Responsible Conduct of Research

Susan Porter
Dean and Vice-Provost,
Graduate & Postdoctoral Studies

Jacqui Brinkman
Manager, *Graduate Pathways to Success*,
Graduate & Postdoctoral Studies

(w/ edits by Wendy Robinson, FoM Asst Dean Graduate & Postdoctoral Education)

The classic scientific paradigm

Generate hypothesis

Design study

Collect data

Analyze data/test hypothesis

Interpret data

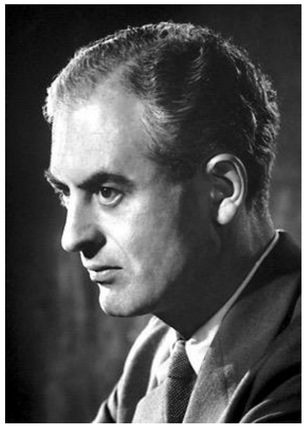
Peer review / Publication

Every stage subject to threats/biases:

Bias
Low power
Poor Quality
P-hacking
Lack of knowledge
Conflicting interests
HARKing

“Everything that a scientist does is a function of what others have done before him; the past is embodied in every new conception and even in the possibility of its being conceived at all.”

Peter Medawar



"father of transplantation”

e.g.
- scholarly integrity (85)
- research (87)
- conflicts of interest (97)

Policies **Regulations**	**Discipline Norms** **Judgment**	**Societal norms** **Shared values**
Fabrication Falsification Plagiarism Misuse of funds Abuse of students	*Sharing of data* *Data interpretation* *Authorship* *Conflicts of interest* *Mentorship*	Honesty Objectivity Efficiency Accuracy Conscientiousness

Objectives

Students will have

- a better understanding of norms and rules for responsible research conduct
- ability to know where and how to ascertain these
- improved ability to make judgments on the ethics of the types of actions and decisions inherent in research

conduct research responsibly!

CITI Collaborative Institutional Training Initiative

2 hours face to face

- Intro to the Responsible Conduct of Research
- Ethics and the Responsible Researcher
- Research Misconduct
- Data Acquisition and Management
- Publication Practices and Responsible Authorship
- Peer Review: Role and Process in Life Sciences Research
- Writing with Integrity
- Responsible Mentoring
- Collaborative Research
- Conflicts of Interest in Research
- Human Participants Research and Ethics
- Animal Care and Use

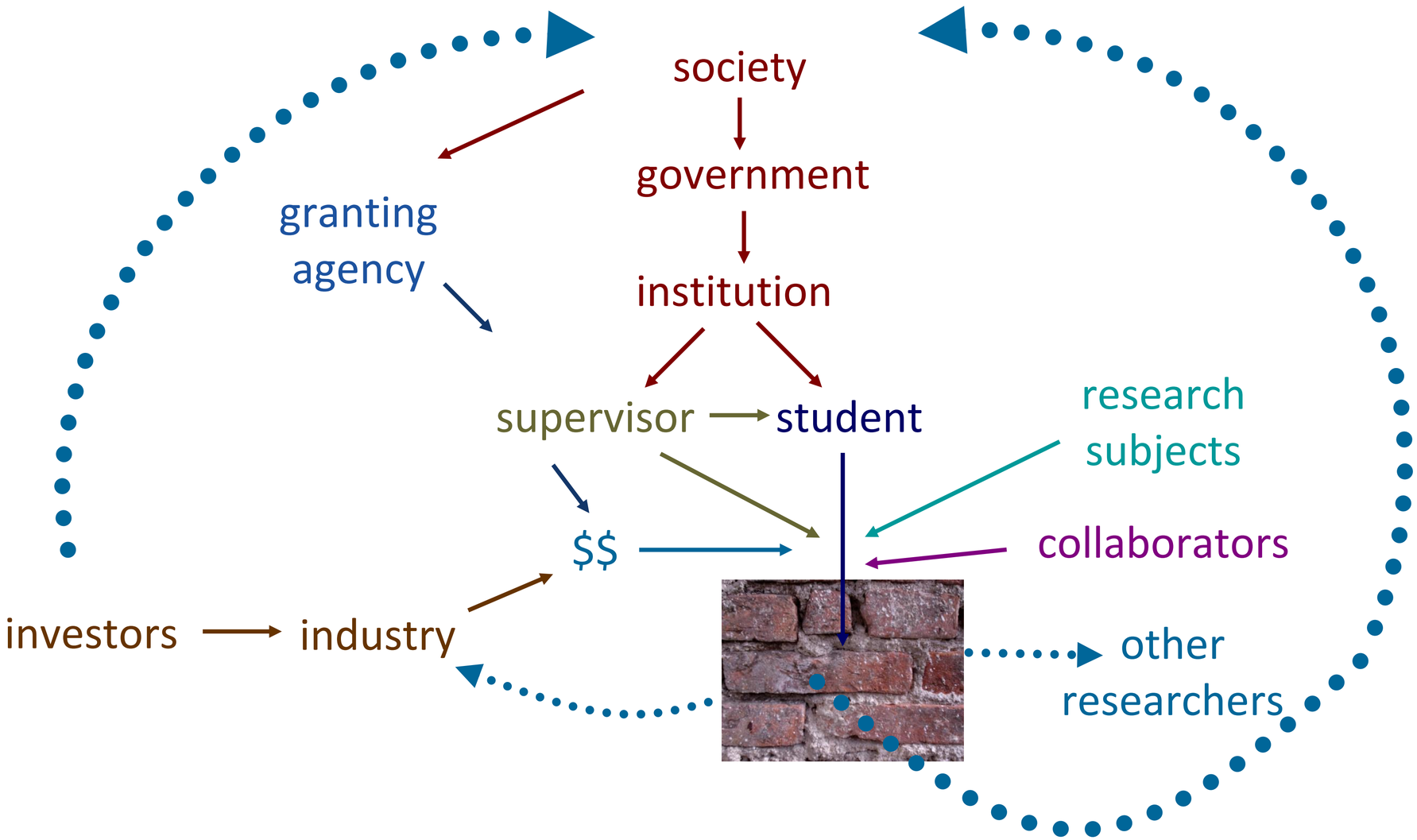
2 hours face to face

Role Play!

