Warum scheitern soviele im Labor hoffnungsvolle Therapien in der klinischen Prüfung?

Center for Stroke Research Berlin

Berin, 24.9.2015

Outline

What is the problem?

Why do we have a problem?

How can we solve the problem?

Outline

What is the problem?

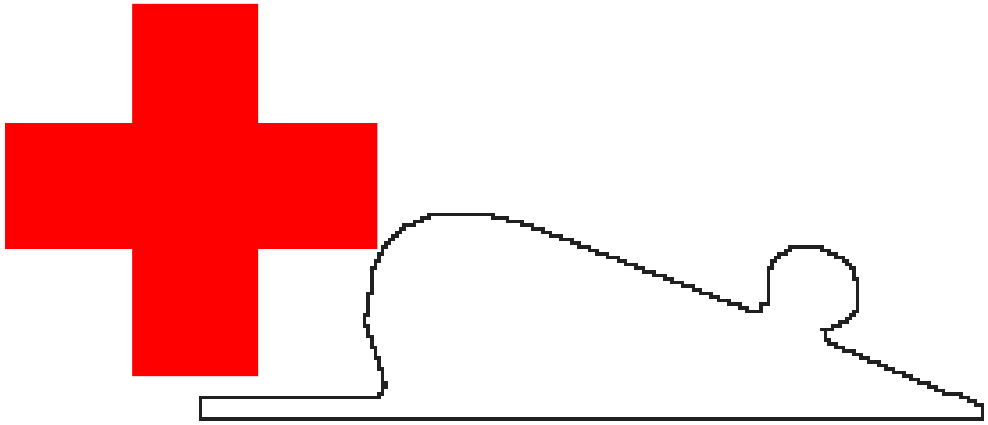
Why do we have a problem?

How can we solve the problem?

Center for Stroke Research Berlin

„The outlook for stroke therapy is excellent ... if you‘re a rat.“

Lindsay Symon, Neurosurgeon

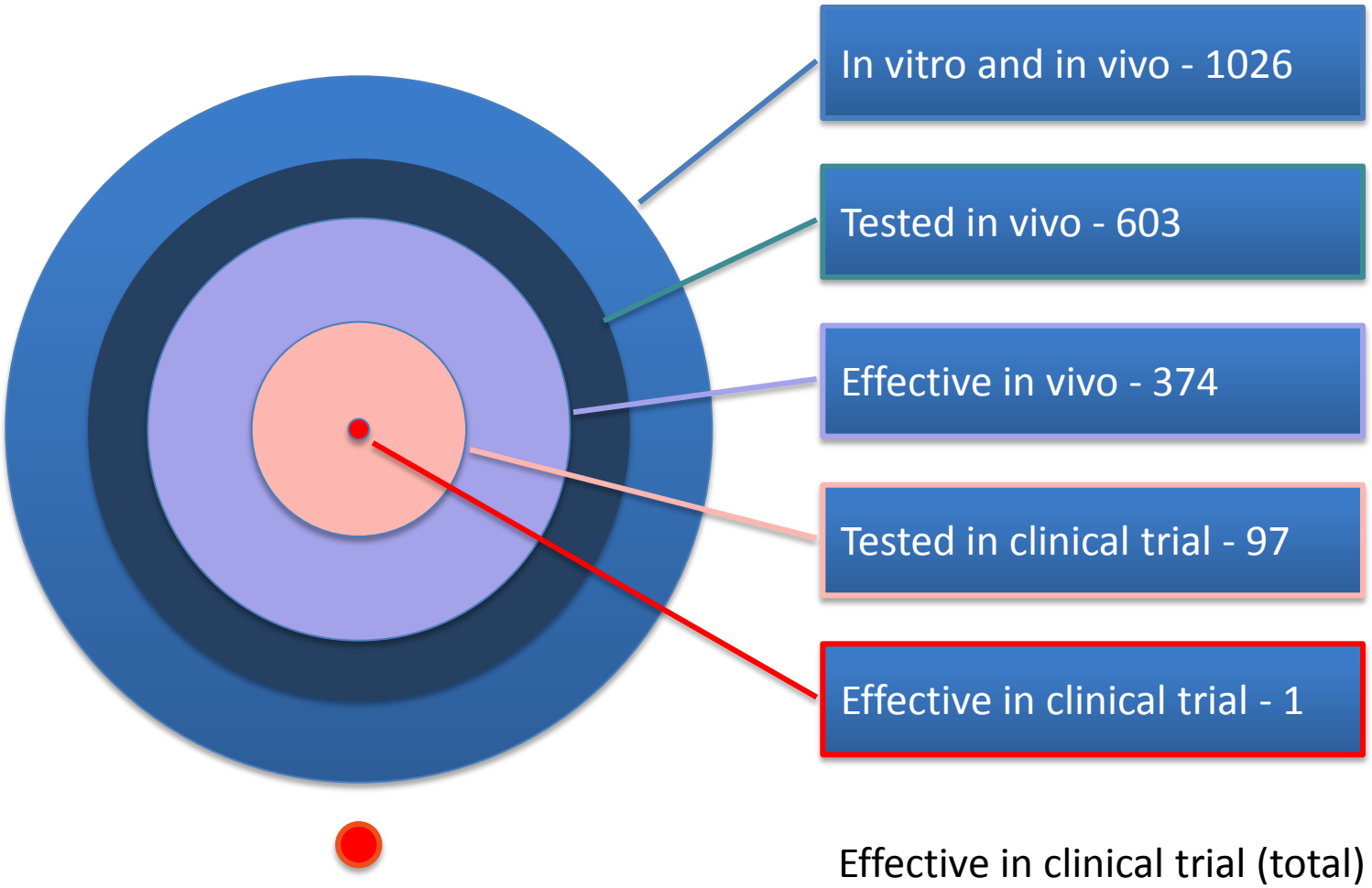


A typical intervention in exp. stroke studies reduces infarct sizes by 30-50 %.

Neuroregenerative strategies (eg. ‘stem cells’) improve functional outcome even after infarct maturation.

