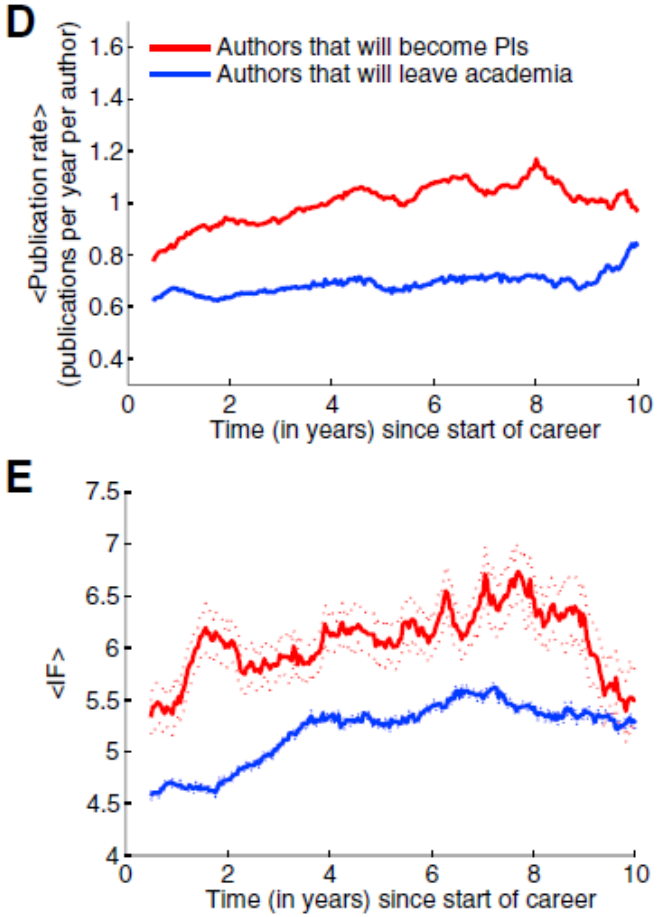
Improving the research method

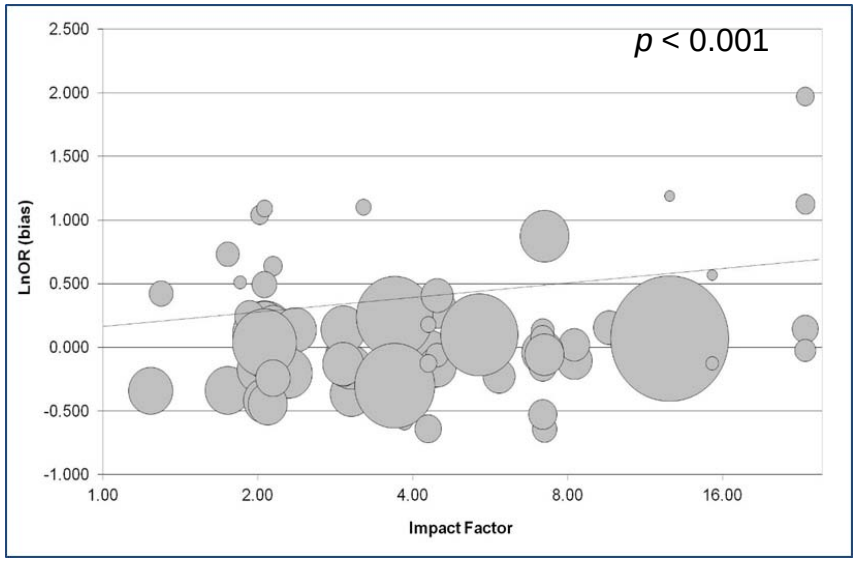
Key issues with the current model

Key issues with the current model

- Rewards (false) positive results over robust methods
- Poor methods increase (false) positive results
  - Low statistical power
  - Lack of control for bias
  - Questionable research practices
- Need to align incentives for career progression with robust methods

Rewarding positive results





Van Dijk et al (2014) Current Biology, 24, R516–R517

Munafò et al (2009). Molecular Psychiatry, 14, 119-120.

Competition increases bias

Figure 1: New faculty positions versus new PhDs.

