R: A Hitchhikers Guide to Reproducible Research

- Hello

Brendan Palmer,

Clinical Research Facility - Cork &

School of Public Health

@B_A_Palmer

HRB Clinical Research Facility Cork

University College Cork, Ireland

Coláiste na hOllscoile Corcaigh

School of Public Health

Reproducible or replicable

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

Reproducible or replicable

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

Reproducible or replicable

<table>
<tr><td colspan="2" rowspan="2"></td><th colspan="2">Data</th></tr>
<tr><th>Same</th><th>Different</th></tr>
<tr><th rowspan="2">Analysis</th><th>Same</th><td>Replicable</td><td>Reproducible</td></tr>
<tr><th>Different</th><td>Robust</td><td>Generalisable</td></tr>
</table>

COMMENTARY

Scientists behaving badly

To protect the integrity of science, we must look beyond falsification, fabrication and plagiarism, to a wider range of questionable research practices, argue **Brian C. Martinson, Melissa S. Anderson** and **Raymond de Vries**.

Table 1 | Percentage of scientists who say that they engaged in the behaviour listed within the previous three years (*n* = 3,247)

Top ten behaviours	All	Mid-career	Early-career
1. Falsifying or 'cooking' research data	0.3	0.2	0.5
2. Ignoring major aspects of human-subject requirements	0.3	0.3	0.4
3. Not properly disclosing involvement in firms whose products are based on one's own research	0.3	0.4	0.3
4. Relationships with students, research subjects or clients that may be interpreted as questionable	1.4	1.3	1.4
5. Using another's ideas without obtaining permission or giving due credit	1.4	1.7	1.0
6. Unauthorized use of confidential information in connection with one's own research	1.7	2.4	0.8 ***
7. Failing to present data that contradict one's own previous research	6.0	6.5	5.3
8. Circumventing certain minor aspects of human-subject requirements	7.6	9.0	6.0 **
9. Overlooking others' use of flawed data or questionable interpretation of data	12.5	12.2	12.8
10. Changing the design, methodology or results of a study in response to pressure from a funding source	15.5	20.6	9.5 ***
Other behaviours			
11. Publishing the same data or results in two or more publications	4.7	5.9	3.4 **
12. Inappropriately assigning authorship credit	10.0	12.3	7.4 ***
13. Withholding details of methodology or results in papers or proposals	10.8	12.4	8.9 **
14. Using inadequate or inappropriate research designs	13.5	14.6	12.2
15. Dropping observations or data points from analyses based on a gut feeling that they were inaccurate	15.3	14.3	16.5
16. Inadequate record keeping related to research projects	27.5	27.7	27.3

Standing on the shoulders of giants?

Research Article

Psychological Science
23(5) 524–532

DOI: 10.1177/0956797611430953
http://pss.sagepub.com
SAGE

Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling

Leslie K. John[1], George Loewenstein[2], and Drazen Prelec[3]
[1]Marketing Unit, Harvard Business School; [2]Department of Social & Decision Sciences, Carnegie Mellon University; and [3]Sloan School of Management and Departments of Economics and Brain & Cognitive Sciences, Massachusetts Institute of Technology

Abstract
Cases of clear scientific misconduct have received significant media attention recently, but less flagrantly questionable research practices may be more prevalent and, ultimately, more damaging to the academic enterprise. Using an anonymous elicitation format supplemented by incentives for honest reporting, we surveyed over 2,000 psychologists about their involvement in questionable research practices. The impact of truth-telling incentives on self-admissions of questionable research practices was positive, and this impact was greater for practices that respondents judged to be less defensible. Combining three different estimation methods, we found that the percentage of respondents who have engaged in questionable practices was surprisingly high. This finding suggests that some questionable practices may constitute the prevailing research norm.

241 shades of grey

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Full Length Articles

The secret lives of experiments: Methods reporting in the fMRI literature

Joshua Carp

University of Michigan, Department of Psychology, 530 Church Street, Ann Arbor, MI, 48109, USA

A R T I C L E I N F O

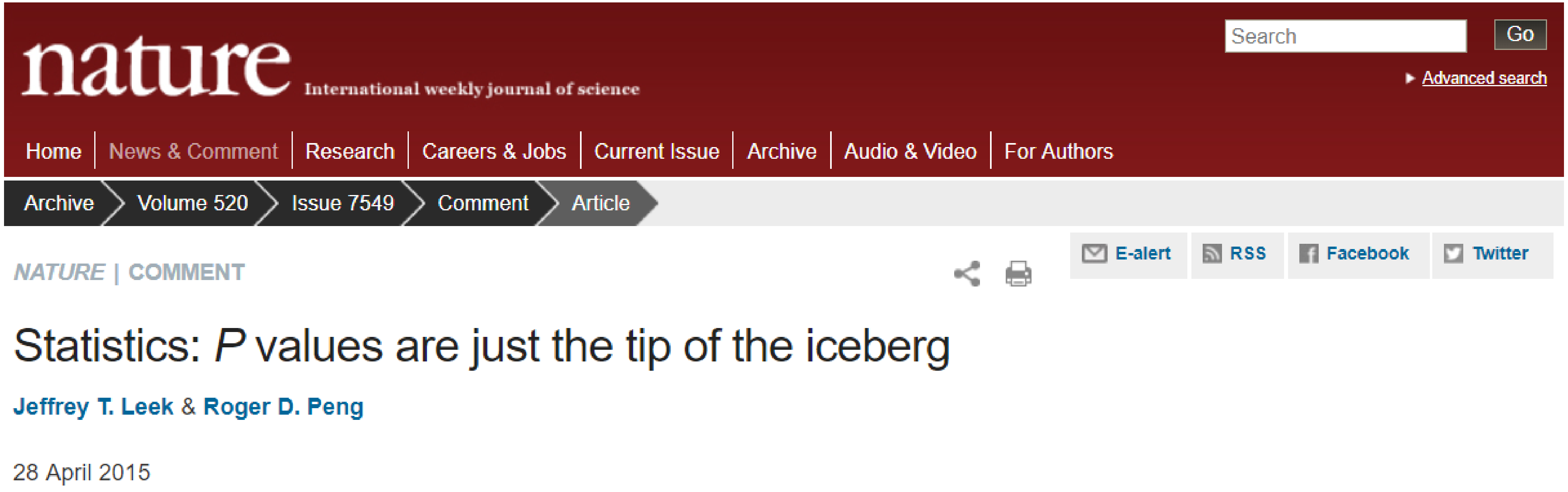
Article history:
Accepted 3 July 2012
Available online 10 July 2012

Keywords:
fMRI
Methods reporting
Reproducibility
Experimental design
Analysis methods
Statistical power

A B S T R A C T

Replication of research findings is critical to the progress of scientific understanding. Accordingly, most scientific journals require authors to report experimental procedures in sufficient detail for independent researchers to replicate their work. To what extent do research reports in the functional neuroimaging literature live up to this standard? The present study evaluated methods reporting and methodological choices across 241 recent fMRI articles. Many studies did not report critical methodological details with regard to experimental design, data acquisition, and analysis. Further, many studies were underpowered to detect any but the largest statistical effects. Finally, data collection and analysis methods were highly flexible across studies, with nearly as many unique analysis pipelines as there were studies in the sample. Because the rate of false positive results is thought to increase with the flexibility of experimental designs, the field of functional neuroimaging may be particularly vulnerable to false positives. In sum, the present study documented significant gaps in methods reporting among fMRI studies. Improved methodological descriptions in research reports would yield significant benefits for the field.

p-values should not define a study



nature International weekly journal of science

Search Go

Advanced search

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive > Volume 520 > Issue 7549 > Comment > Article