Center for Stroke Research Berlin

1026 interventions in experimental stroke

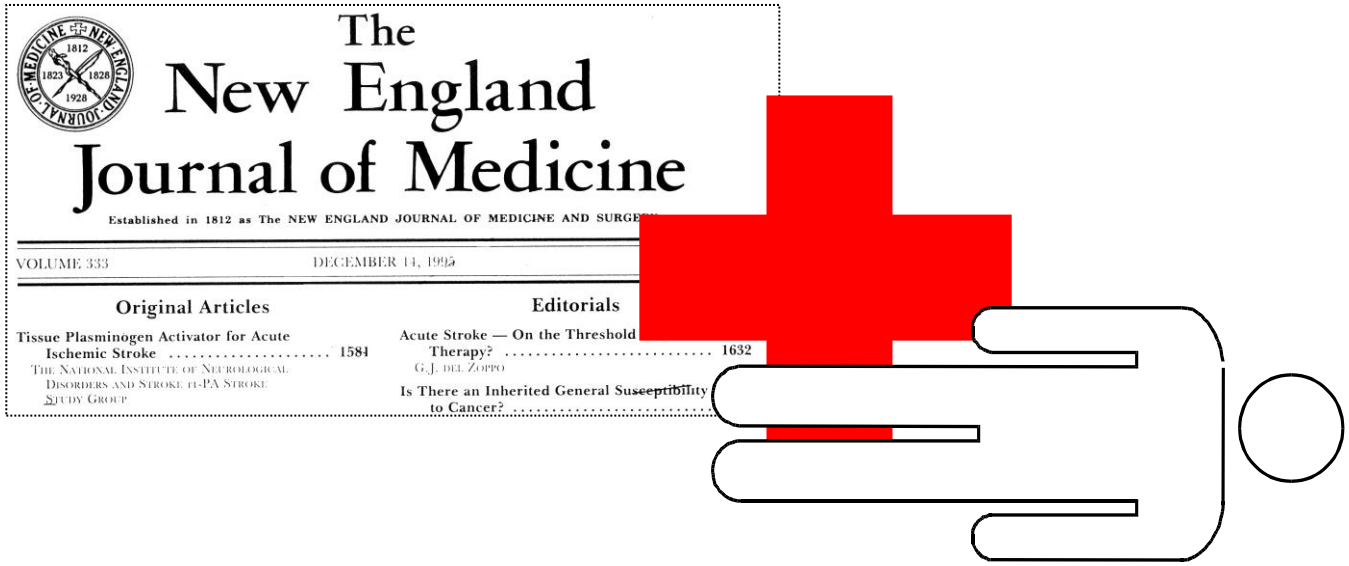


Effective in clinical trial (total) - 3

O'Collins et al, 2006

Center for Stroke Research Berlin

Only thrombolysis clinically effective!



I.v. thrombolysis is the only clinically proven pharmacological therapy of acute ischemic stroke.

Benefit only for a small percentage of stroke victims.

There is no therapeutic 'neuroprotection' or 'neuroregeneration' in human stroke.

Costly failures

The NEW ENGLAND JOURNAL of MEDICINE

N Engl J Med 2007;357:562-71.

ORIGINAL ARTICLE

NXY-059 for the Treatment of Acute Ischemic Stroke

Public release date: 17-Oct-2011
[Print | E-mail | Share] [Close Window]

Back To EurekAlert!

Contact: David Cameron
david_cameron@hms.harvard.edu
617-432-0441
Harvard Medical School

Harvard Medical School launches major initiative to address crisis in drug development

ELSEVIER

www.elsevier.com/locate/ynbdi
Neurobiology of Disease 26 (2007) 1–13

Review

Lost in translation: Treatment trials and in human ALS

Michael Benatar*

Department of Neurology, Emory University School of Medicine, Woodruff Memor

Received 20 October 2006; revised 12 December 2006; accepted 20 December 20
Available online 3 January 2007

Therapeutic success in the superoxide dismutase (SOD1) mouse model of amyotrophic lateral sclerosis (ALS) has not translated into effective therapy for human ALS, calling into question the utility of such

with an anti-inflammatory mechanism of action, and anti-oxidative agents such as creatine or the manganese porphyrin AEOL-10150, appear to be the most promising for preventative and therapeutic trials respectively in patients with familial ALS. These conclusions should be tempered by the methodological limitations of the relevant literature.

OPEN ACCESS Freely available online

PLoS CLINICAL TRIALS

Essay

Evolution and Translation of Research Findings: From Bench to Where?

John P. A. Ioannidis

JAMA®
The Journal of the American Medical Association

Commentary

Lost in Translation

Bumps in the Road Between Bench and Bedside

Cummings *et al. Alzheimer's Research & Therapy* 2014, 6:37
http://alzres.com/content/6/4/37

alzheimer's research&therapy

RESEARCH Open Access

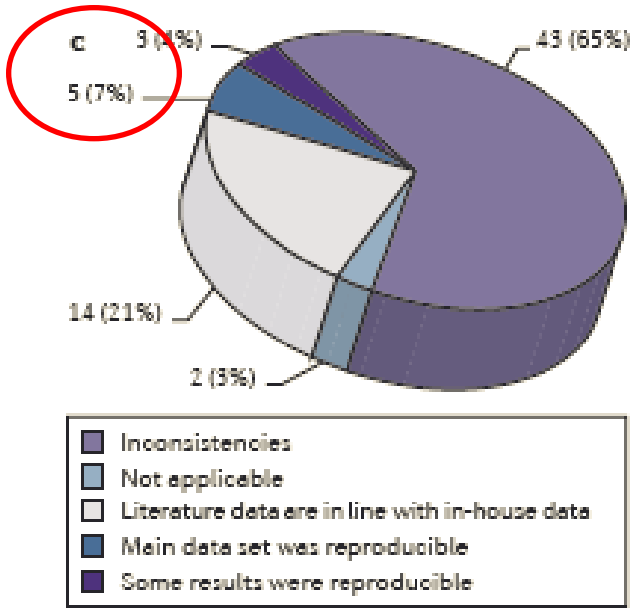
Alzheimer's disease drug-development pipeline: few candidates, frequent failures

Jeffrey L Cummings[1*], Travis Morstorf[2] and Kate Zhong[1]

Results: During the 2002 to 2012 observation period, 413 AD trials were performed: 124 Phase 1 trials, 206 Phase 2 trials, and 83 Phase 3 trials. Seventy-eight percent were sponsored by pharmaceutical companies. The United States of America (U.S.) remains the single world region with the greatest number of trials; cumulatively, more non-U.S. than U.S. trials are performed. The largest number of registered trials addressed symptomatic agents aimed at improving cognition (36.6%), followed by trials of disease-modifying small molecules (35.1%) and trials of disease-modifying immunotherapies (18%). The mean length of trials increases from Phase 2 to Phase 3, and the number of participants in trials increases between Phase 2 and Phase 3. Trials of disease-modifying agents are larger and longer than those for symptomatic agents. A very high attrition rate was found, with an overall success rate during the 2002 to 2012 period of 0.4% (99.6% failure).

Believe it or not: how much can we rely on published data on potential drug targets?