From
The missing piece to changing the university culture
Maximiliaan Schillebeeckx, Brett Maricque & Cory Lewis
Nature Biotechnology **31**, 938–941 (2013) | doi:10.1038/nbt.2706
Published online 08 October 2013

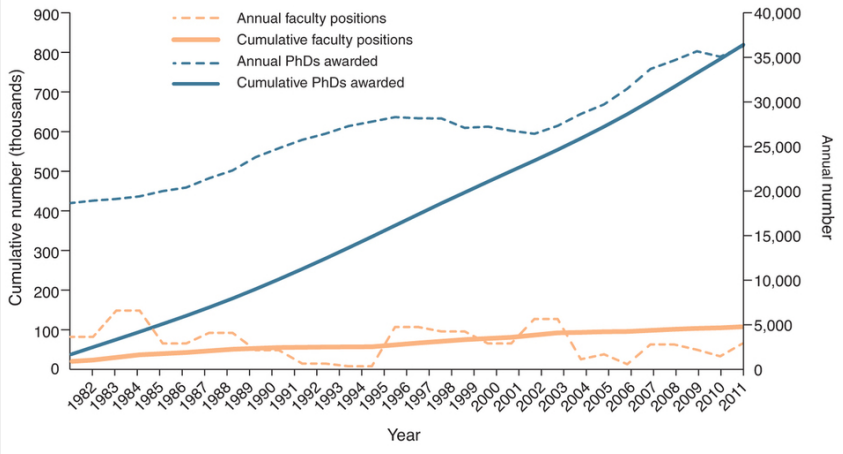
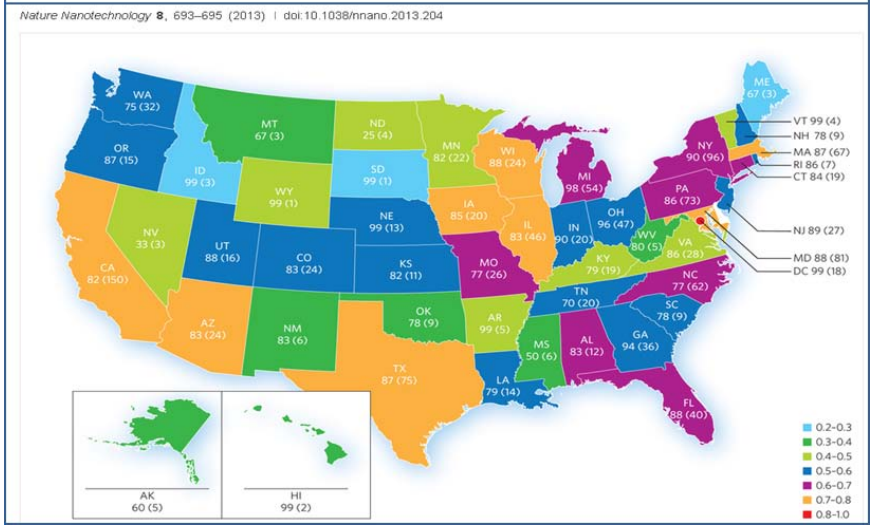


Figure 1: A competitive research environment might increase the positive-outcome bias.

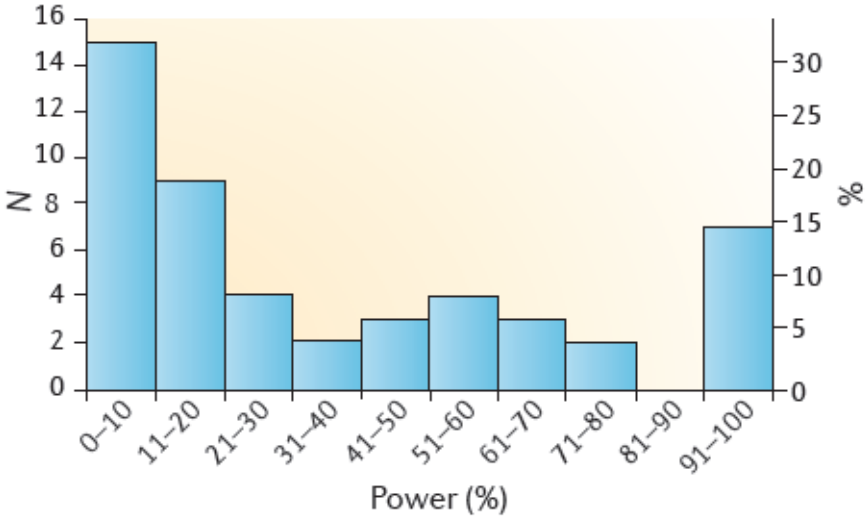
Nature Nanotechnology **8**, 693–695 (2013) | doi:10.1038/nnano.2013.204



Poor methods get results!