Third time “lucky”…

- What are the conflicting interests here?
- How does one differentiate between experiments that “don’t work” vs. those that have negative or unanticipated results?
- Who are the stakeholders?
- What human factors or supervision practices might improve scientific rigour in this lab

Who are we responsible to?



What are we responsible for?

Creation of new knowledge that is:

- true (accurate, conclusions are well-justified)
- accessible to other researchers
- able to be independently verified by other researchers
- of benefit to society

Carried out in a manner that:

- makes efficient use of resources
- minimizes impact on people, animals, and the environment
- involves honest, caring, and fair relationships with peers and mentees

Ethical Dilemma: Conflict between competing moral imperatives or values

Need and duty to publish ⟷ Duty to facilitate research of others

Need to publish in a timely fashion ⟷ Need to be thorough and careful

Financial gain ⟷ Objectivity, honesty

Recognition, tenure, etc. ⟷ Careful/courageous? research

Decisions can be difficult, and may be related to individual context

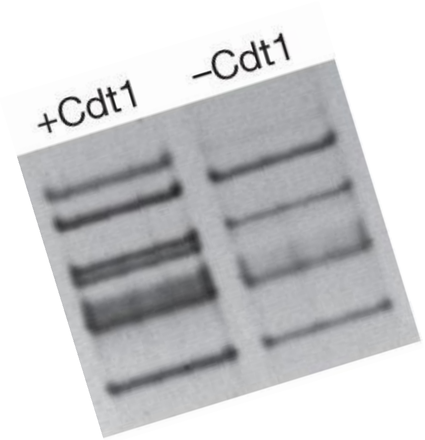
A Framework for Ethical Decision-Making

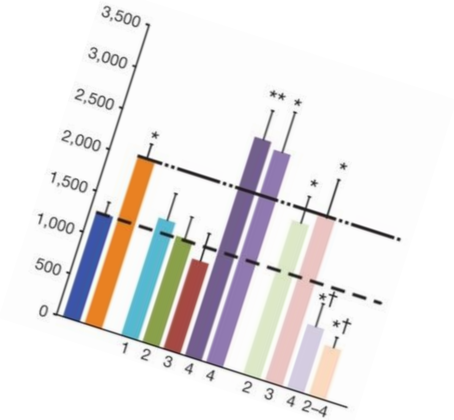
- Identify the problem(s), determine whether it’s an ethical issue.
- Identify the individuals/groups with a stake in the outcome
  - What are their legitimate rights, responsibilities and expectations?
- Get the facts and assess the situation
- Identify options for acting and evaluate them
  - What are the consequences for all affected parties?
  - What if everyone in these circumstances did this?
  - What will do the most good with the least harm?
- Make a decision and test it
- Act and reflect on the outcome

(End Part 1)

Case discussions:
—Karen Klein
—Arun Patel

Responsible Data….





Obtaining

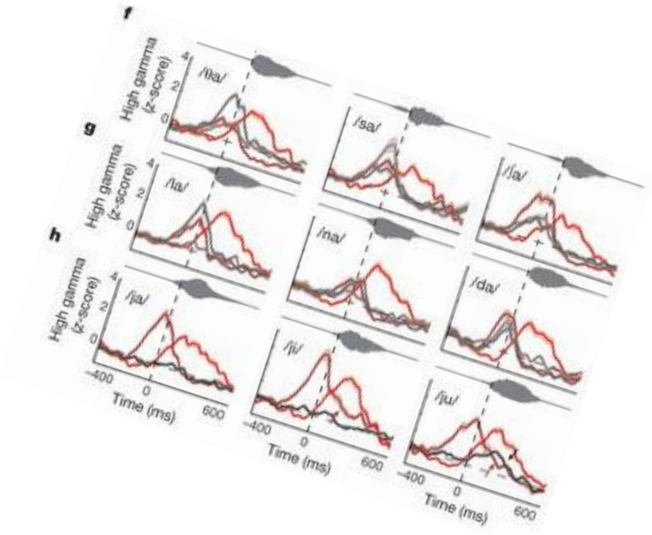
Recording

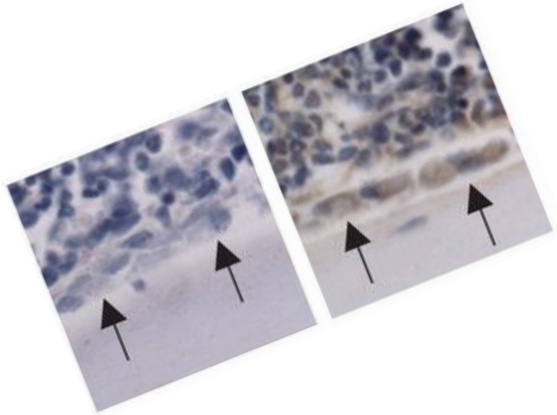
Analyzing

Reporting

Ownership

Sharing



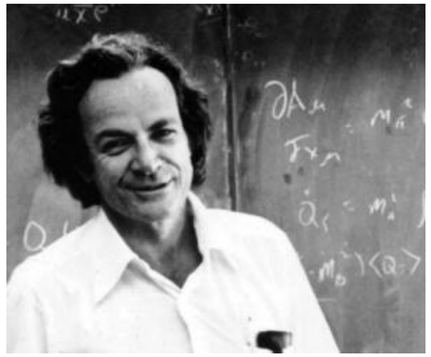


Obtaining Data

“ The first principle is that you must not fool yourself, and you are the easiest person to fool so you have to be careful about that.