Florian Prinz, Thomas Schlange and Khusru Asadullah



'Indeed, our analysis revealed that the reproducibility of published data did not significantly correlate with journal impact factors, the number of publications on the respective target or the number of independent groups that authored the publications. '

Raise standards for preclinical cancer research

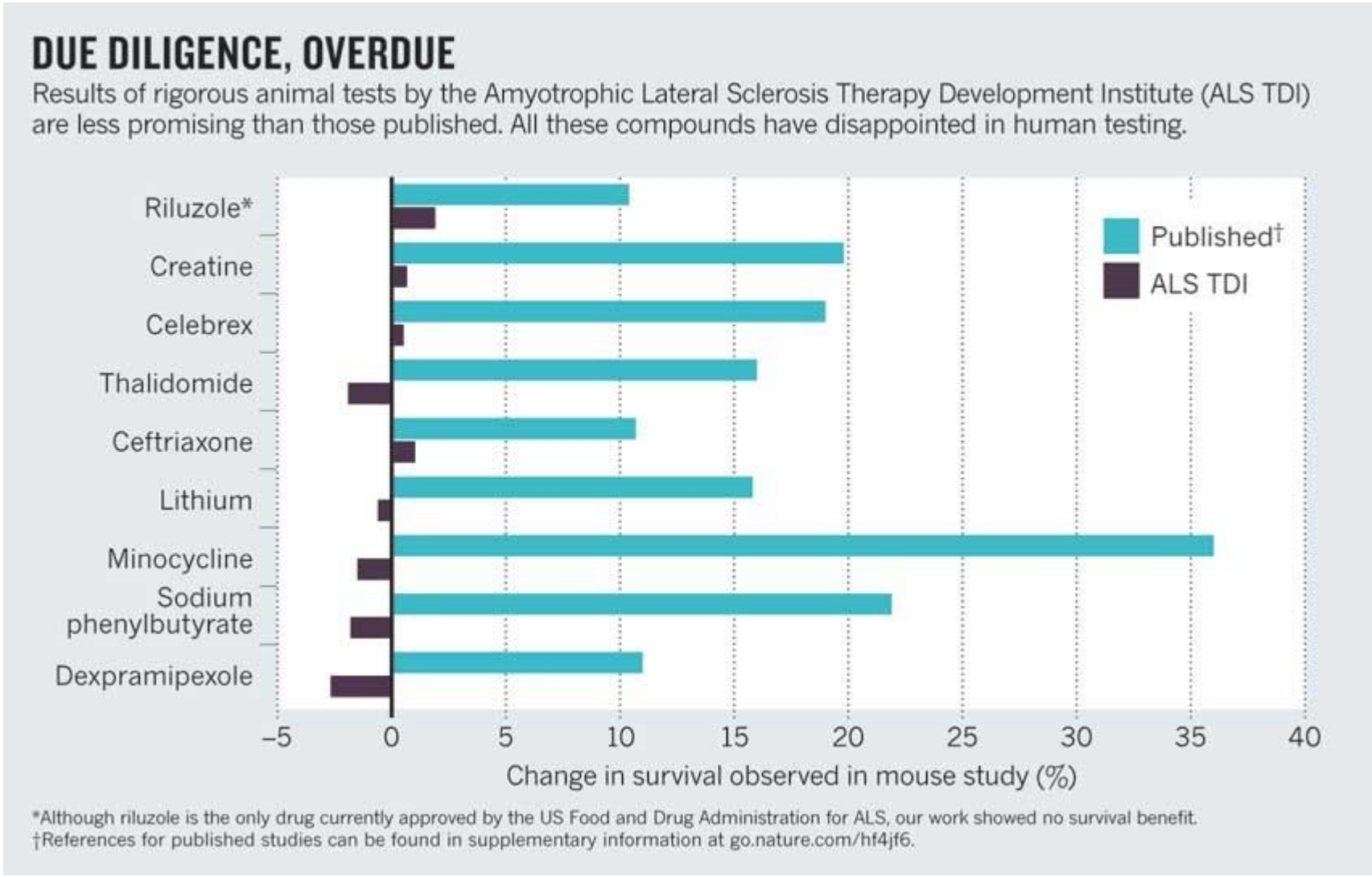
C. **Glenn Begley** and **Lee M. Ellis** propose how methods, publications and incentives must change if patients are to benefit.

related to that work. Fifty-three papers were deemed 'landmark' studies (see 'Reproducibility of research findings'). It was acknowledged from the outset that some of the data might not hold up, because papers were deliberately selected that described something completely new, such as fresh approaches to targeting cancers or alternative clinical uses for existing therapeutics. Nevertheless, scientific findings were confirmed in only 6 (11%) cases. Even knowing the limitations of preclinical research, this was a shocking result.

REPRODUCIBILITY OF RESEARCH FINDINGS
Preclinical research generates many secondary publications, even when results cannot be reproduced.

Journal impact factor	Number of articles	Mean number of citations of non-reproduced articles*	Mean number of citations of reproduced articles
>20	21	248 (range 3–800)	231 (range 82–519)
5–19	32	169 (range 6–1,909)	13 (range 3–24)

'Replication crisis' (= 'Nonreplication epidemic')



Perrin S (2014) Nature 407:423-425

'Replication crisis' (= 'Nonreplication epidemic')

Special Issue: NIH Replication Studies

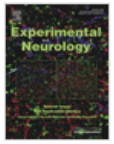
Edited By Oswald Steward and Phillip Popovich

ELSEVIER

Experimental Neurology

Volume 233, Issue 2, February 2012, Pages 597–605

Special Issue: NIH Replication Studies



Editorial

Replication and reproducibility in spinal cord injury research

Oswald Steward[a, b, c, d], Phillip G. Popovich[e, f], W. Dalton Dietrich[g, h], Naomi Kleitman[i]

Show more

doi:10.1016/j.expneurol.2011.06.017

Get rights and content

Abstract

This special issue of Experimental Neurology compiles a series of papers that either explicitly replicate published studies or retest phenomena reported in previous publications. The explicit replications were carried out as part of the "Facilities of Research Excellence—Spinal Cord Injury" (FORE—SCI) program launched by the National Institute of Neurological Disorders and Stroke (NINDS) in 2003. Here, we review the FORE—SCI replication experiments published prior to those in this special issue. We then discuss emerging issues regarding replication and reproducibility in spinal cord injury research, especially in terms of potential translation to clinical trials.