Low statistical power

Neuroscience



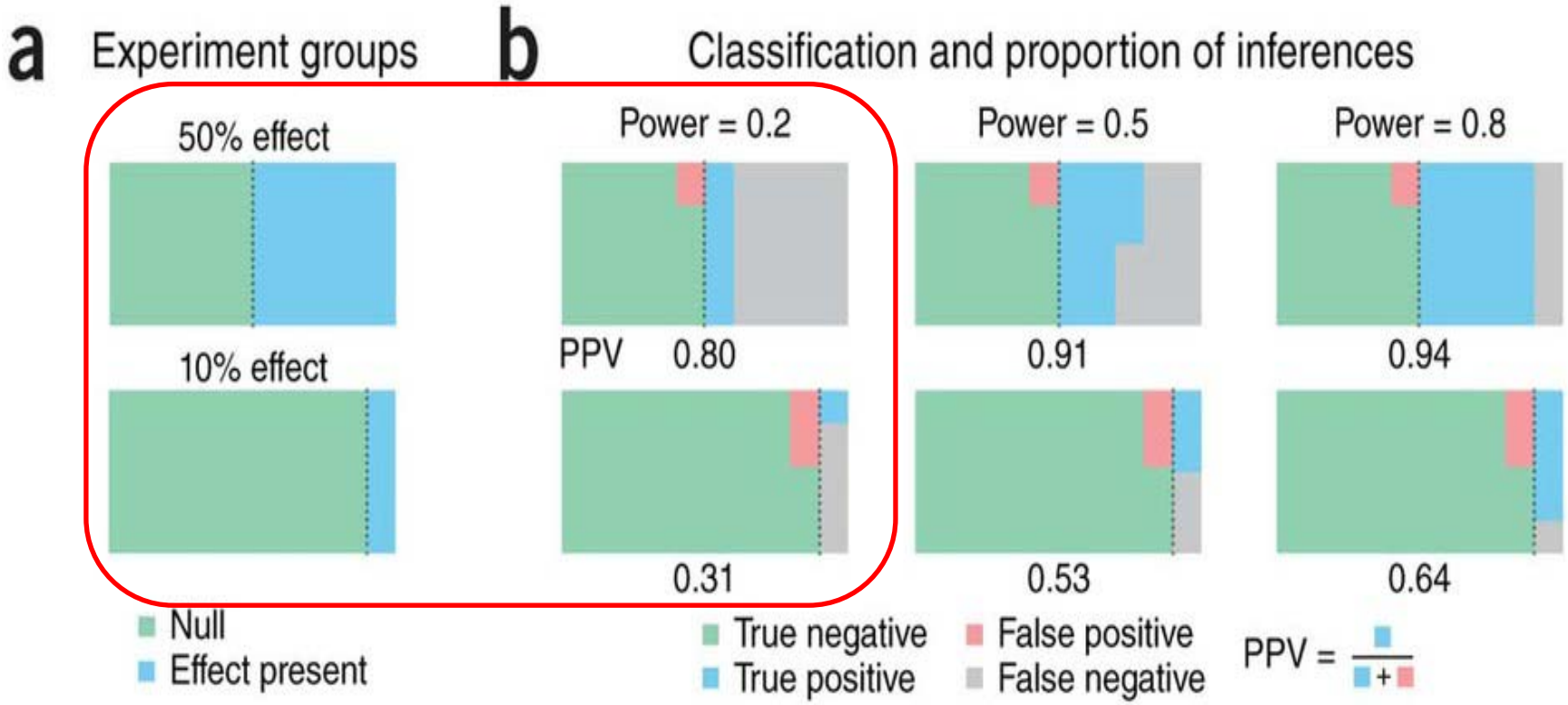
- *k* = 730 studies
- **Median power = 21%**
- Expected 254, observed 349 "significant" studies, $p < 0.0001$

Brain imaging

- Excess of significance in brain volume abnormalities (Ioannidis (2011). Archives of General Psychiatry, 14, 1105-1107).
- 41 meta-analyses pub. 2006 - 2009
- 461 contributing studies
- **Median power = 8%**