I'm talking about a specific, extra type of integrity…..bending over backwards to show how you're maybe wrong, that you ought to do when acting as a scientist. This is our responsibility as scientists….”

- Richard Feynman, Nobel Laureate



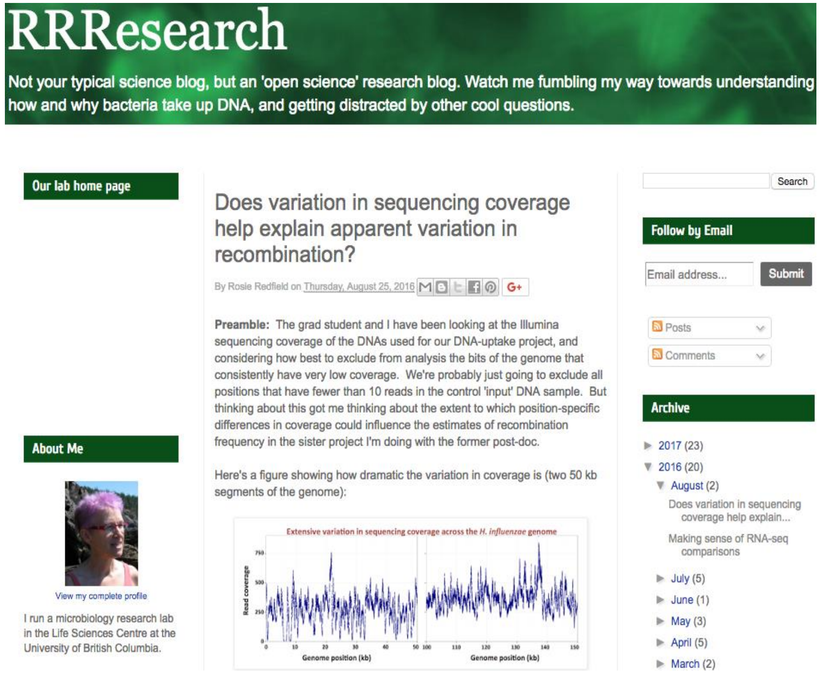
www.thefamouspeople.com/profiles/images/richard-feynman-3.jpg

- experimental design is free from bias
- sample is appropriate
- conducted carefully with proper controls
- understand the statistics to be used from the start
- get independent advice (mentor as well as others)
- consider preregistering study

Sharing research and analysis plans makes for stronger science

Study preregistration

Open research blog



Recording Data

Data books should inform:

- WHY, WHEN, & HOW
- RESULTS
- CONCLUSIONS
- IDEAS for next steps

Why is this important?

Legacy for the lab.

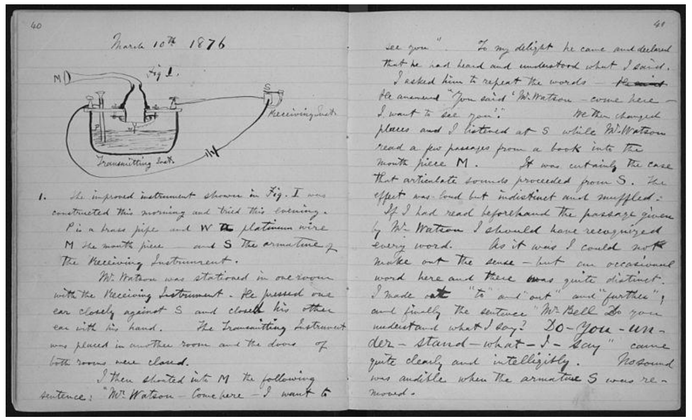
Others should be able to repeat what you did.

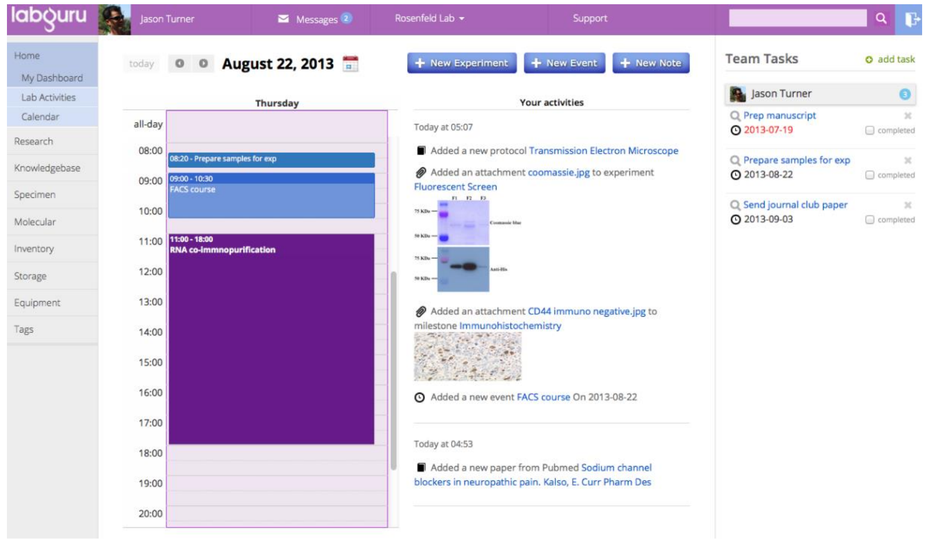
Provides legal support for patents or defense against fraud.