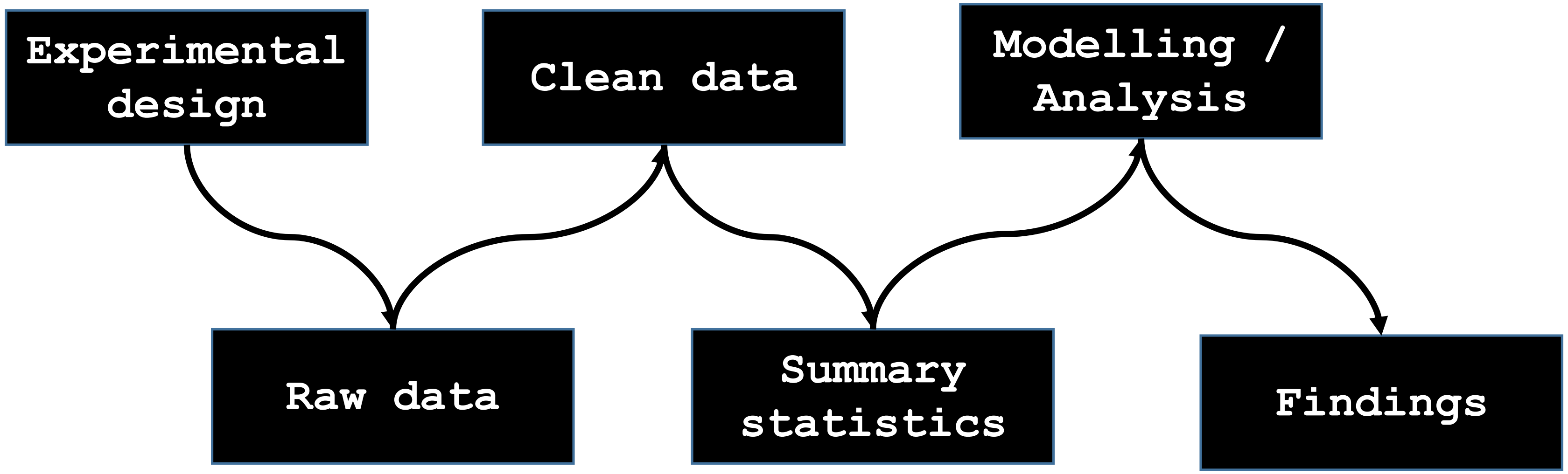
E-alert RSS Facebook Twitter

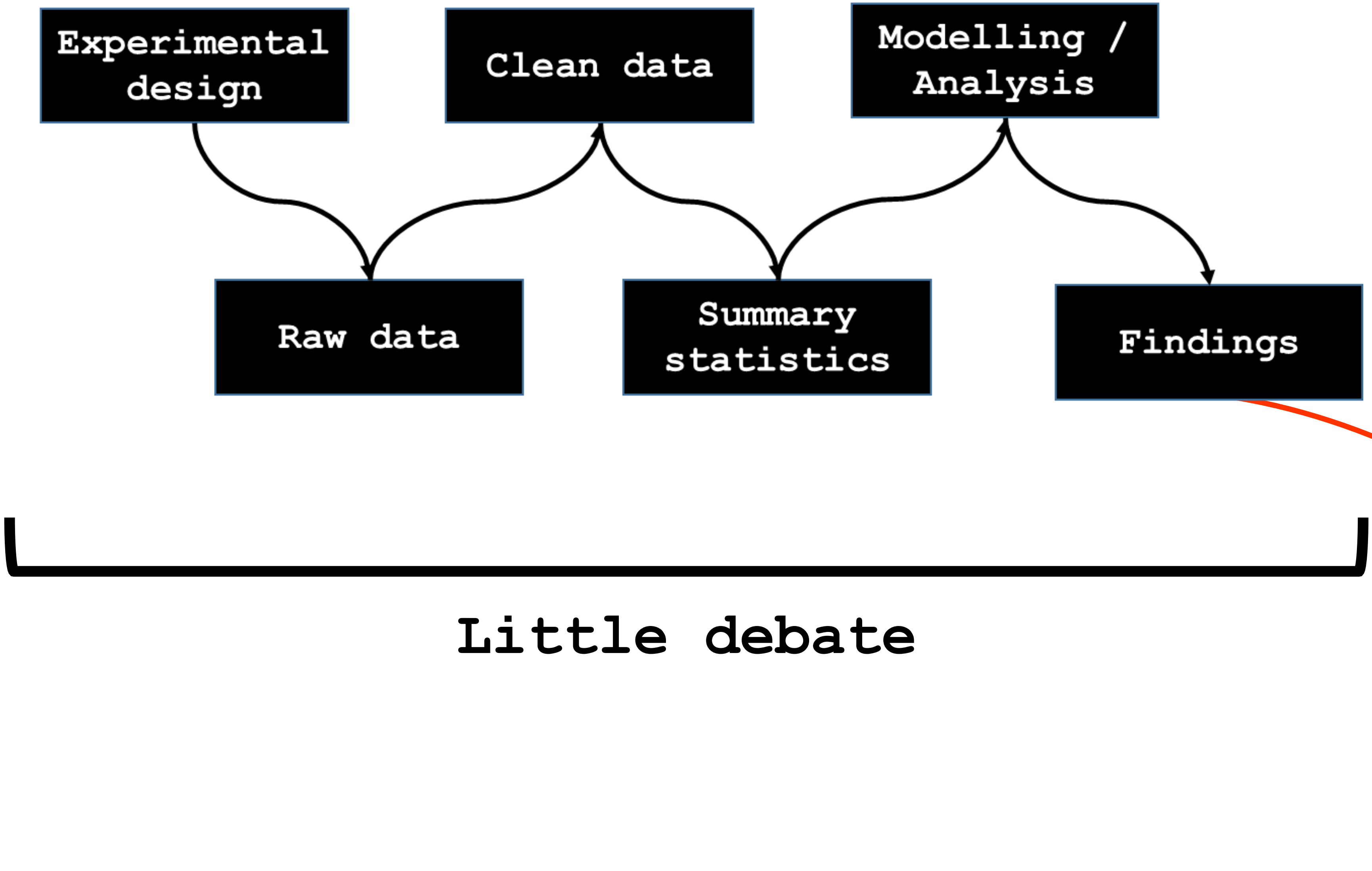
NATURE | COMMENT

Statistics: *P* values are just the tip of the iceberg

Jeffrey T. Leek & Roger D. Peng

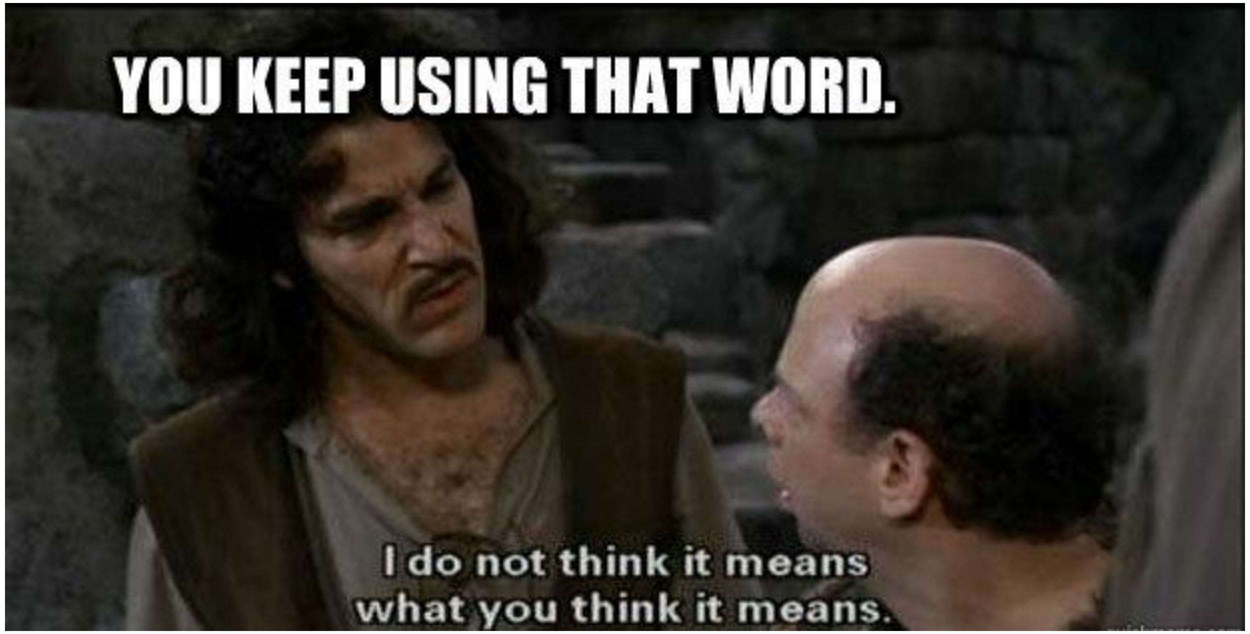
28 April 2015

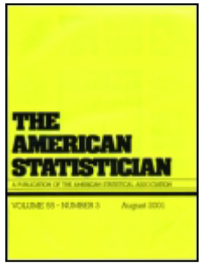



Experimental design
Clean data
Modelling / Analysis
Raw data
Summary statistics
Findings
Little debate

Extreme scrutiny

Significant





Taylor & Francis
Taylor & Francis Group

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on *p*-Values: Context, Process, and Purpose

Ronald L. Wasserstein & Nicole A. Lazar

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

CrossMark

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]

Received: 9 April 2016 / Accepted: 9 April 2016 / Published online: 21 May 2016

Abstract Misinterpretation and abuse of statistical tests, confidence intervals, and statistical power have been decried for decades, yet remain rampant. A key problem is that there are no interpretations of these concepts that are at once simple, intuitive, correct, and foolproof. Instead, correct use and interpretation of these statistics requires an attention to detail which seems to tax the patience of working scientists. This high cognitive demand has led to an epidemic of shortcut definitions and interpretations that are simply wrong, sometimes disastrously so—and yet these misinterpretations dominate much of the scientific literature. In light of this problem, we provide definitions and a discussion of basic statistics that are more general and critical than typically found in traditional introductory expositions. Our goal is to provide a resource for instructors, researchers, and consumers of statistics whose knowledge of statistical theory and technique may be limited but who wish to avoid and spot misinterpretations. We emphasize how violation of often unstated analysis protocols (such as selecting analyses for presentation based on the *P* values they produce) can lead to small *P* values even if the declared test hypothesis is correct, and can lead to large *P* values even if that hypothesis is incorrect. We then provide an explanatory list of 25 misinterpretations of *P* values, confidence intervals, and power. We conclude with guidelines for improving statistical interpretation and reporting.

Editor's note This article has been published online as supplementary material with an article of Wasserstein RL, Lazar NA. The ASA's statement on p-values: context, process and purpose. The American Statistician 2016.

Albert Hofman, Editor-in-Chief EJE.