Button et al (2013). Nature Reviews Neuroscience, 14, 365 – 376

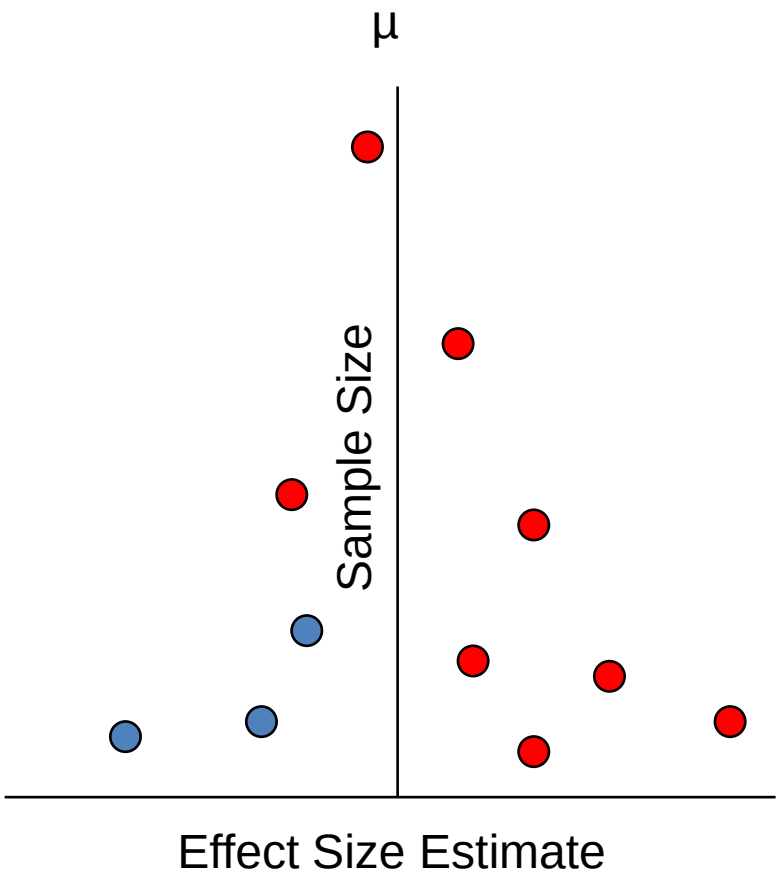
Low statistical power



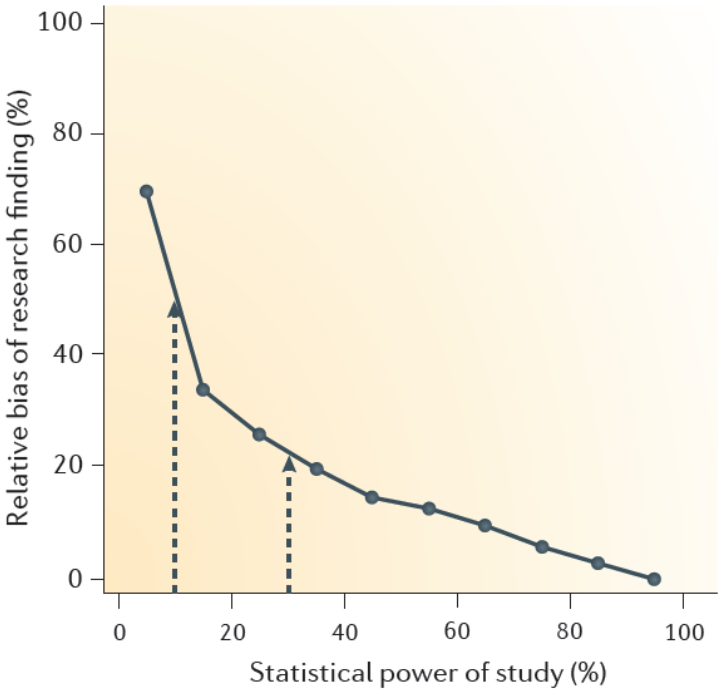
Lower positive predictive value

Krzywinski & Altman (2013). Nat Methods, 10, 1139-1140.