Classic Notebook

Electronic notebook

- Bound
- Pen
- Legible
- Numbered pages

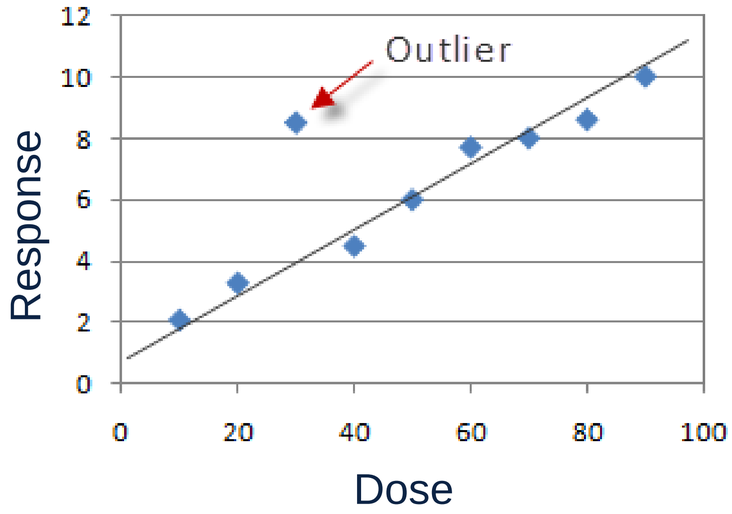




https://www.labguru.com/tour

1. Recorded and dated as data collected
2. Include raw data + calculations or links to location
3. Errors corrected without obscuring original recording
4. Include relevant correspondence
5. Include interpretations

Analyzing and Reporting Data: Outliers/data trimming



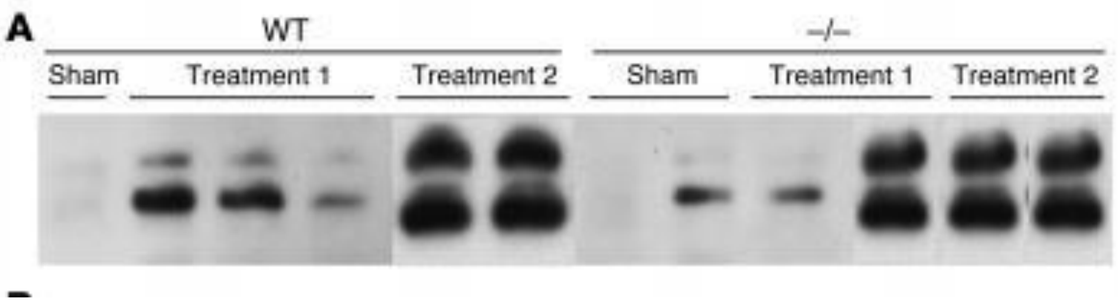
It is NOT acceptable to drop an observation just because it is an outlier

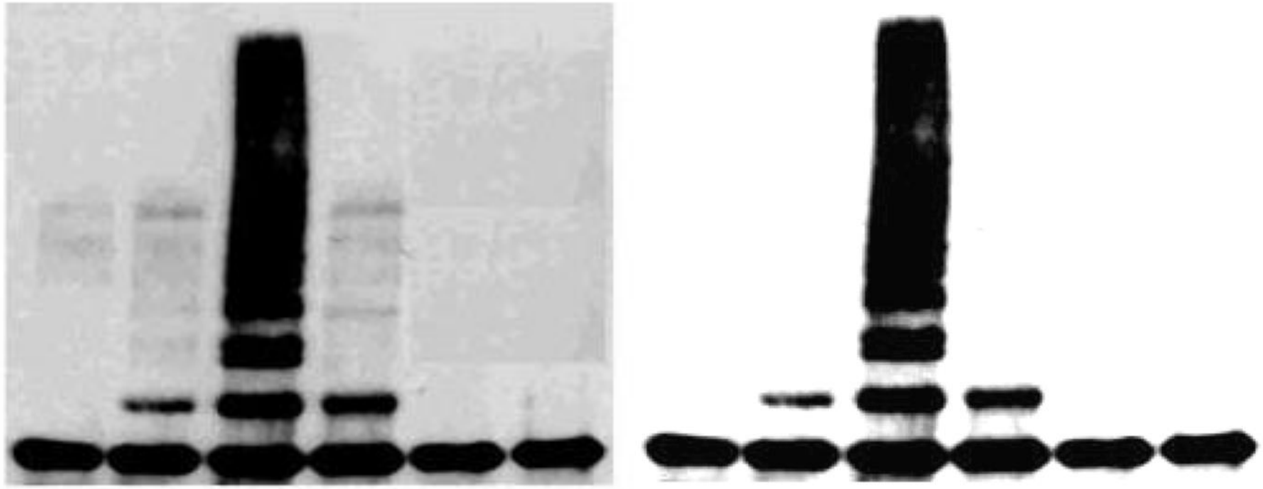
- Decide data parameters before hand

(if you suspect a potential problem don't wait to decide whether to include the data point(s) only after seeing the results!)

- Do data QC checks prior to analysis
- Determine likely cause if possible
- Assess using statistical measures.
- If it is clearly due to a measurement/recording error, can drop, but <u>always disclose the basis</u> for dropping or modifying data

Analyzing and Reporting Data: Image data





Label graphic data clearly

Don’t splice together images, crop images, use touch-up tools

Take high-resolution pictures from start

Don’t doctor image inappropriately!

Do not reuse previously published figures

guidelines about images: *Nature 439, 891-892(23 February 2006)*

http://www.nature.com/authors/policies/image.html

“The most dangerous of all falsehoods is a slightly distorted truth.”

G.C. Lichtenberg (1742–1799)

“Beautification is a form of misrepresentation. Slightly dirty images reflect the real world.”