✉ Sander Greenland
lesdomes@ucla.edu

Stephen J. Senn
stephen.senn@lih.lu

John B. Carlin
john.carlin@mcri.edu.au

Charles Poole
cpoole@unc.edu

Steven N. Goodman
steve.goodman@stanford.edu

Douglas G. Altman
doug.altman@csm.ox.ac.uk

[1] Department of Epidemiology and Department of Statistics, University of California, Los Angeles, CA, USA

[2] Competence Center for Methodology and Statistics, Luxembourg Institute of Health, Strassen, Luxembourg

[3] RTI Health Solutions, Research Triangle Institute, Research Triangle Park, NC, USA

[4] Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, School of Population Health, University of Melbourne, Melbourne, VIC, Australia

[5] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

[6] Meta-Research Innovation Center, Departments of Medicine and of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA

[7] Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

Everything is awesome!

thebmj | Research | Education | News & Views | Campaigns | Archive | For authors | Jobs | Hosted | Search

Research » Christmas 2015: The Publication Game

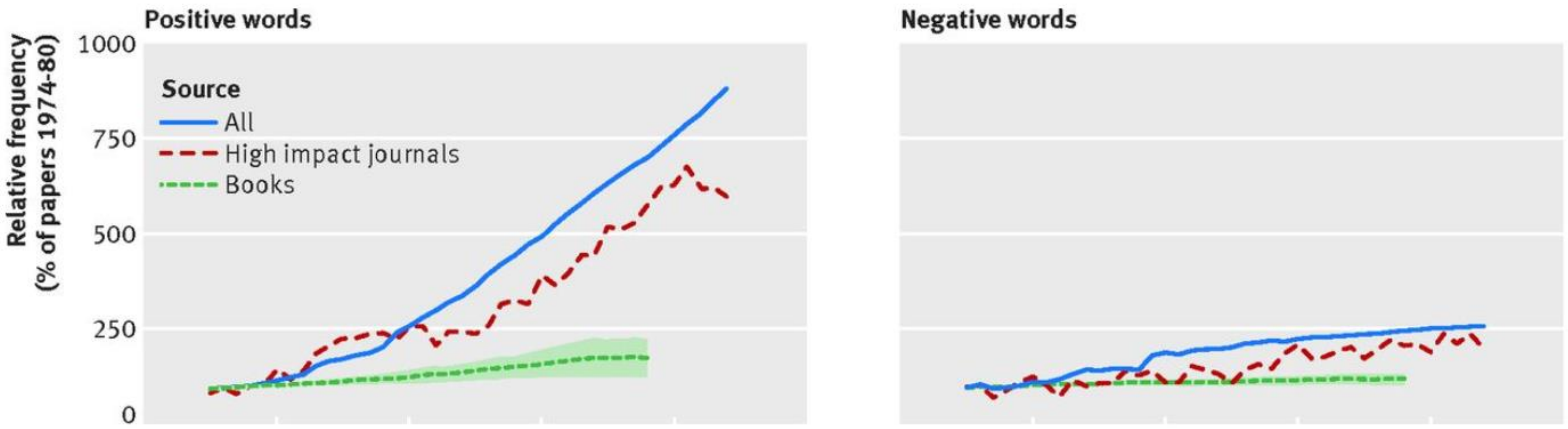
Open access

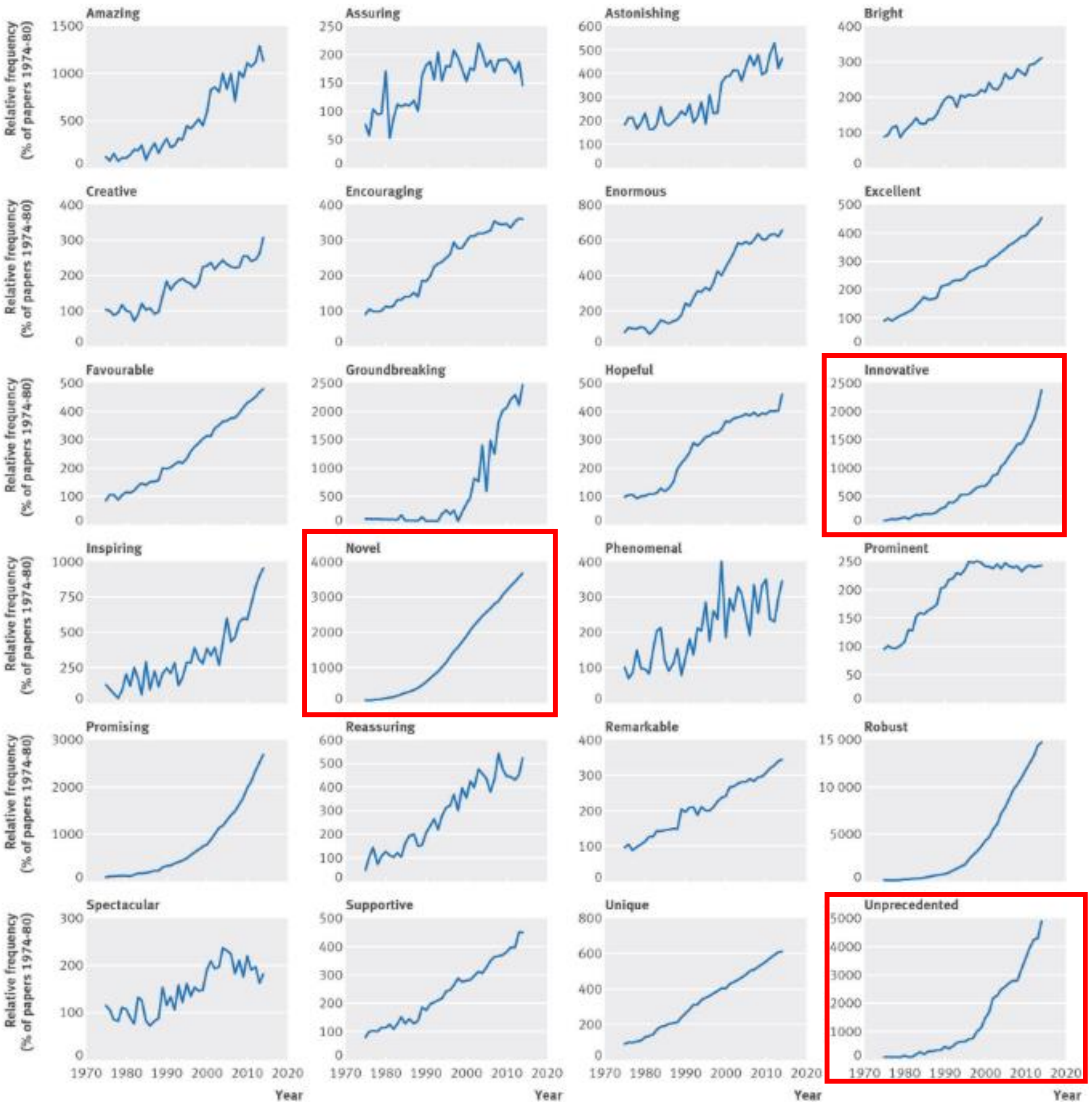
Tweet | Like 17

Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis

BMJ 2015 ; 351 doi: https://doi.org/10.1136/bmj.h6467 (Published 14 December 2015)
Cite this as: *BMJ* 2015;351:h6467