Low statistical power

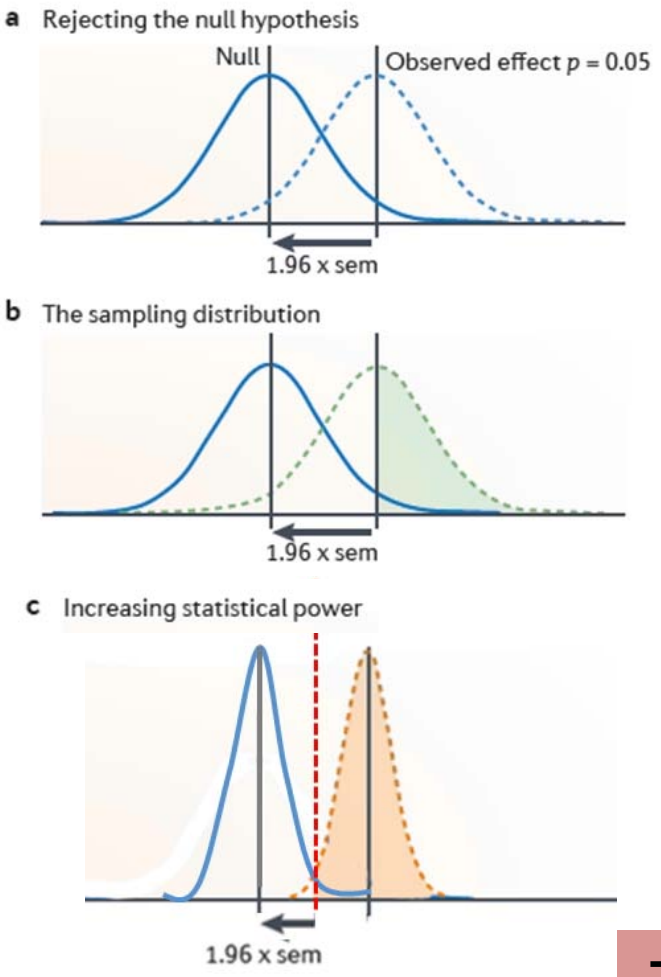


P < 0.05



Bigger (but biased) effects

Low statistical power



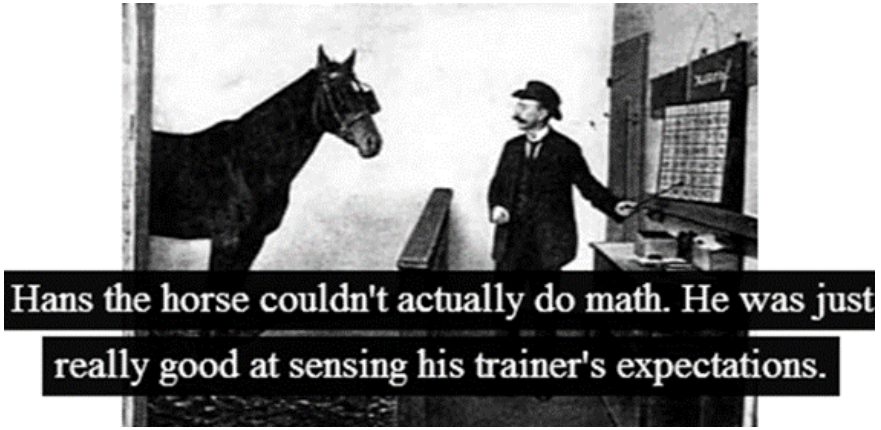
- To achieve adequate power, the replication study will need a sample size twice that of the original (where original p ~ 0.05)

		Discovery P-Value		
		0.05	0.01	0.005
Replication Sample	n	50%	74%	81%
	2n	80%	96%	98%
	4n	98%	100%	100%