Nature 439, 891-892(23 February 2006)

PUBPEER
The online Journal club

PubPeer is a website that allows users to discuss and review scientific research.

It has highlighted shortcomings in several high-profile papers, in some cases leading to retractions and to accusations of scientific fraud,

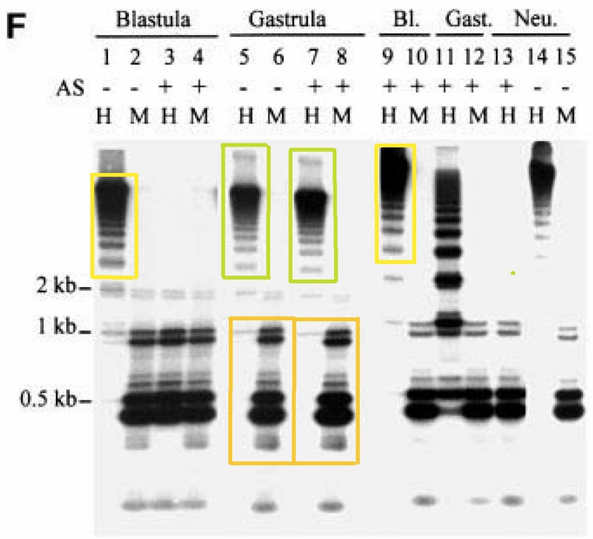


Figure 1F EMBO J 20,1963.

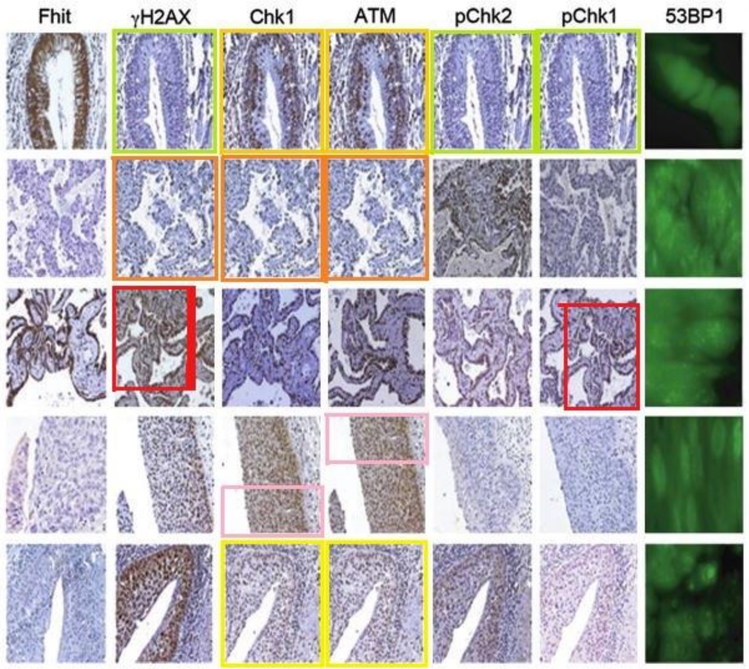
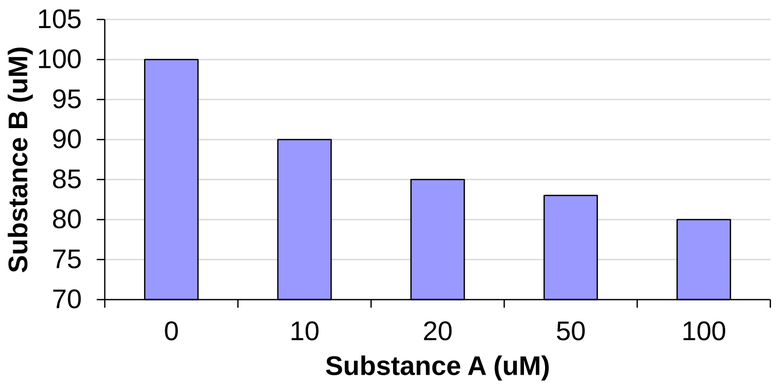
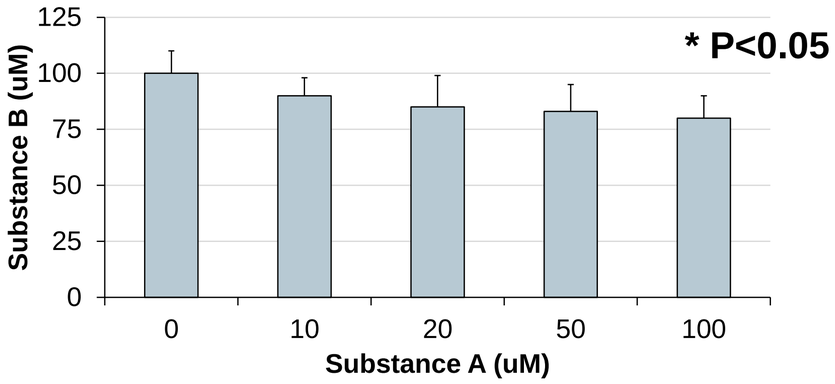
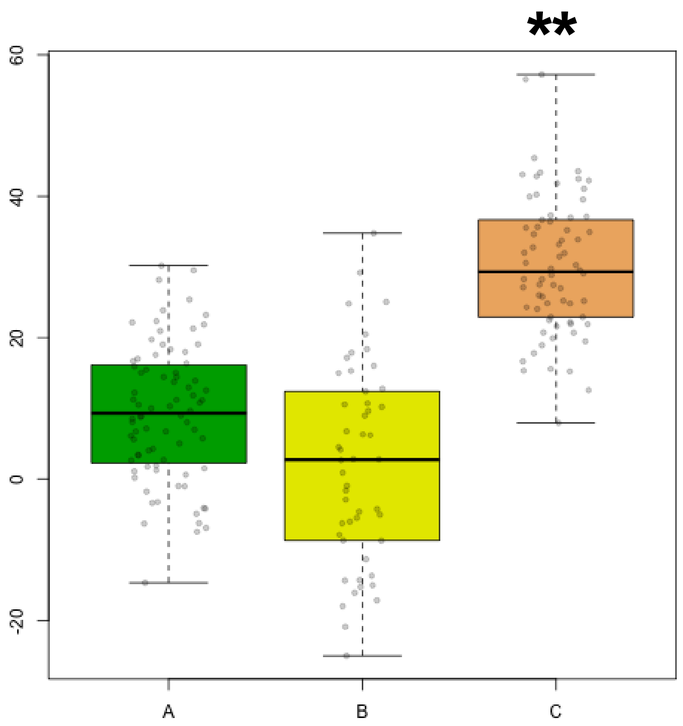