Keywords

Replication; Regeneration

Costly nonreproducibility

PLOS | BIOLOGY

PERSPECTIVE

The Economics of Reproducibility in Preclinical Research

Leonard P. Freedman[1]*, Iain M. Cockburn[2], Timothy S. Simcoe[2,3]

1 Global Biological Standards Institute, Washington, D.C., United States of America, 2 Boston University School of Management, Boston, Massachusetts, United States of America, 3 Council of Economic Advisers, Washington, D.C., United States of America

* lfreedman@gbsi.org

OPEN ACCESS

Citation: Freedman LP, Cockburn IM, Simcoe TS (2015) The Economics of Reproducibility in

Abstract

Low reproducibility rates within life science research undermine cumulative knowledge production and contribute to both delays and costs of therapeutic drug development. An analysis of past studies indicates that the cumulative (total) prevalence of irreproducible preclinical research exceeds 50%, resulting in approximately US$28,000,000,000 (US $28B)/year spent on preclinical research that is not reproducible—in the United States alone. We outline a framework for solutions and a plan for long-term improvements in reproducibility rates that will help to accelerate the discovery of life-saving therapies and cures.

Summary I: What is the problem

- Great progress in the lab (and academic careers...),
- but little of this gets translated into efftive new therapies.
- 'Replication crisis'
- Waste of resources, potential harm to patients.

Outline

What is the problem?

Why do we have a problem?

How can we solve the problem?

A quality problem? Evidence from meta-research (research on research)

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 relations field. In th

OPEN ACCESS Freely available online

PLOS | BIOLOGY

Evaluation of Excess Significance Bias in Animal Studies of Neurological Diseases

Konstantinos K. Tsilidis[1*], Orestis A. Panagiotou[1*], Emily S. Sena[2,3], Eleni Aretouli[4,5], Evangelos Evangelou[1], David W. Howells[3], Rustam Al-Shahi Salman[2], Malcolm R. Macleod[2], John P. A. Ioannidis[6*]

1 Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, 2 Department of Clinical Neurosciences, University of Edinburgh,

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button[1,2], John P. A. Ioan
Jonathan Flint[5], Emma S. J. Robinson[6]

Nature Reviews Neuroscience | AOP, publishe

Empirical Evidence of Bias in the Design of Experimental Stroke Studies

A Metaepidemiologic Approach

Nicolas A. Crossley, MSc; Emily Sena, BSc; Jos Goehler; Jannekke Horn, MD; Bart van der Worp, MD; Philip M.W. Bath, MD; Malcolm Macleod, PhD; Ulrich Dirnagl, MD

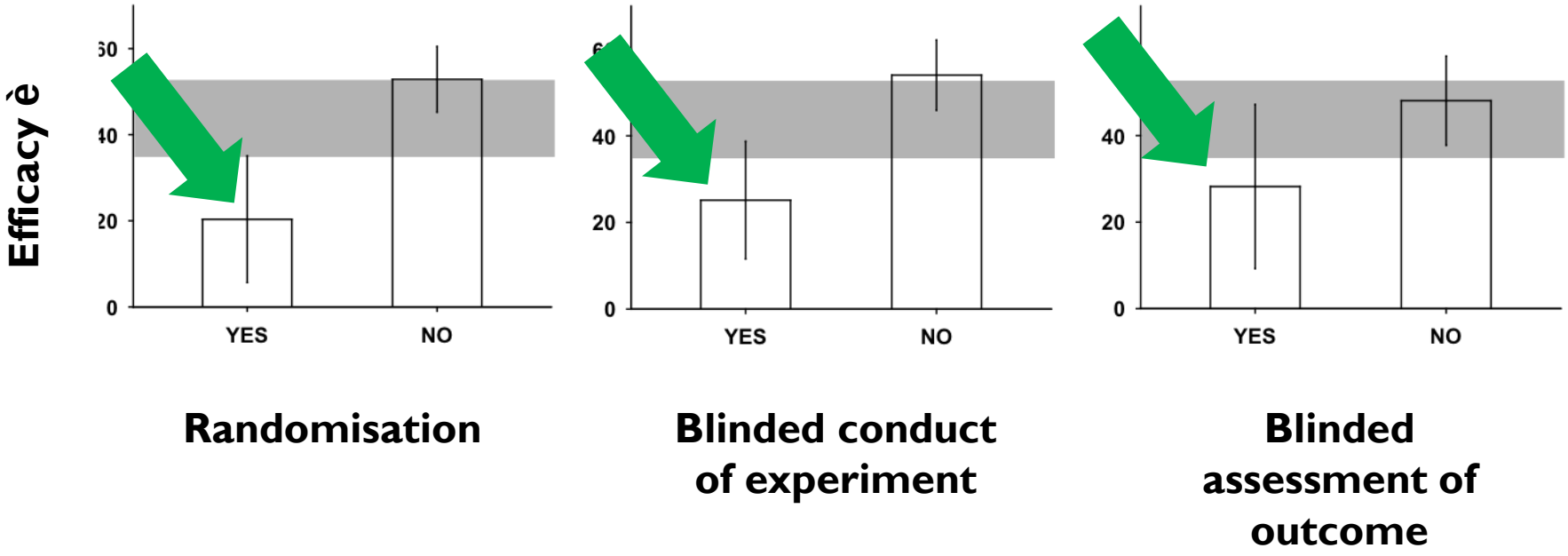
Background and Purpose—At least part of the failure in the transition from experimental to clinical studies in stroke has been attributed to the imprecision introduced by problems in the design of experimental stroke studies. Using a metaepidemiologic approach, we addressed the effect of randomization, blinding, and use of comorbid animals on the estimate of how effectively therapeutic interventions reduce infarct size.

Methods—Electronic and manual searches were performed to identify meta-analyses that described interventions in experimental stroke. For each meta-analysis thus identified, a reanalysis was conducted to estimate the impact of various quality items on the estimate of efficacy, and these estimates were combined in a meta–meta-analysis to obtain a summary measure of the impact of the various design characteristics.