This is assuming the original effect estimate is accurate

Lack of control for bias

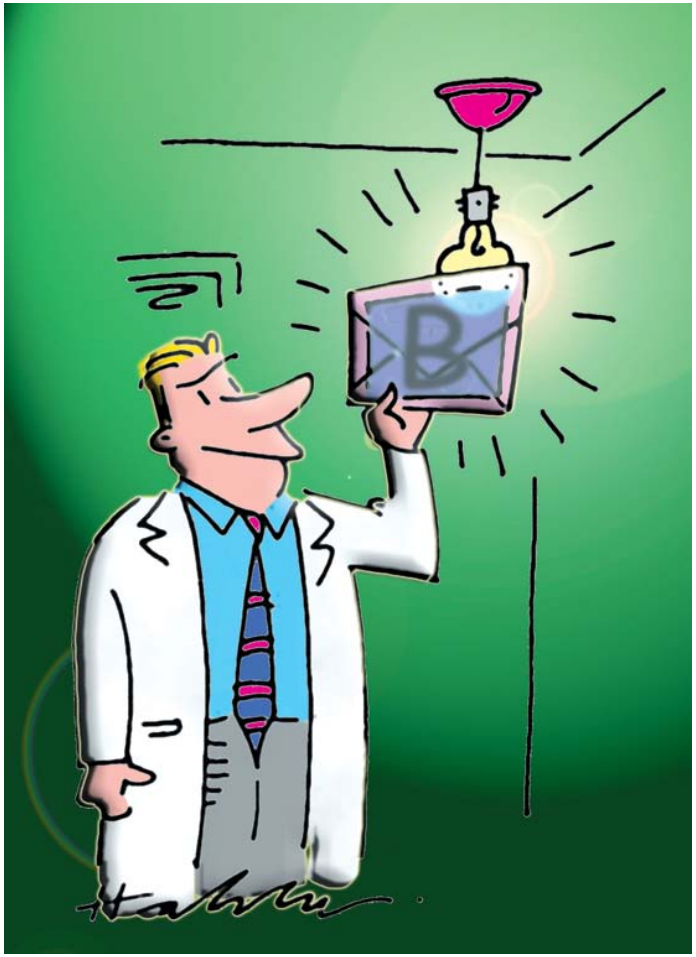
"Flaws in the design, conduct, analysis, and reporting of randomised trials can cause the effect of an intervention to be underestimated **or overestimated**."



Source	Control
Selection bias (confounding)	Randomisation + concealment
Performance bias	Blinding
Detection bias	Blinding
Attrition bias	ITT
Reporting bias	Pre-registration

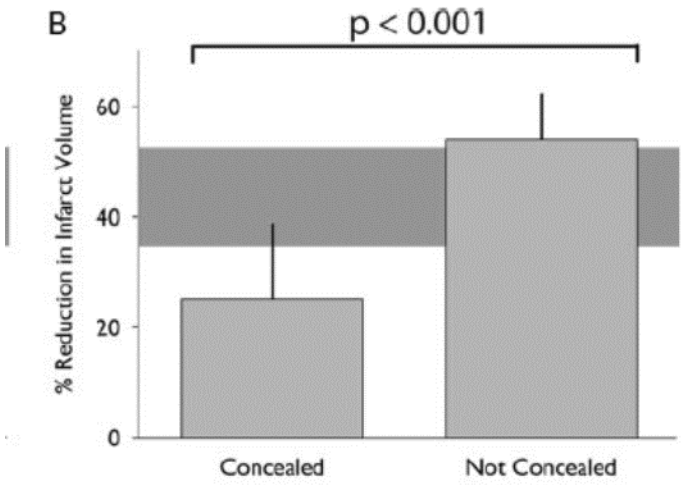
Higgins et al. BMJ 2011; 343 doi: http://dx.doi.org/10.1136/bmj.d5928

Lack of control for bias



Evidence for the Efficacy of NXY-059 in Experimental Focal Cerebral Ischaemia Is Confounded by Study Quality

Malcolm R. Macleod, PhD, FRCP; H. Bart van der Worp, MD, PhD; Emily S. Sena, BSc; David W. Howells, PhD; Ulrich Dirnagl, MD, PhD; Geoffrey A. Donnan, MD, FRACP