Figure 1 Cancer Lett 291,230. 2010

Data Visualization: Graphical data





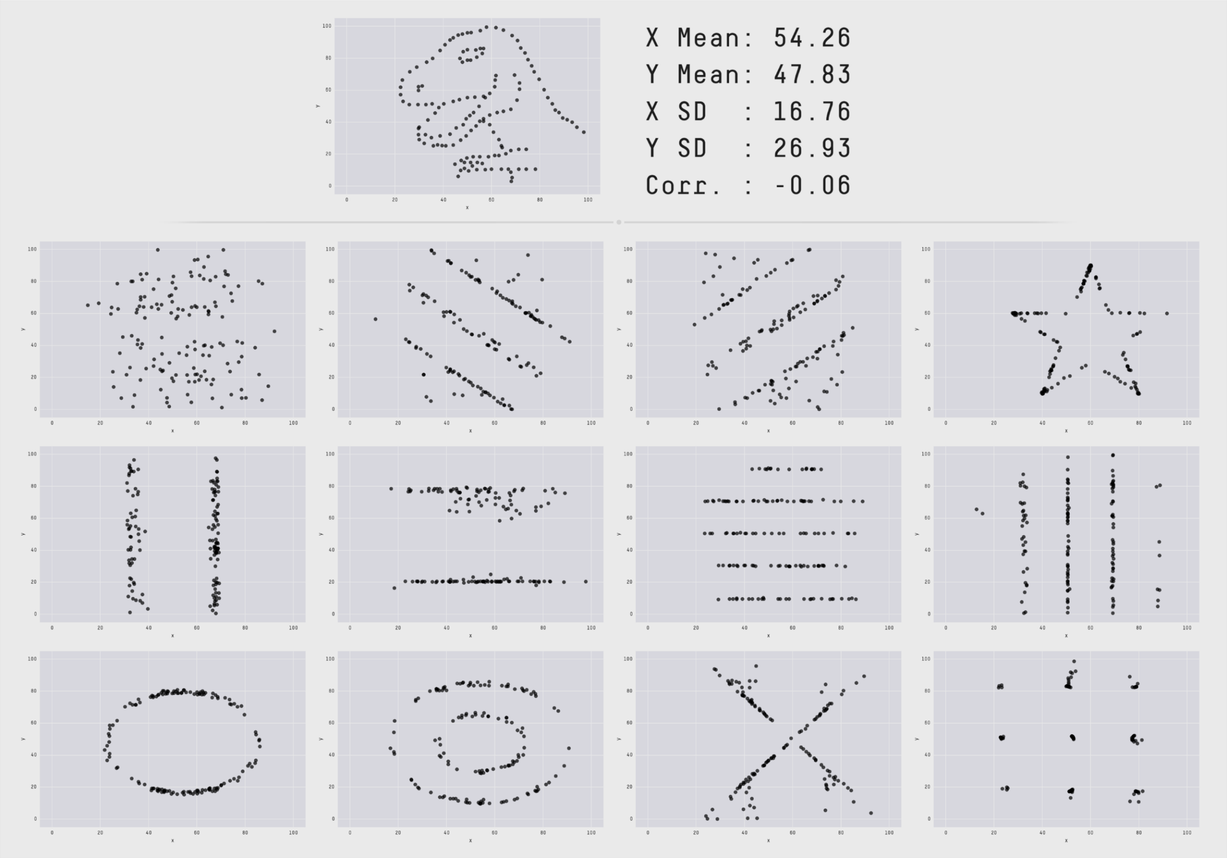


www.r-graph-gallery.com/boxplot/

convey information clearly, do not mislead, show data points if possible

Data Visualization: The value of showing the data

All of these plots have the same summary statistics



https://www.autodeskresearch.com/publications/samestats

Analyzing and Reporting Data: Statistics

- Use methodologies appropriate to the data
- Remain current and understand the theory, the data, and the methods (or enlist statistical help)
- Consider statistical methods before research is undertaken
- Report all assumptions and data cleaning/screening procedures used
- Address potential confounding variables
- **Correct for multiple comparisons**!

In today's environment, our talent for jumping to conclusions makes it all too easy to find false patterns in randomness, to ignore alternative explanations for a result or to accept 'reasonable' outcomes without question — that is, to ceaselessly lead ourselves astray without realizing it. – *Nusso, Nature 2015*

Analyzing and Reporting Data: Interpretation

Common errors (honest or otherwise)

- association ≠ causation
- *a posteriori* hypothesis presented as *a priori*
- not taking into account all data
- failing to account for technical confounders
- incomplete description of methods or materials
- Don’t over-interpret p-values
- not considering alternative explanations

Eur J Epidemiol (2016) 31:337–350
DOI 10.1007/s10654-016-0149-3

ESSAY

Statistical tests, *P* values, confidence intervals, and power: a guide to misinterpretations

Sander Greenland[1] · Stephen J. Senn[2] · Kenneth J. Rothman[3] · John B. Carlin[4] · Charles Poole[5] · Steven N. Goodman[6] · Douglas G. Altman[7]

25 common misconceptions about P-values

1) The P value is the probability that the test hypothesis is true; i.e, if a test of the null hypothesis yields P = 0.01, the null hypothesis has only a 1 % chance of being true;

No! The P value assumes the null hypothesis is true—it is not a hypothesis probability.