Results—Thirteen meta-analyses that described outcomes in 15 635 animals were included. Studies that included

Effect size inversely correlates with study quality

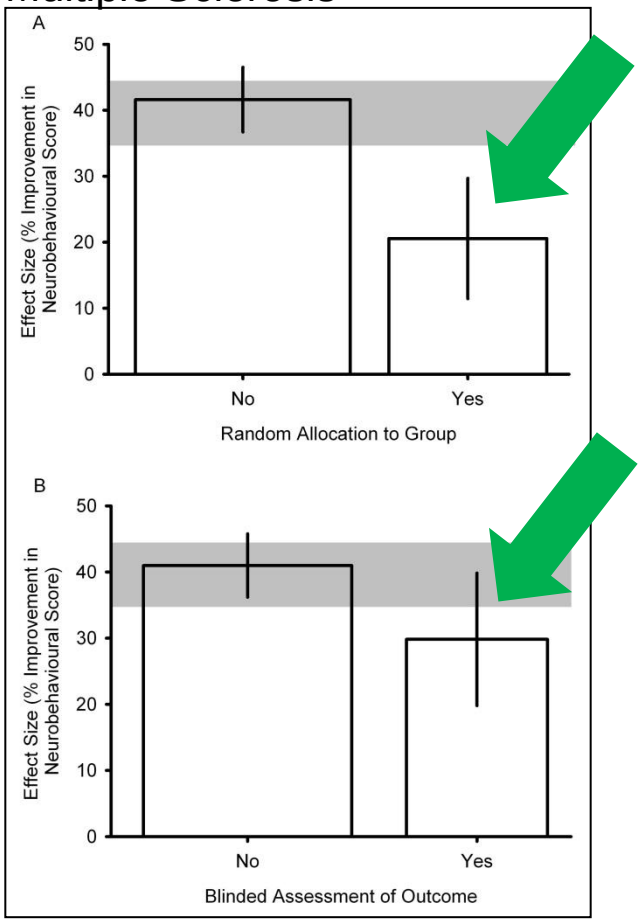
- Treatment with NXY-059. Outcome: Infarct Volume
  - 11 publications, 29 experiments, 408 animals
  - Improved outcome by 44% (35-53%)



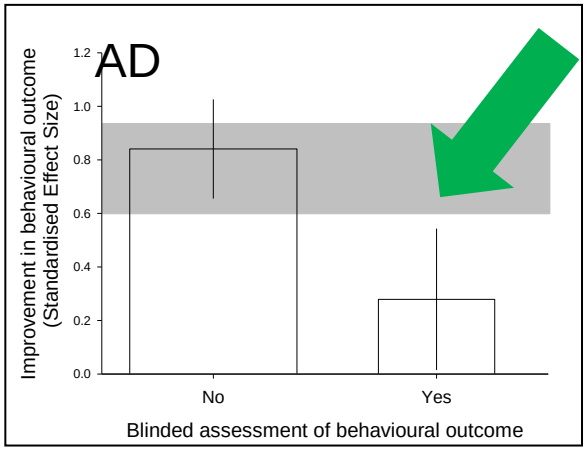
(Stroke 2008; 39:2824-9.)

Low internal validity: Meta-research exposes exposes selection and perfomance bias