See other articles in issue 8038

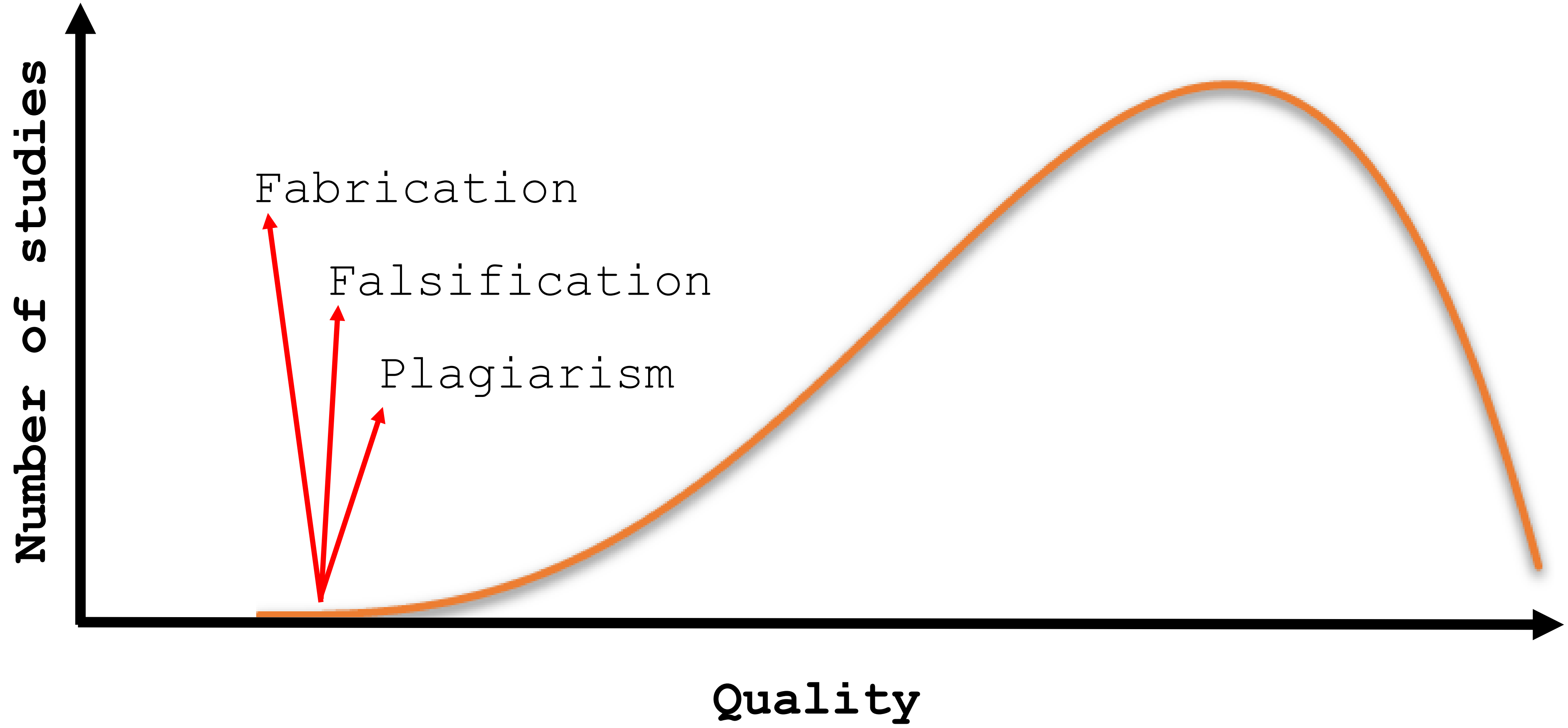




Relative frequencies of individual words as used in PubMed between 1975 and 2014

Vinkers et al., 2015

Today

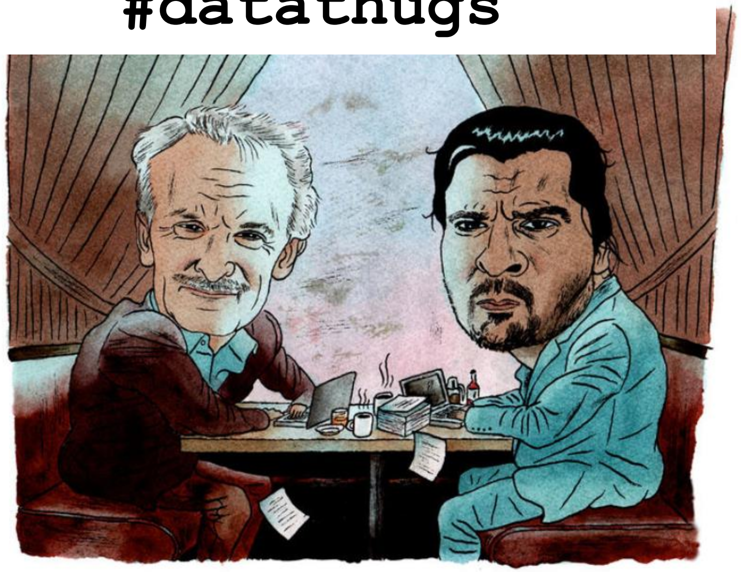

Number of studies
Fabrication
Falsification
Plagiarism
Quality

Don’t do what Donny Dont does!

“In short, peer review misses all the hard stuff, and a worrying amount of the easy stuff”

James Heathers,
Northwestern University

#datathugs



Brian Wansink: The grad student who never said no

“Every day we would scratch our heads, ask "Why," and come up with another way to reanalyze the data with yet another set of plausible hypotheses. Eventually we started discovering solutions”

Install the Chrome plugin PubPeer


NCBI Resources How To Sign in to NCBI

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed | Is the Power Threshold of 0.8 Applicable to Surgical Science? | Search

Create RSS Create alert Advanced Help

Format: Abstract

Send to

See 1 citation found by title matching your search:

J Surg Res. 2019 Apr 26;241:235-239. doi: 10.1016/j.jss.2019.03.062. [Epub ahead of print]

23 comments on PubPeer (by: Andrew D. Althouse, Thom Baguley, Guillaume A. Rousselet, Timothy Feeney, Paul M Brown, Frank E. Harrell, David Nunan, Samantha R. Seals, Raj Mehta, Yevgeniy Feyman, Ionomidotis Irregularis, Andrew Gelman, Aleksi Reito, Daniel E. Leisman, Pavlos Msaouel, Ryan Miller, Maarten Van Smeden, Zad Rafi Chow)

Is the Power Threshold of 0.8 Applicable to Surgical Science?-Empowering the Underpowered Study.

Bababekov YJ[1], Hung YC[2], Hsu YT[2], Udelsman BV[2], Mueller JL[2], Lin HY[2], Stapleton SM[2], Chang DC[2].

Author information

Abstract

BACKGROUND: Many articles in the surgical literature were faulted for committing type 2 error, or concluding no difference when the study was "underpowered". However, it is unknown if the current power standard of 0.8 is reasonable in surgical science.

Full text links

ELSEVIER FULL-TEXT ARTICLE

Save items

Similar articles

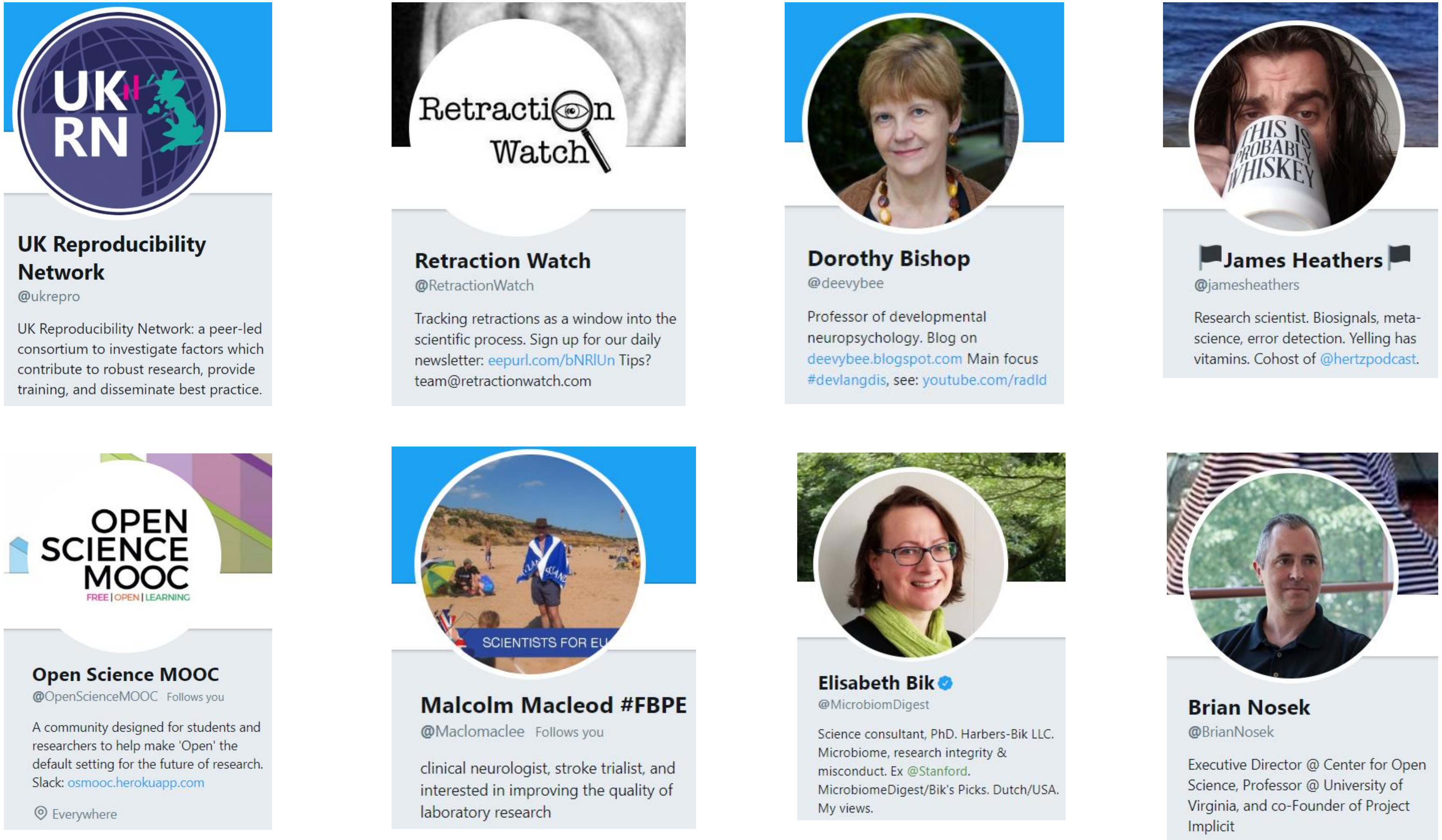
Review Interventions to Prevent Falls in Community-[[Agency for Healthcare Research...]