Schulz, Grimes (2002) THE LANCET, 359

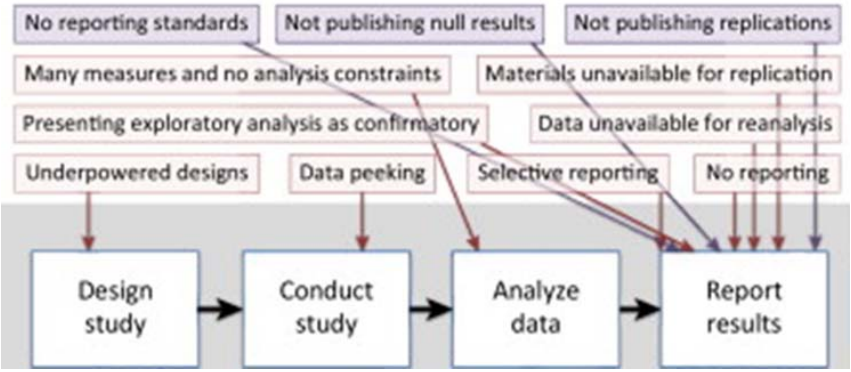
Questionable research practices

Table 1. Likelihood of Obtaining a False-Positive Result

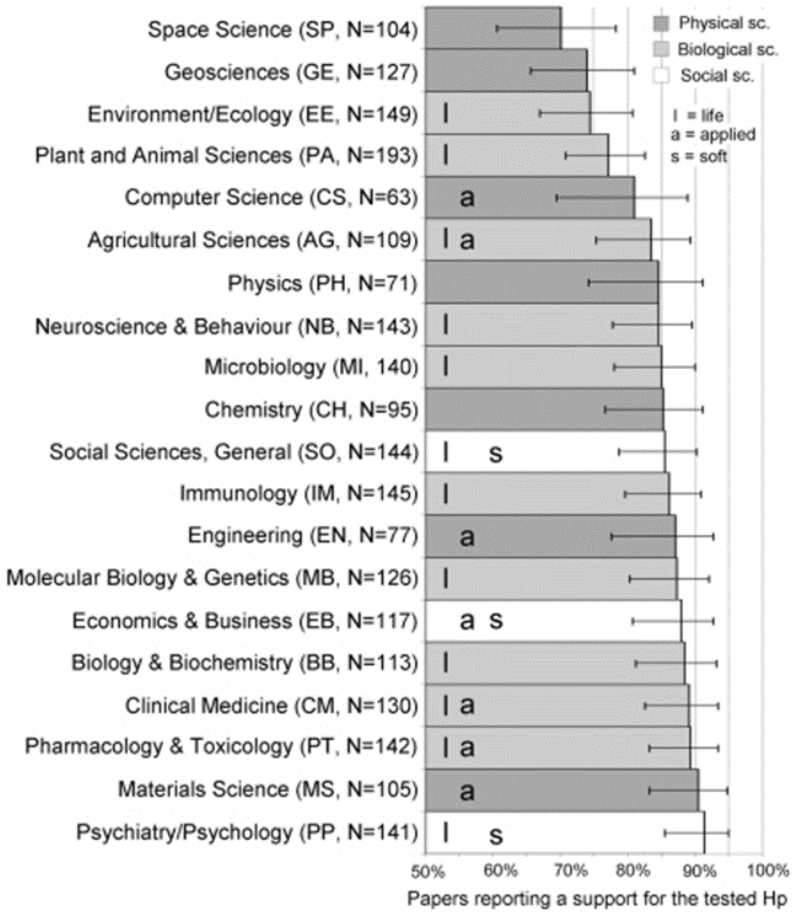
Researcher degrees of freedom	Significance level $p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

Note: The table reports the percentage of 15,000 simulated samples in which at least one of a set of analyses was significant. Observations were drawn independently from a normal distribu-

Simmons et al. (2011). Psychological Science, 22, 1359-1366.

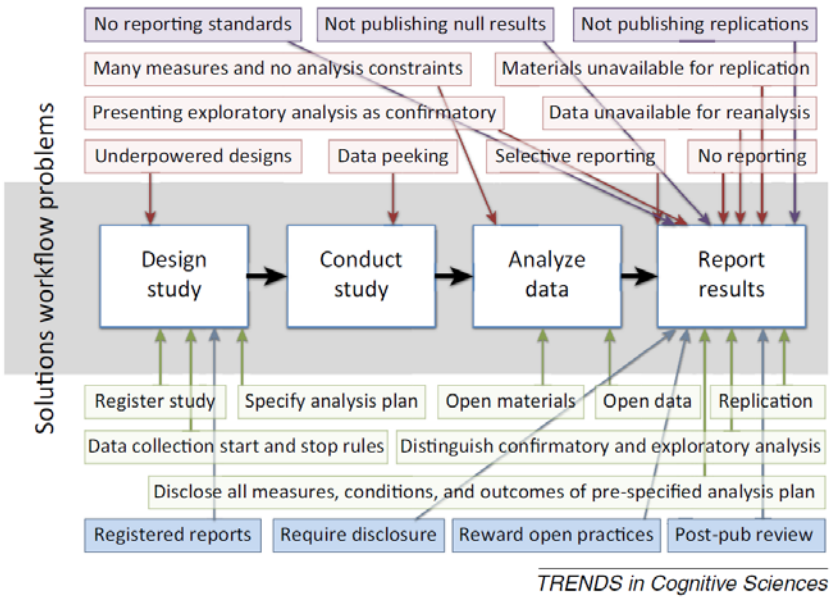


Ioannidis et al (2014) Trends in Cognitive Sciences, 18, 235 – 241.

