent failures. Degradation of underground pipes, for example, is influenced by the compositions, microstructures and designs of materials, as well as by a raft of environmental conditions such as soil oxy-

“The Chinese government has invested nearly 200 million yuan (US$30 million) since 2006 on a platform for sharing corrosion data from 30 field-testing stations covering standard materials in environments (air, soil and water) typical of different parts of the country.”

—

Li, X. G. *et al. Nature* **527,** 441–442 (2015).

But is the world ready to open up their data?

Or even share it on request?

Has open data arrived at the *British Medical Journal (BMJ)*? An observational study

Objective To quantify data sharing trends and data sharing policy compliance at the *British Medical Journal (BMJ)* by analysing the rate of data sharing practices, and investigate attitudes and examine barriers towards data sharing.

Results … only 7/157 research articles shared their data sets, 4.5% …

Conclusions Despite the *BMJ*'s strong data sharing policy, sharing rates are low. Possible explanations for low data sharing rates could be: the wording of the *BMJ* data sharing policy … requests [ending] up in researchers spam folders; and that researchers are not rewarded for sharing their data.

— Anisa Rowhani-Farid & Adrian G Barnett
BMJ Open **6,** e011784 (2016).

Encourage more sharing of data

Nature now requires authors to tell readers how they can get hold of their data

In March 2016 we ran a pilot requiring authors of papers at five *Nature* journals to declare how readers can obtain the ‘minimal data set’ needed to interpret, replicate and build on the findings they report.

By the end of 2016 we rolled this out to all journals in the *Nature* family.

As well as being an inherent good, we hope to:

- Learn more about how different disciplines share data differently.
- Learn how we can help them meet the challenges they.
- Raise awareness and encourage more research into the open.

ANNOUNCEMENT

Where are the data?

As the research community embraces data sharing, academic journals can do their bit to help. Starting this month, all research papers accepted for publication in *Nature* and an initial 12 other Nature titles will be required to include information on whether and how others can access the underlying data.

These statements will report the availability of the ‘minimal data set’ necessary to interpret, replicate and build on the findings reported in the paper. Where applicable, they will include details about publicly archived data sets that have been analysed or generated during the study. Where restrictions on access are in place — for example, in the case of privacy limitations or third-party control — authors will be expected to make this clear.

Nature **536,** 138 (20016). To learn more go to http://go.nature.com/2bf4vqn

And not just data — materials are just as important!

"A condition of publication in a Nature journal is that authors are required to make unique materials promptly available to others without undue qualifications."
http://www.nature.com/authors/policies/availability.html

"After publication, all data and materials necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of Science."
http://www.sciencemag.org/authors/science-editorial-policies?_ga=1.200322094.1663650950.1484202254

"One of the terms and conditions of publishing with Cell Press is that authors be willing to distribute any materials and protocols used in the published experiments to qualified researchers for their own use."
http://www.cell.com/cell/authors

And finally … pre-registration

What is Preregistration?

"Preregistration separates hypothesis testing from hypothesis generating research. Both are important. However, the same data cannot be used to generate and test a hypothesis, which often happens unintentionally. With preregistration, confirmatory analyses are planned in advance in order to retain the validity of their statistical inferences, and exploratory analyses are reported as post hoc investigations that might inspire confirmatory tests in future studies."

— from https://cos.io/prereg/

Thank you!　谢谢！

Ed Gerstner 印格致 — e.gerstner@nature.com

Further reading

- ‘Turning Point’ white paper on Chinese research landscape (2015) — http://www.nature.com/press_releases/turning_point.pdf
- Jane Qiu, Safeguarding research integrity in China. *Natl Sci. Rev.* **2,** 122–125 (2015); doi: 10.1093/nsr/nwv002
- Munafò, M. R. *et al*. A manifesto for reproducible science. *Nature Human Behaviour* **1,** 0021 (2017); doi:10.1038/s41562-016-002

natureresearch

“Good Research Practice”

February 20, 2017
Jun Fudano, Ph.D.
(Tokyo Institute of Technology)

Three senses of “Goodness”

Traditional Theories of Ethics

Agent → Conduct → Consequences

↓ ↓ ↓

Virtue Ethics

Deontological Ethics

Utilitarian Ethics

Line Drawing/Demarcation Problem

"How good is good enough?"

"How good is good enough for what purpose, for whom?"

The same line for every discipline?

Well-being Model 1

What is welfare/well-being/happiness?

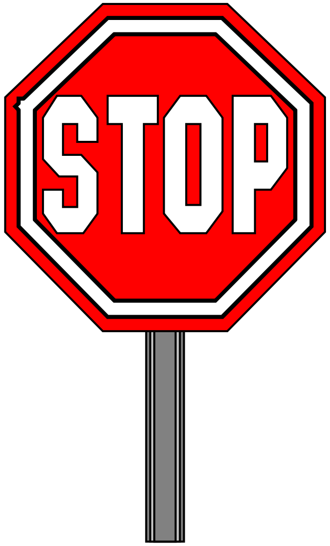
T

“Hold pa ount the safety, health, and we are of the public”

Two aspects of Ethics

Aspire to do	Ought not to do
Aspirational Ethics	Preventive Ethics
Positive	Negative
External orientation	Internal orientation
Tend to be uplifting	Tend to be intimidating

Contact



Jun Fudano, Ph.D.

Professor
Institute for Liberal Arts
Tokyo Institute of Technology

W-9 63 2-12-1 Ookayama

Meguro-ku, Tokyo 152-8552, JAPAN

Tel.: +81-3-5734-2370 Fax: +81-3-5734-2844

E-mail: fudano.j.aa@m.titech.ac.jp

Quantity, Quality and Quality

- Publish or Perish - Quantity
- Impact factors and citations - Quality
- Research impact - Excellent Quality
- Promotions and Ranking

Academic Excellence:

Quality Research is Responsible Research

Research Data Management

Researcher:

- Data management plan
- Reliable data collection, usage and interpretation
- Practice good data recording
- Data sharing

Institution:

- Research data management policy
- Data storage and archive

APRI, Prof. Mai Har Sham, The University of Hong Kong