Review Is There Truly "No Significant Difference"? Under| [J Bone Joint Surg Am. 2015]

Review Randomized controlled trials and neurosurgery: the ideal fit or : [J Neurosurg. 2016]

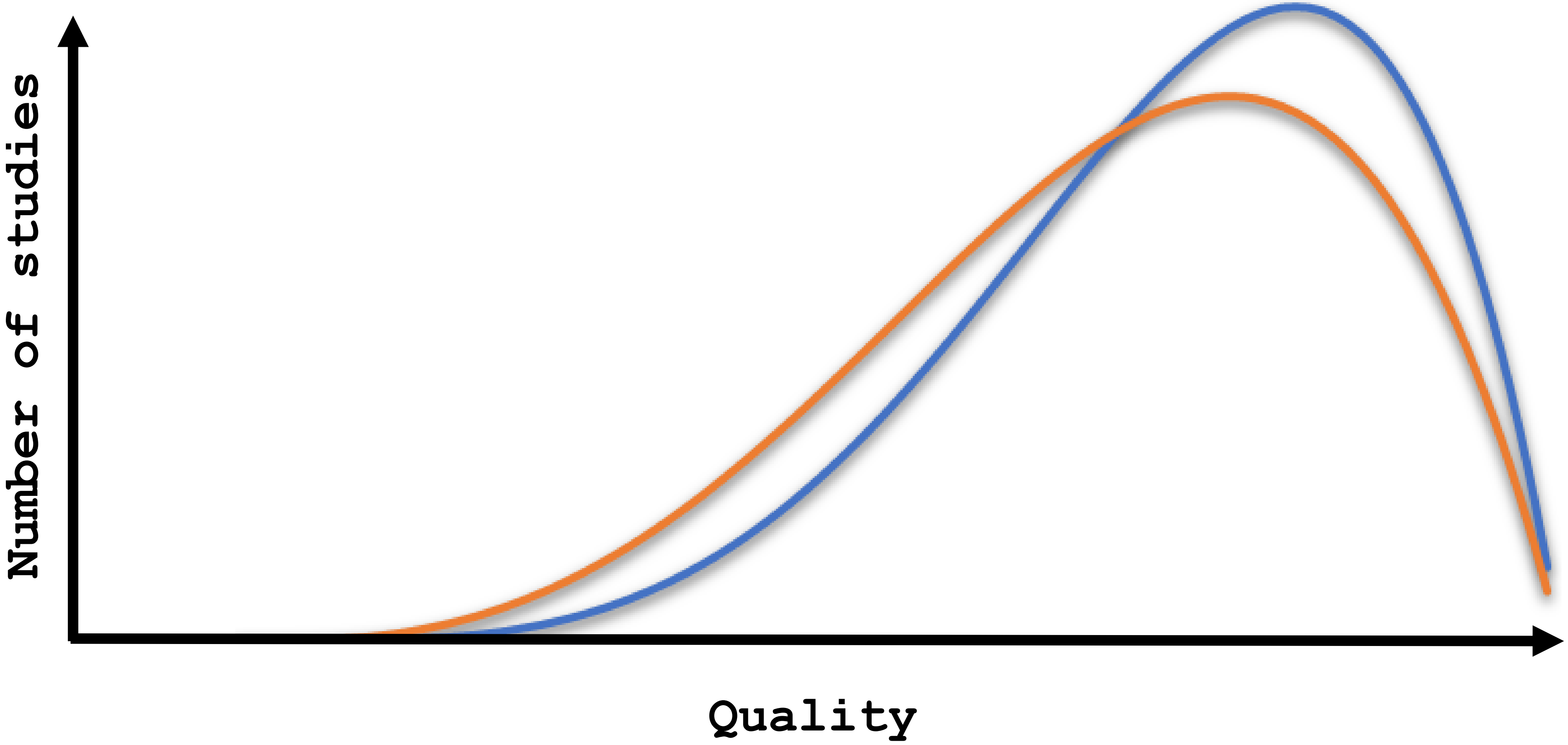
Review Low-Dose Aspirin for the Prevention of Morbidity anc [Agency for Healthcare Research...]

Twitter



UK RN
UK Reproducibility Network
@ukrepro
UK Reproducibility Network: a peer-led consortium to investigate factors which contribute to robust research, provide training, and disseminate best practice.
Retraction Watch
Retraction Watch
@RetractionWatch
Tracking retractions as a window into the scientific process. Sign up for our daily newsletter: eepurl.com/bNRlUn Tips? team@retractionwatch.com
Dorothy Bishop
@deevybee
Professor of developmental neuropsychology. Blog on deevybee.blogspot.com Main focus #devlangdis, see: youtube.com/radld
THIS IS PROBABLY WHISKEY
James Heathers
@jamesheathers
Research scientist. Biosignals, meta-science, error detection. Yelling has vitamins. Cohost of @hertzpodcast.
OPEN SCIENCE MOOC
FREE | OPEN | LEARNING
Open Science MOOC
@OpenScienceMOOC Follows you
A community designed for students and researchers to help make 'Open' the default setting for the future of research. Slack: osmooc.herokuapp.com
Everywhere
SCIENTISTS FOR EU
Malcolm Macleod #FBPE
@Maclomaclee Follows you
clinical neurologist, stroke trialist, and interested in improving the quality of laboratory research
Elisabeth Bik
@MicrobiomDigest
Science consultant, PhD. Harbers-Bik LLC. Microbiome, research integrity & misconduct. Ex @Stanford. MicrobiomeDigest/Bik's Picks. Dutch/USA. My views.
Brian Nosek
@BrianNosek
Executive Director @ Center for Open Science, Professor @ University of Virginia, and co-Founder of Project Implicit

Tomorrow



Who benefits most from reproducibility?

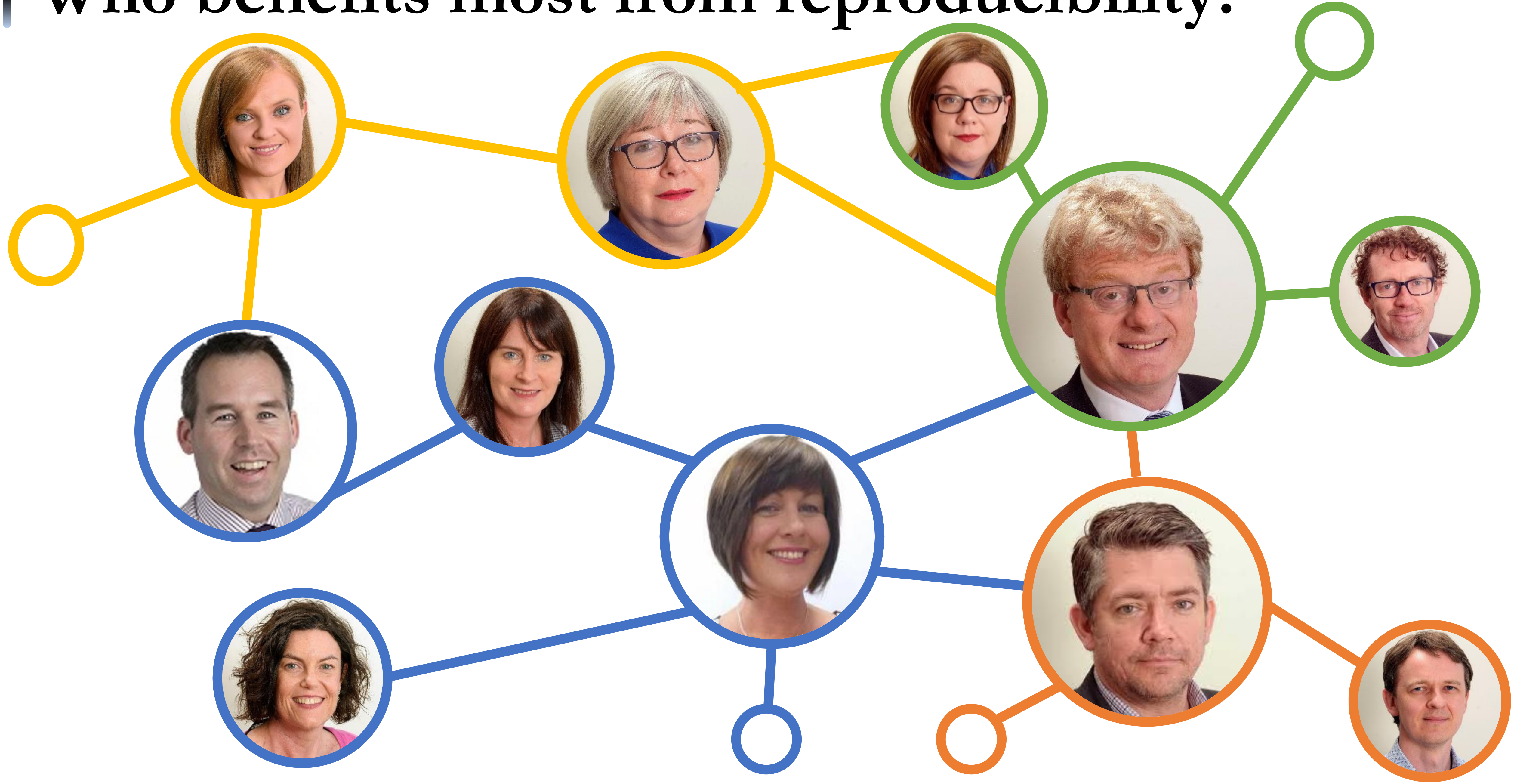
Casey Greene
@GreeneScientist

Follow

Reproducibility is important because the you of 3 months ago is terrible at answering email! - @tracykteal at #2016dssummit