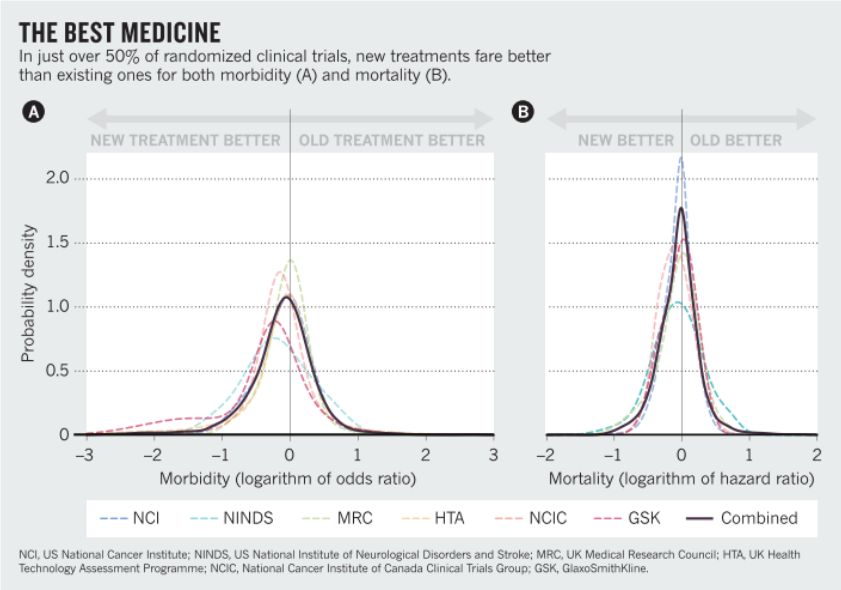


Fanelli (2012). PLOS ONE, 5, e10068.

Solutions



Ioannidis et al (2014) Trends in Cognitive Sciences, 18(5), 235 - 241



Djulbegovic et al (2013). Nature, 500, 395-396

But applied to pre-clinical = fewer publishable results?

Solutions

NIHR HTA Funded Trial (3 – 4 years)

Design study
- Grant peer reviewed
- Panel reviewed
- Protocol registered and published

→

Conduct study
- Steering committee
- Stopping rules

→

Analyse data
- Primary outcome per-protocol
- Trial statistician

→

Report results
- Contractual
- HTA report open access

12 – 18 months

Student human laboratory study (4 – 6 months)

Design study
- No peer review
- Protocol may be published (OSF)
- Low power

→

Conduct study

→

Analyse data
- Primary outcome per-protocol

→

Report results
- Not-contractual
- Try and publish in peer-review journal otherwise results on OSF

1 month!

Concluding remarks

- Poor methods and QRPs means most published findings are likely false
- There are solutions but robust, reproducible research takes more time and resources
- We need systemic change

  to prevent good-practice

  being penalised

University of BRISTOL

Acknowledgements

Collaborators:

Marcus Munafò	University of Bristol
John Ioannidis	Stanford University
Claire Mokrysz	UCL
Brian Nosek	University of Virginia
Jonathan Flint	University of Oxford
Emma Robinson	University of Bristol

Thanks to:

Tobacco and Alcohol Research Group

Kate.Button@bristol.ac.uk

@ButtonKate

@BristolTARG

Katherine button is funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR). The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health. The NIHR SPCR is a partnership between the Universities of Birmingham, Bristol, Keele, Manchester, Nottingham, Oxford, Southampton and University College London.

School for Primary Care Research

The National Institute for Health Research School for Primary Care Research (NIHR SPCR) is a partnership between the Universities of Birmingham, Bristol, Keele, Manchester, Nottingham, Oxford, Southampton and University College London.

NHS National Institute for Health Research