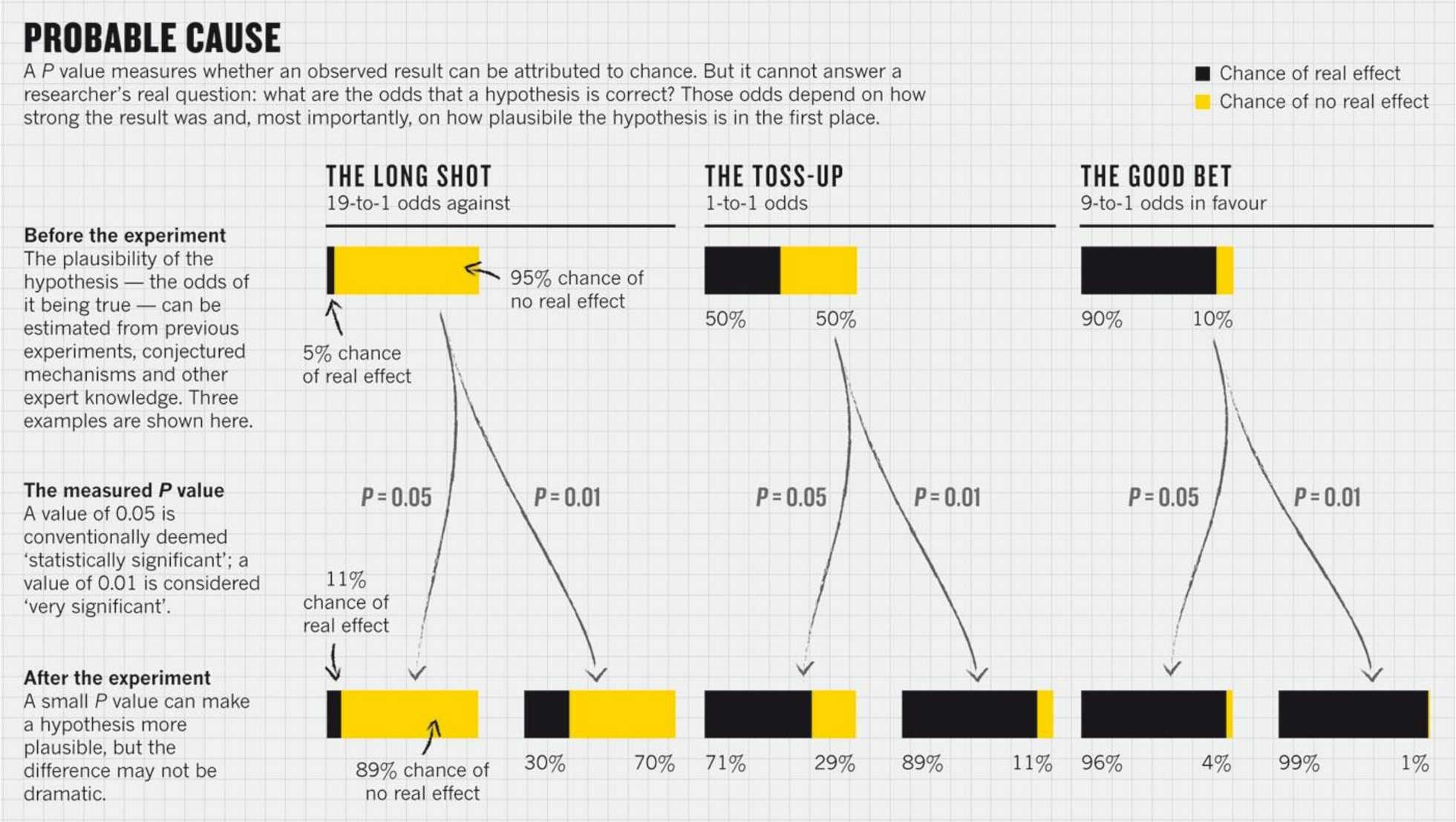
3) A significant test result (P < 0.05) means that the test hypothesis is false or should be rejected.

No! A small P value simply flags the data as being unusual if all the assumptions used to compute it were correct; it may be small because there was a large random error or because some assumption other than the test hypothesis was violated.