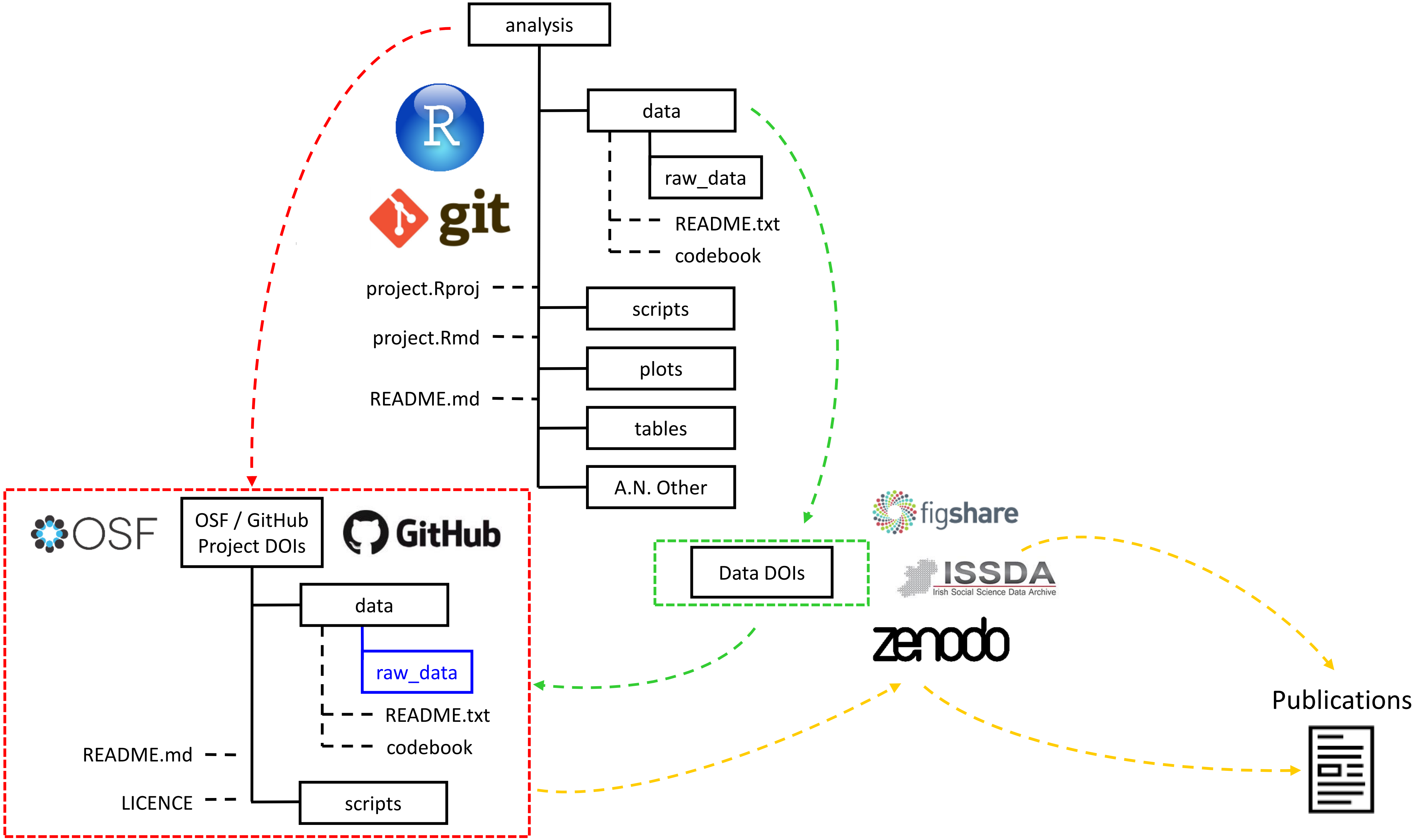
1:17 PM - 26 Oct 2016 from Manhattan, NY

Who benefits most from reproducibility?

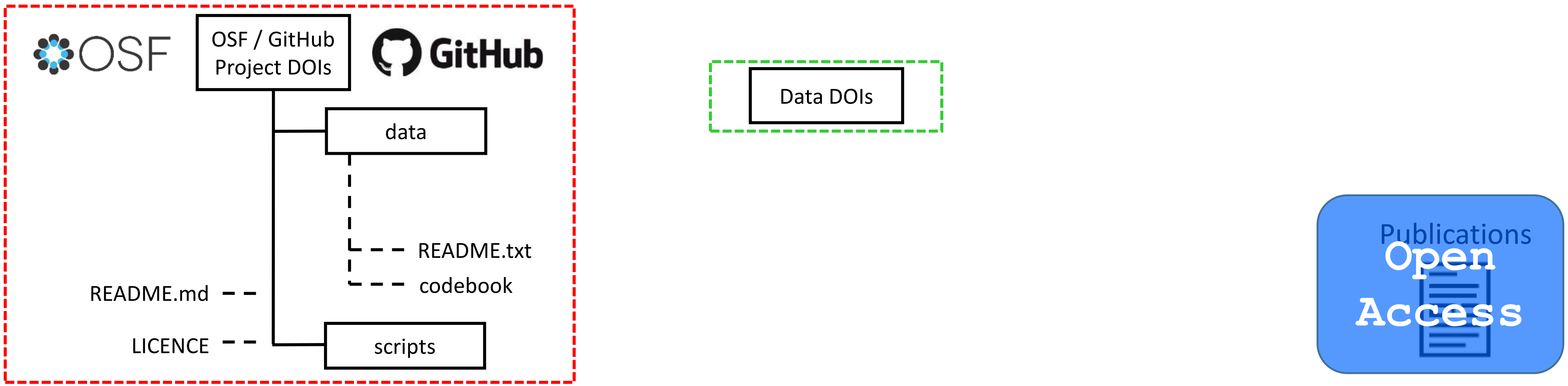


Where to begin…

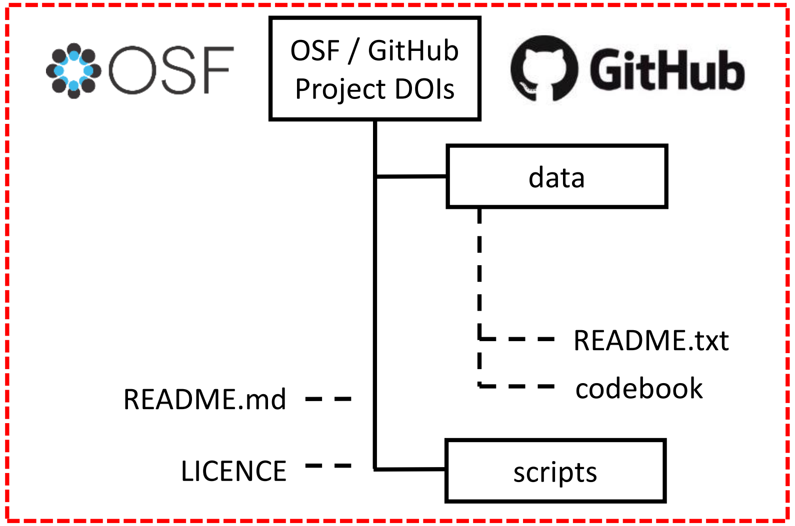
What does this allow us to do?



What does this allow us to do?



What does this allow us to do?