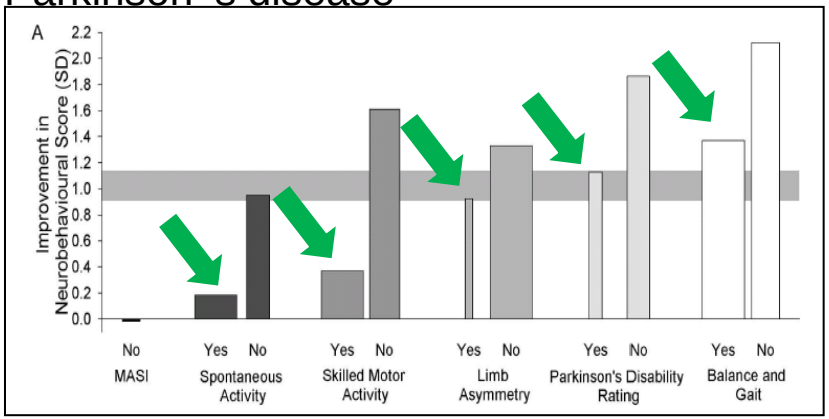
Multiple Sclerosis



Alzheimer´s disease



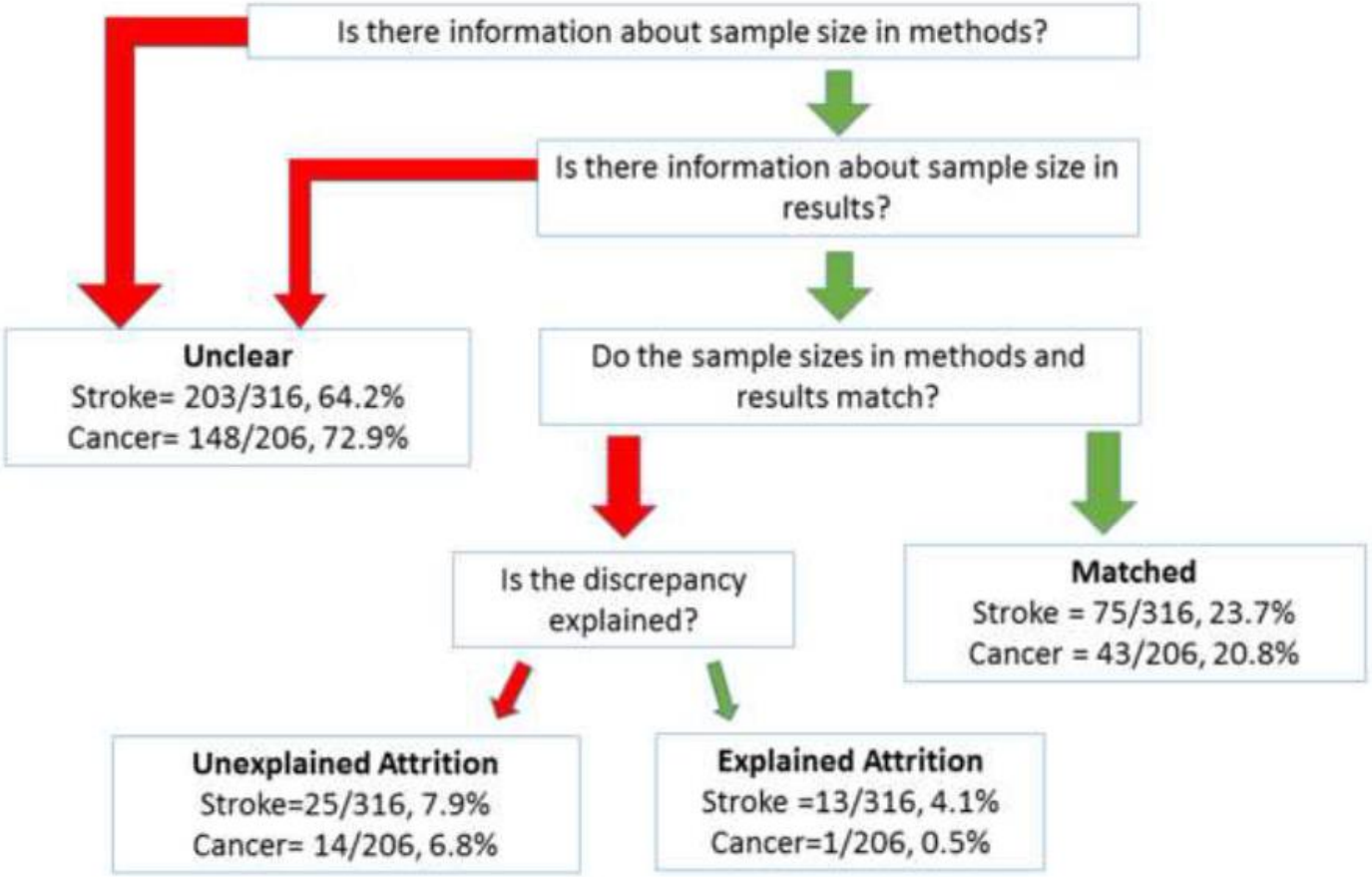
Parkinson´s disease



Low internal validity: Meta-research exposes exposes selection and perfomance bias

Disease modelled		Number of Publications	Sample Size Calculation (%)	Random Allocation to Group (%)	Blinded conduct of experiment (%)	Blinded Assessment of Outcome (%)
Alzheimer's Disease		428	0	16	n/a	22
Multiple Sclerosis		1117	<1	9	n/a	16
Parkinson's Disease		252	<1	16	n/a	15
Intracerebral Haemorrhage		88	0	31	8	49
Pain		160	0	12	n/a	26
Focal Ischaemia	NXY 059	9	22	33	56	44
	Hypothermia	101	0	36	4	38
	Erythropoietin	19	0	37	21	42
	Tirilazad	18	0	67	6	72
	Alteplase	113	7	37	20	21

Attrition in preclinical research (stroke, cancer)

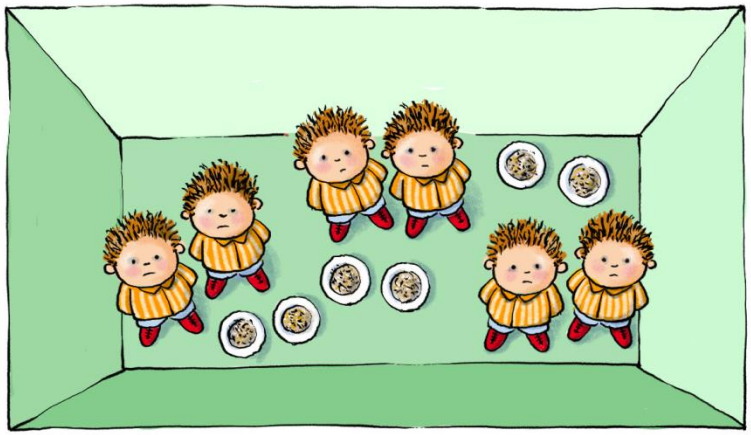


Holman et al. (submitted)

Internal validity is low

- Selection bias (creating groups with different confounders; solved by randomization)
- Performance bias and detection bias (investigators respectively treating or assessing more positively those subjects on the treatment arm; controlled by blinding interventions and outcome assessments);
- Attrition bias (dropouts of subjects with a negative outcome not included in the final result)

External validity is low



Healthy, pubertal male twins raised in 6 m^2 isolator tents on an enriched granola diet

vs.

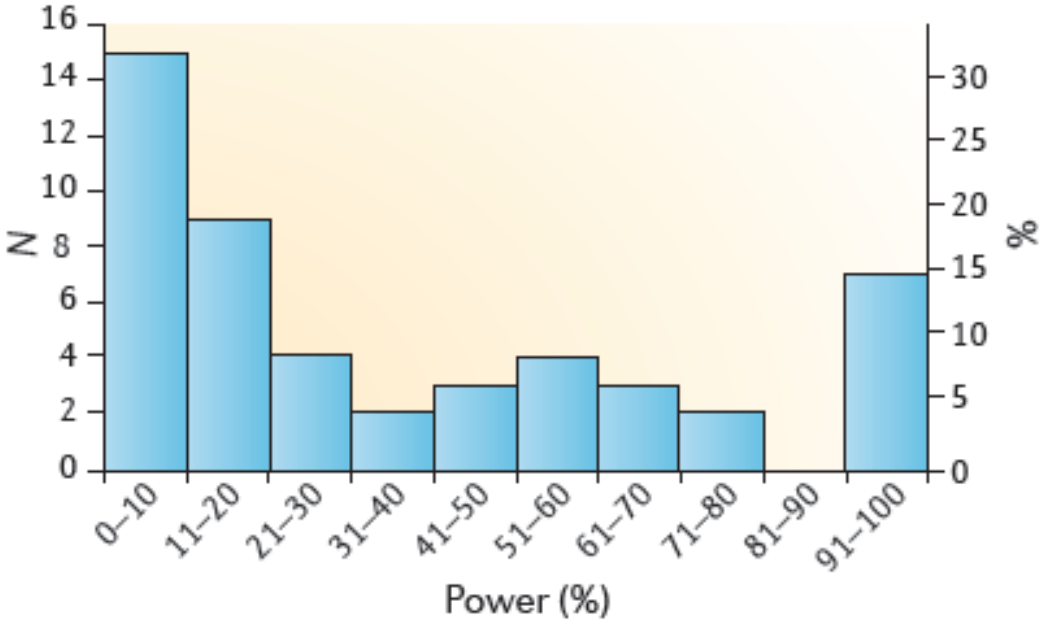
Patients of both sexes, elderly, comorbid, multiple medications, exposed to multiple pathogens and antigens throughout life

Statistical power is low, false positives abound

Power failure: why small sample size undermines the reliability of neuroscience

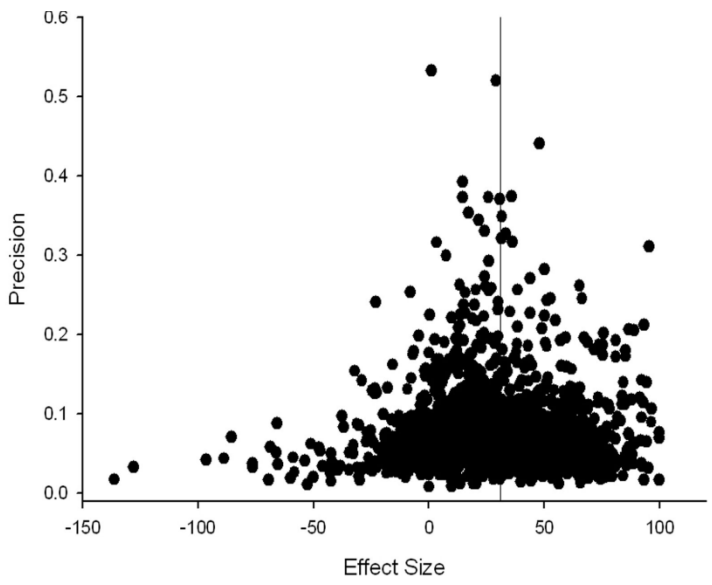
Katherine S. Button[1,2], John P. A. Ioannidis[3], Claire Mokrysz[1], Brian A. Nosek[4], Jonathan Flint[5], Emma S. J. Robinson[6] and Marcus R. Munafò[1]