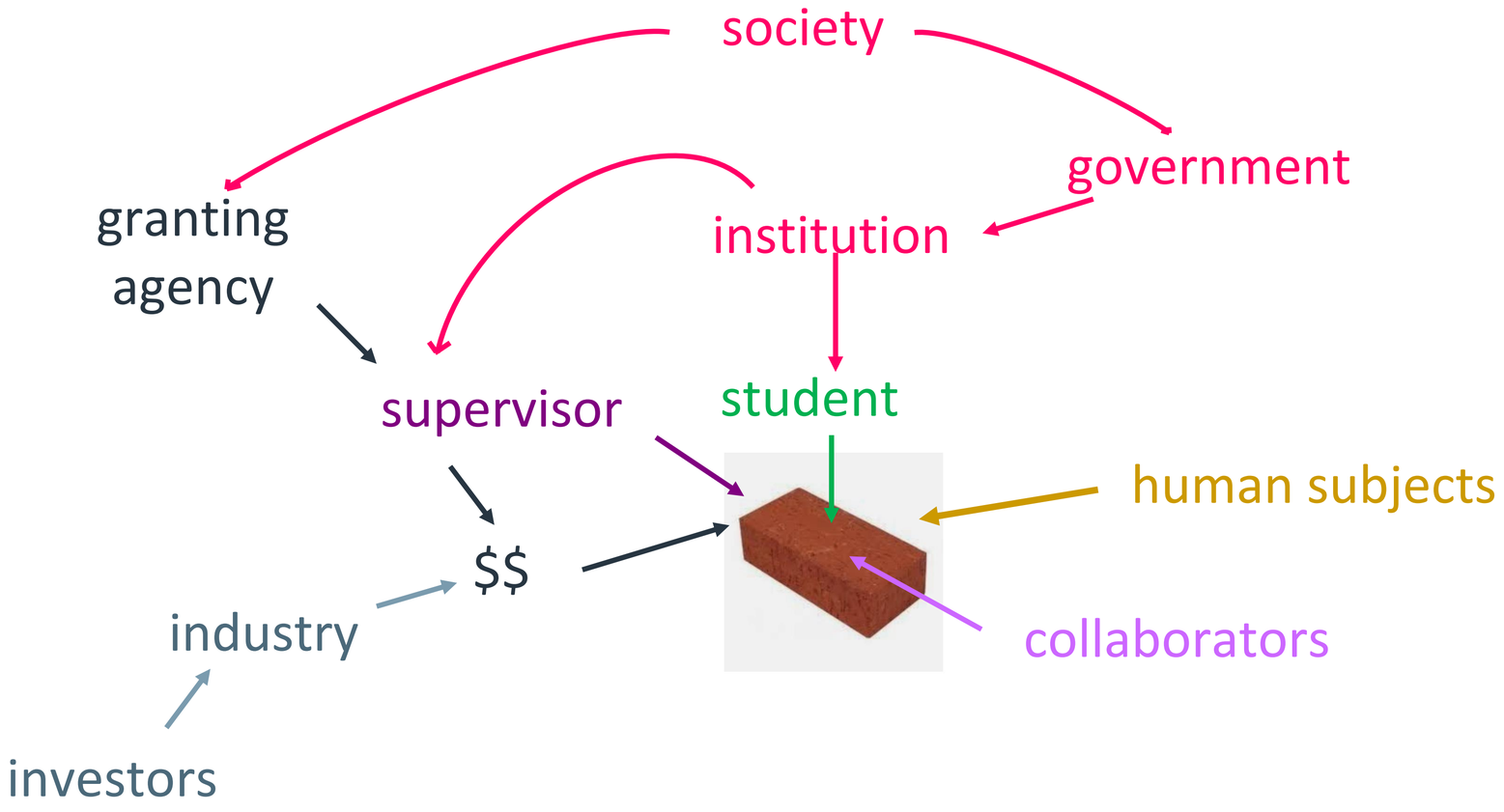
7) Statistical significance indicates a scientifically or substantively important relation has been detected.

No! Especially when a study is large, very minor effects or small assumption violations can lead to statistically significant tests of the null hypothesis.

Analyzing and Reporting Data: Statistics



Ownership of Data & Intellectual Property



Ownership of Data & Intellectual Property

- Research data usually belong to the institution or jointly by the institution and researcher(s)

Norms

- Individual researchers are generally permitted to have a copy of the study data (after subject identifiers are removed)

Sharing of Data & Materials

"An author's obligation is…to release data and materials to enable others to verify or replicate published findings"

- NATIONAL RESEARCH COUNCIL OF *THE NATIONAL ACADEMIES*

"Data sharing is essential for expedited translation of research results into knowledge, products, and procedures to improve human Health"

- NATIONAL INSTITUTES OF HEALTH STATEMENT ON SHARING RESEARCH DATA

Sharing : principles and recommendations

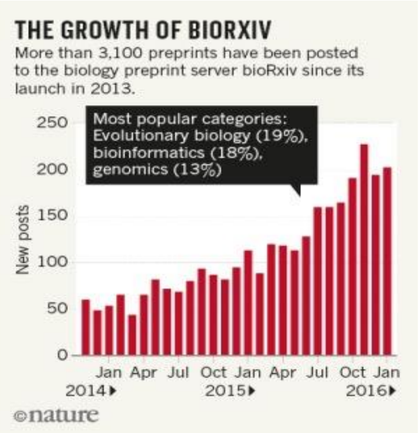
- In publication, include all data necessary to support claims and enable others to replicate them.
- If applicable, use publicly accessible repositories by the time of publication.
- Anticipate requests and indicate how to obtain them.
- If integral material is patented, provide under a licence for research use.
- Provider of material cannot demand exclusive license to commercialize a new substance from the material or require collaboration or coauthorship
- Transfer should occur within 60 days. If this doesn’t happen, the requestor may contact the journal, then the author’s institution or funder.
- Recipients should acknowledge the source in publications

bioRχiv beta

THE PREPRINT SERVER FOR BIOLOGY

Search

Advanced Search



nature.com > scientific data > policies > recommended data repositories

a natureresearch journal

SCIENTIFIC DATA

Search | E-alert | Submit | Login

Recommended Data Repositories

Scientific Data mandates the release of datasets accompanying our Data Descriptors, but we do not ourselves host data. Instead, we ask authors to submit datasets to an appropriate public data repository. Data should be submitted to discipline-specific, community-recognized repositories where possible, or to generalist repositories if no suitable community resource is available.

Repositories included on this page have been evaluated to ensure that they meet our requirements for data access, preservation and stability. Please be aware, however, that some repositories on this page may only accept data from those funded by specific sources, or may charge for hosting data. Please ensure you are aware of any deposition policies for your chosen repository. If your repository of choice is not listed please see our guidelines for suggesting additional repositories.