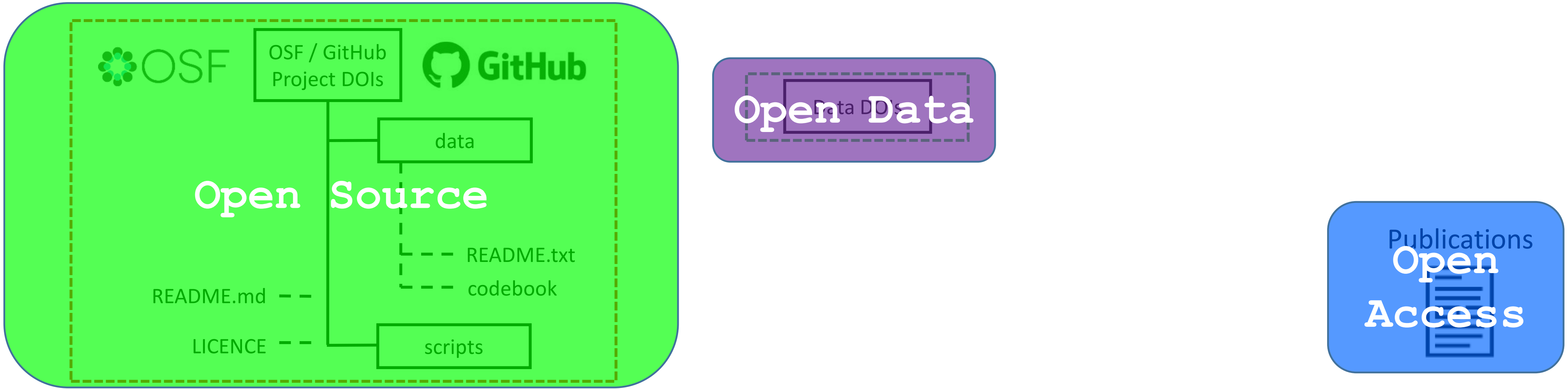
Data DOIs

Open Data

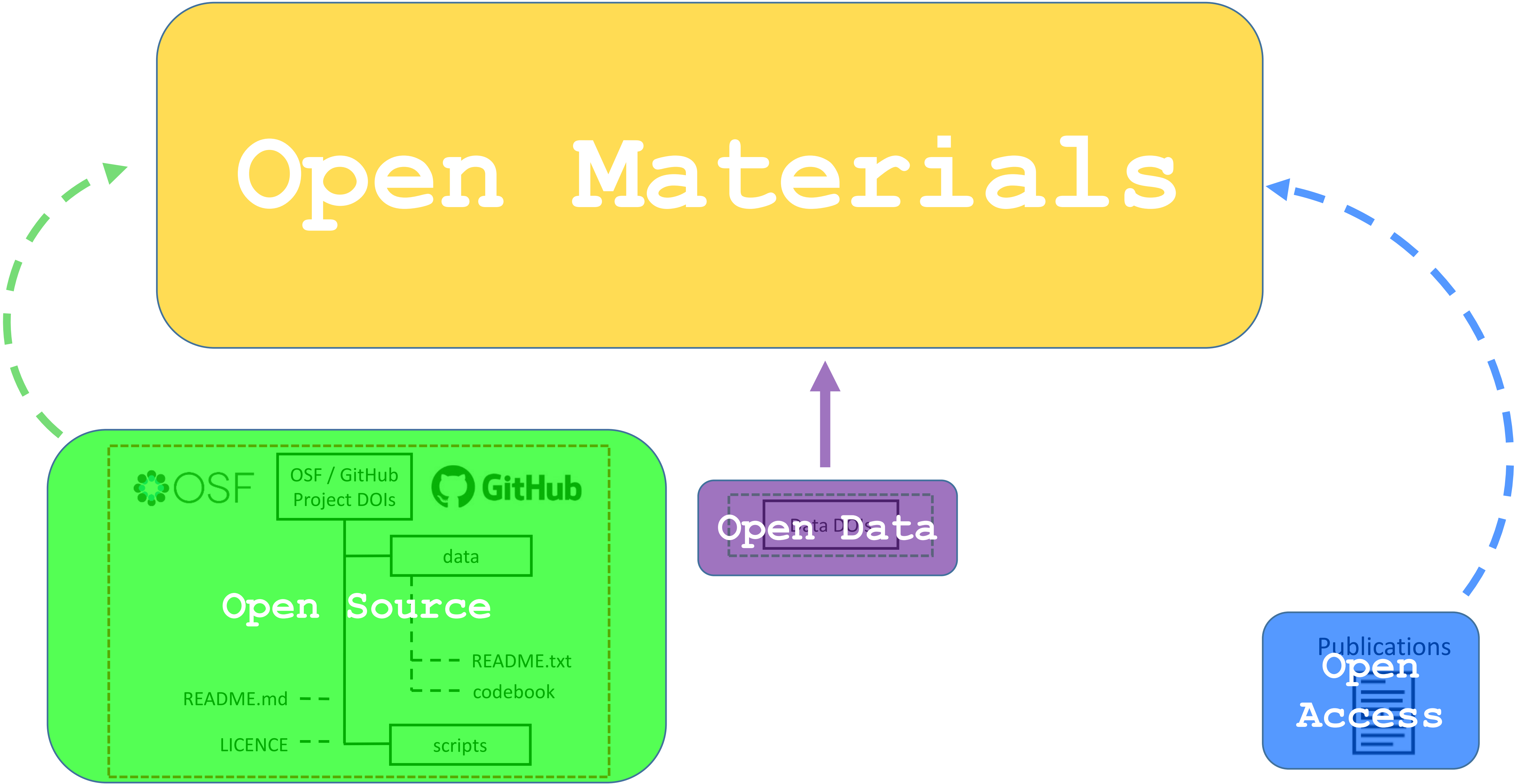
Publications

Open Access

What does this allow us to do?



What does this allow us to do?



The butterfly has started flapping its wings

Plan S
Making full and immediate Open Access a reality

Why Plan S | 10 Principles | Funders & support | Implementation | About | Contact

"After 1 January 2020 scientific publications on the results from research funded by public grants provided by national and European research councils and funding bodies, must be published in compliant Open Access Journals or on compliant Open Access Platforms."

EUROPEAN COMMISSION
Directorate-General for Research & Innovation

H2020 Programme

Guidelines on

FAIR Data Management in Horizon 2020

IRISH RESEARCH COUNCIL
An Chomhairle um Thaighde in Éirinn

f) Data Management Plan – 2 pages max.
- Applicants should address the following issues:
  - What standards will be applied?
  - How will data be exploited and/or shared/made accessible for verification and reuse? If data cannot be made available, why?
  - How data will be curated & preserved?
  - If applicable, how does the applicant plan to make the research data FAIR (findable, accessible, interoperable and reusable).

HRB Health Research Board

Funding | Data collections & evidence | Publications | Success stories | News | About

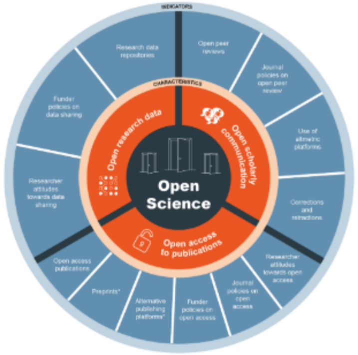
Home > Funding > Policies and principles > Open Research

Funding schemes
EU funding support
Manage a grant
Funding awarded
Evaluation
GDPR guidance for researchers
Policies and principles
EU legislation
Gender
Good research practice
Open Research

Open Research

The HRB is committed to ensuring that its funded research is open, accessible and usable, so it can have the greatest possible impact.

There is a fundamental shift across Europe towards making research more transparent, collaborative, accessible and efficient. This Open Science movement is a strategic priority for the European Commission in research and innovation policy and an EU high-level Expert Group, the Open Science Policy Platform (OSPP 2016-2018) has been established to consider key implementation areas.



Science Foundation Ireland SFI For what's next

Funding | Engagement | Events | Research News | SFI Research Centres

Science Foundation Ireland joins DORA

14th February 2019, Dublin – Science Foundation Ireland has become a signatory to the San Francisco Declaration of Research Assessment (DORA), making a formal commitment to assessing the quality and impact of research through means other than journal impact factors.

The movement towards FAIR data

F + **A** + **I** = **R**

Findable **Accessible** **Interoperable** **Reusable** (Replicable, Reproducible)

www.nature.com/scientificdata

SCIENTIFIC DATA

Amended: Addendum

OPEN

SUBJECT CATEGORIES
» Research data
» Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*#

Three main take home messages

- **FAIR is not prescriptive**
- **Varying levels of FAIR**
- **Use of existing tools**

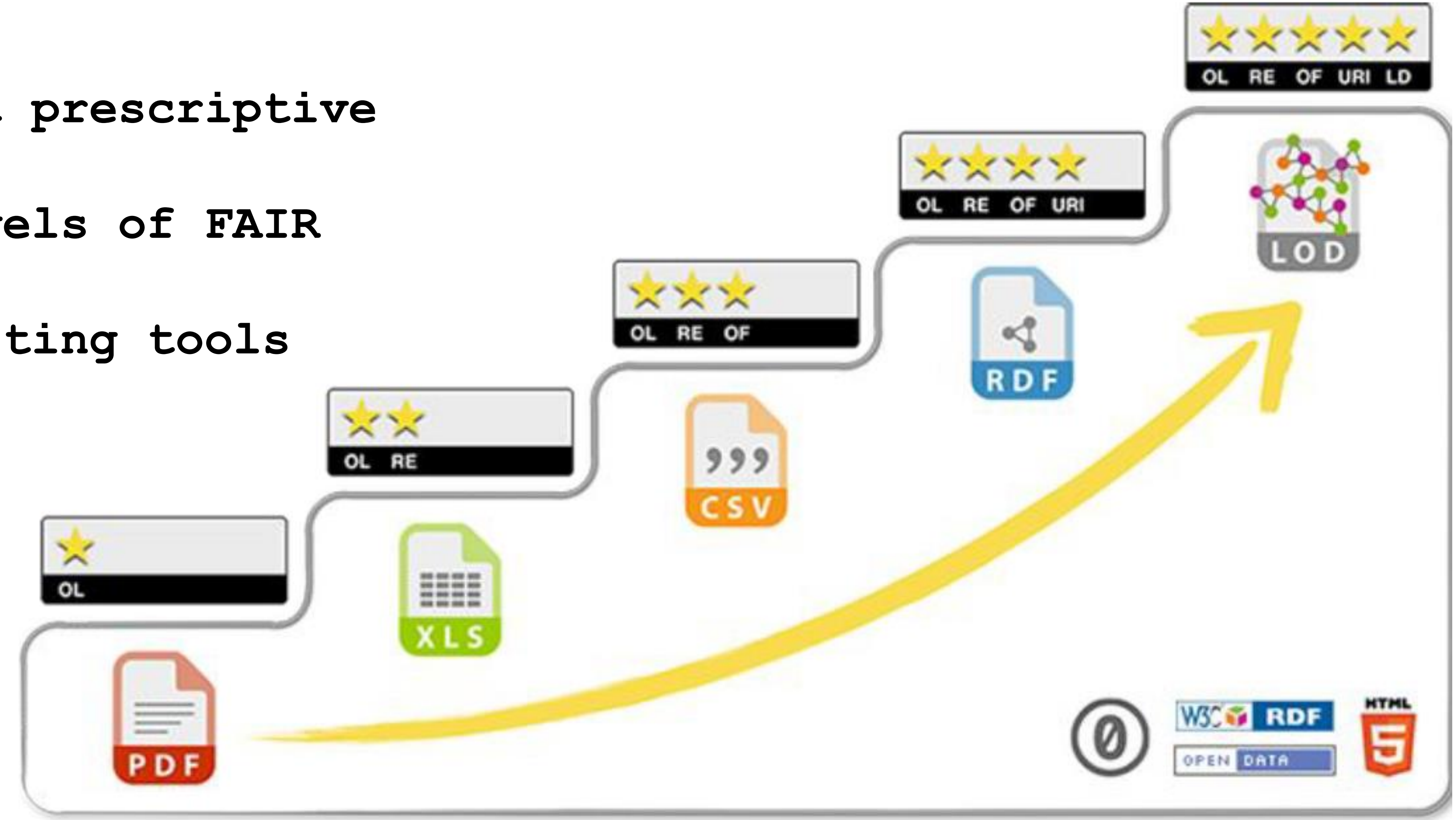


Image: 5 ★ OPEN DATA

Ask questions

Live feedback and resources

HackMD New! Name your version!