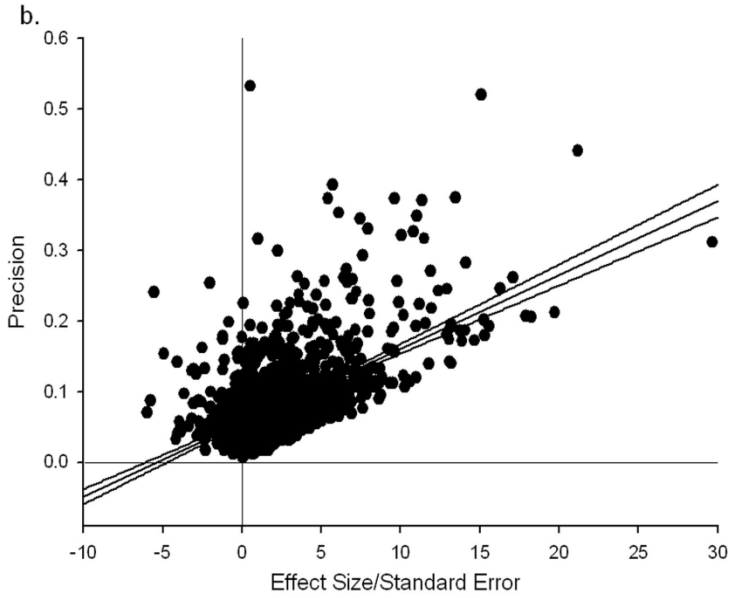
Nature Reviews Neuroscience | AOP, published online 10 April 2013; doi:10.1038/nrn3475



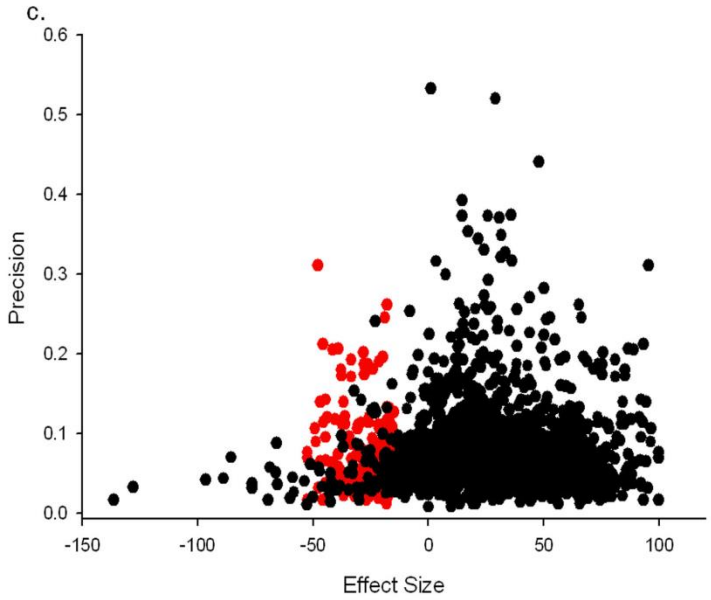
Overall median power of 730 primary neuroscience studies: 21 %