CHANGED A DAY AGO

Watch

R: A Hitchhikers Guide to Reproducible Research

University College Cork Summer 2019

Instructors

Brendan Palmer
Darren Dahly

Course notes

- GitHub repository: https://github.com/bapalmer/R-A_Hitchhikers_Guide_to_Reproducible_Research

Who's here

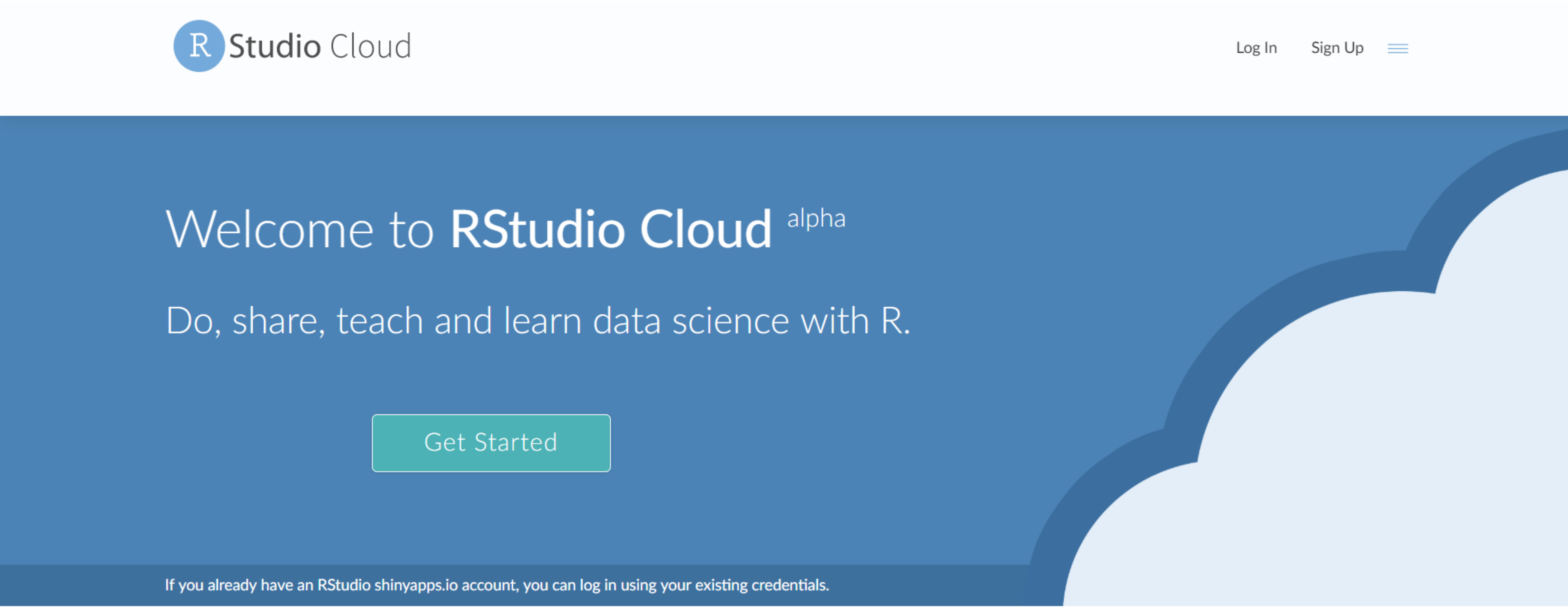
If you'd like to, please let us know who you are. This hackmd notepad is public, so please don't share any information you don't feel comfortable making publicly available.

Name / Group / Twitter

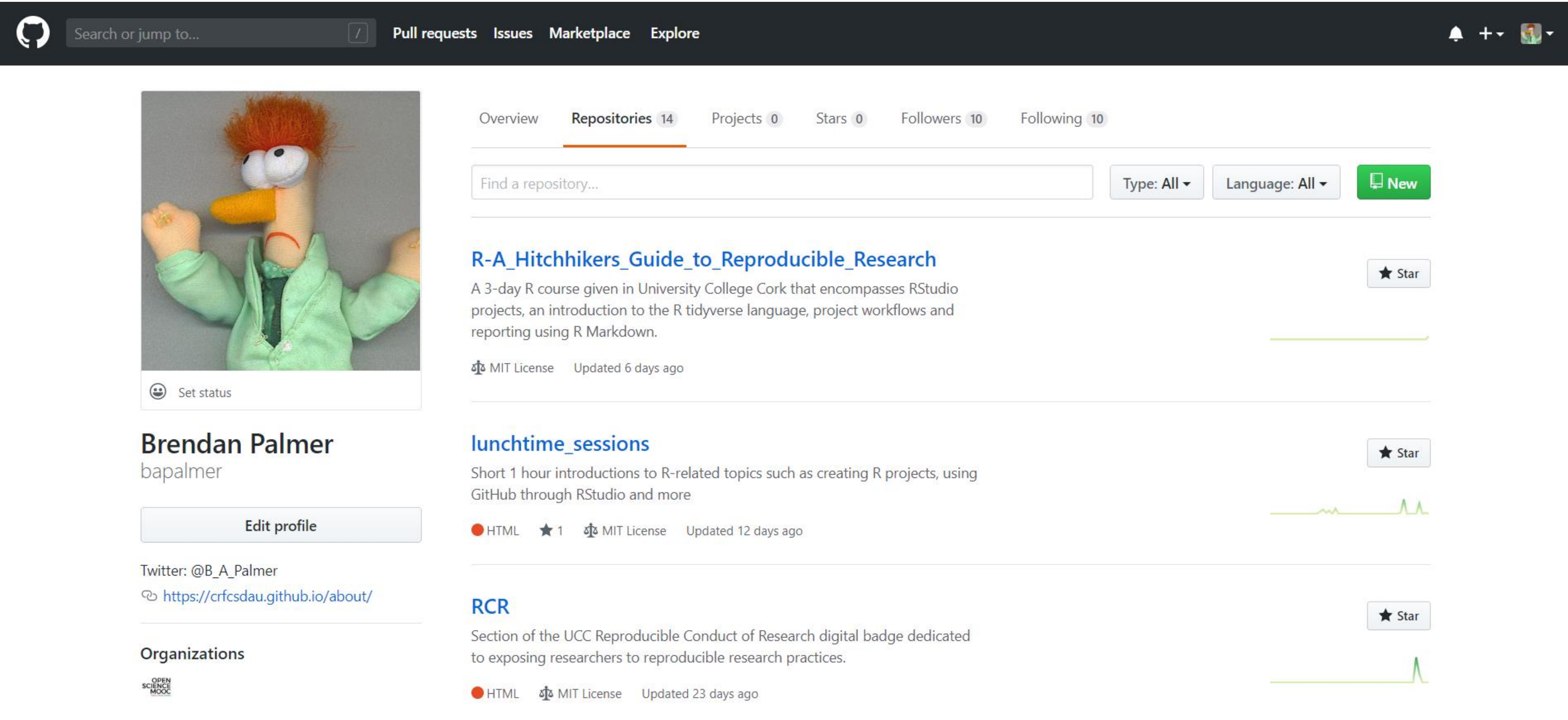
- Brendan Palmer / CRF-C | School of Public Health / @B_A_Palmer

Expand all
Back to top
Go to bottom

RStudio Cloud



GitHub



Search or jump to...
Pull requests
Issues
Marketplace
Explore
Overview
Repositories 14
Projects 0
Stars 0
Followers 10
Following 10
Find a repository...
Type: All
Language: All
New
Set status
Brendan Palmer
bapalmer
Edit profile
Twitter: @B_A_Palmer
https://crfcsdau.github.io/about/
Organizations
R-A_Hitchhikers_Guide_to_Reproducible_Research
A 3-day R course given in University College Cork that encompasses RStudio projects, an introduction to the R tidyverse language, project workflows and reporting using R Markdown.
MIT License
Updated 6 days ago
Star
lunchtime_sessions
Short 1 hour introductions to R-related topics such as creating R projects, using GitHub through RStudio and more
HTML
1
MIT License
Updated 12 days ago
Star
RCR
Section of the UCC Reproducible Conduct of Research digital badge dedicated to exposing researchers to reproducible research practices.
HTML
MIT License
Updated 23 days ago
Star

Central theme of this workshop

Workshop content