Only "positive" results are published



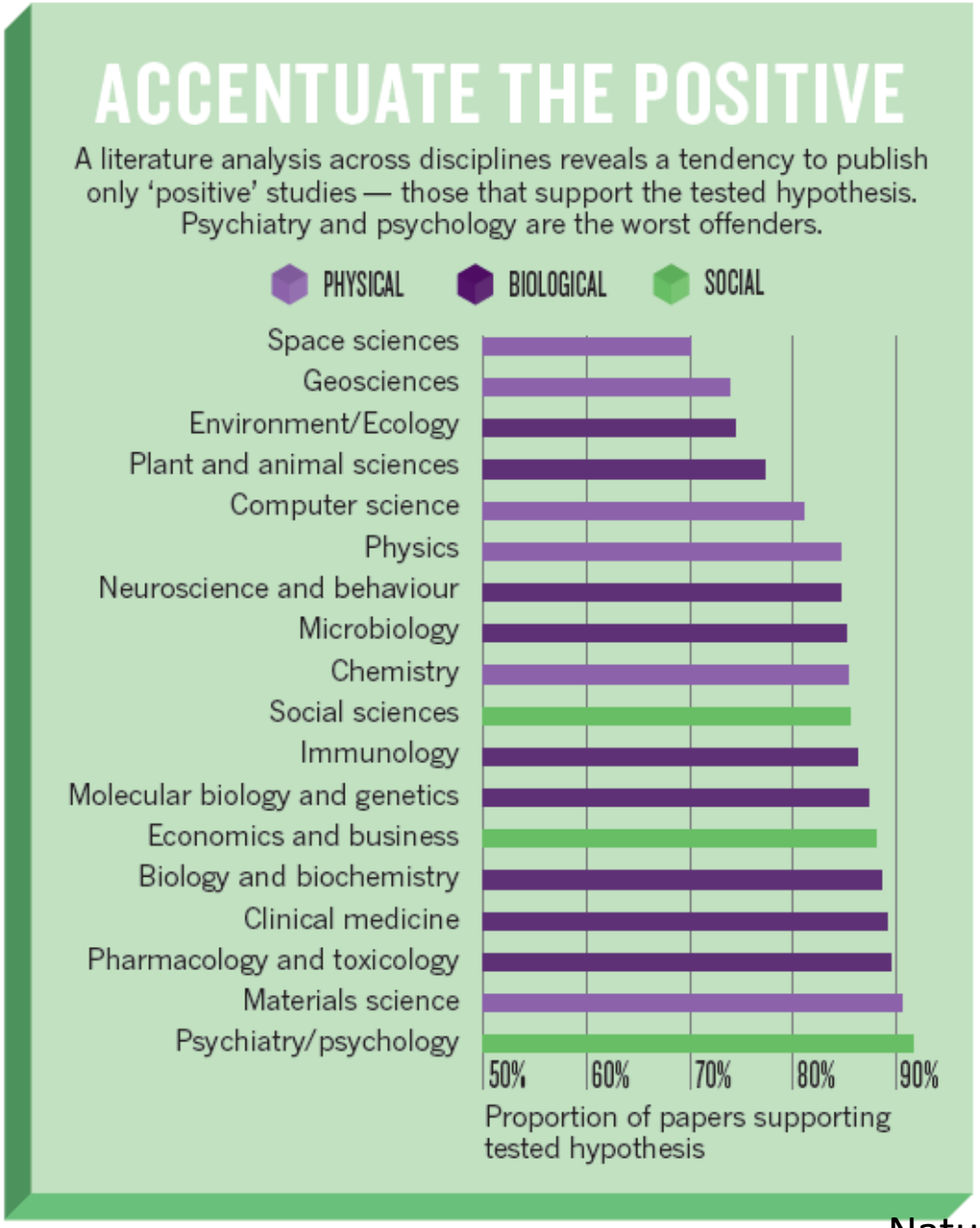


"Publication bias is highly prevalent (present in the literature describing the efficacy of at least 16 of 18 interventions) and accounts for around 30% of the reported efficacy of candidate neuroprotective interventions."



Publication bias in reports of animal stroke studies leads to major overstatement of efficacy
Emily S Sena, H. Bart van der Worp, Philip M.W. Bath, David W Howells and Malcolm R Macleod (PLoS Biol. 2010 Mar 30;8(3):e1000344)

Only "positive" results are published



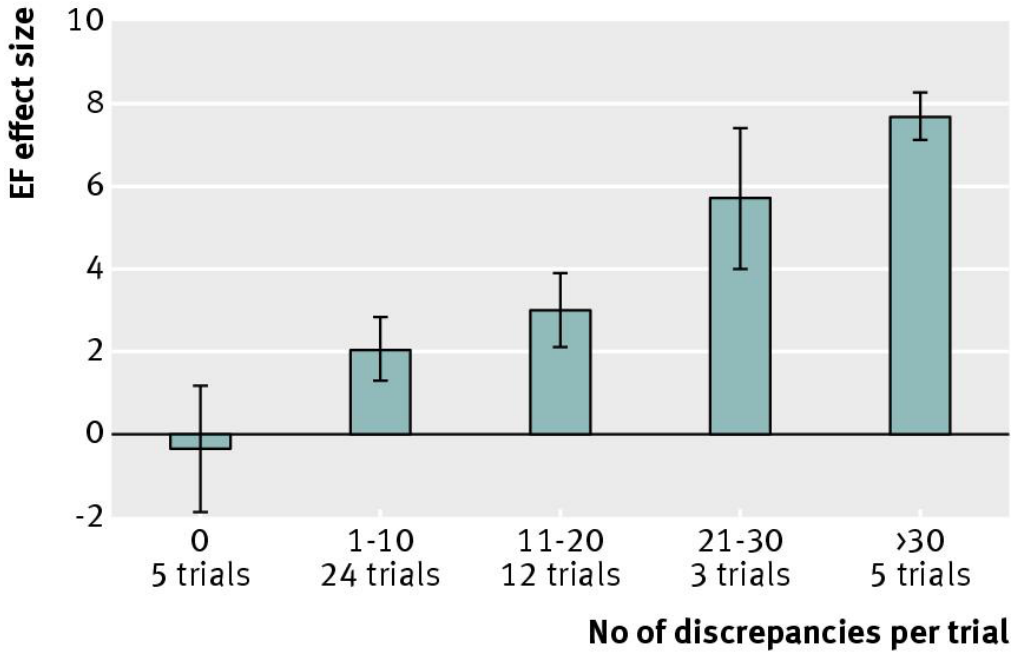
Nature (2012) 485:298-300

BMJ

BMJ 2014;348:g2688 doi: 10.1136/bmj.g2688 (Published 29 April 2014)

RESEARCH

Discrepancies in autologous bone marrow stem cell trials and enhancement of ejection fraction (DAMASCENE): weighted regression and meta-analysis



Quality problems in cinical research

RESEARCH

BMJ. 2015 Sep 16;351:h4320.

OPEN ACCESS

Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence

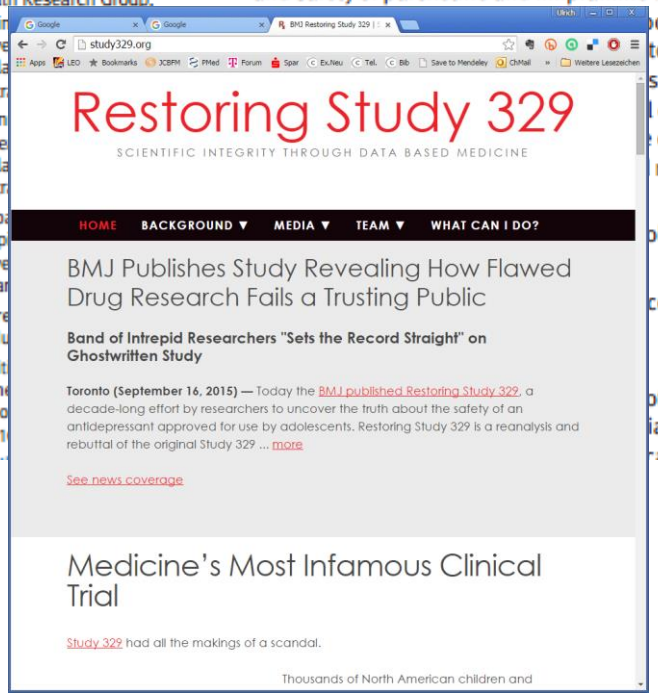
Joanna Le Noury,[1] John M Nardo,[2] David Healy,[1] Jon Jureidini,[3] Melissa Raven,[3] Catalin Tufanaru,[4] Elia Abi-Jaoude[5]

[1]School of Medical Sciences, Bangor University, Bangor, Wales, UK
[2]Emory University, Atlanta, Georgia, USA
[3]Critical and Ethical Mental Health Research Group,

ABSTRACT

OBJECTIVES

To reanalyse SmithKline Beecham's Study 329 (published by Keller and colleagues in 2001), the primary objective of which was to compare the effic and safety of paroxetine and imipramine with place



Summary II: Why do we have a problem?

- Low internal validity (bias due to lack of randomization, blinding, attrition etc.)
- Low external validity (gender, age, comorbidities)
- Low statistical power (exceedingly small group sizes)
- Positive publication bias

consequently

- False positives
- Inflated effect sizes
- Non-replicability
- Waste

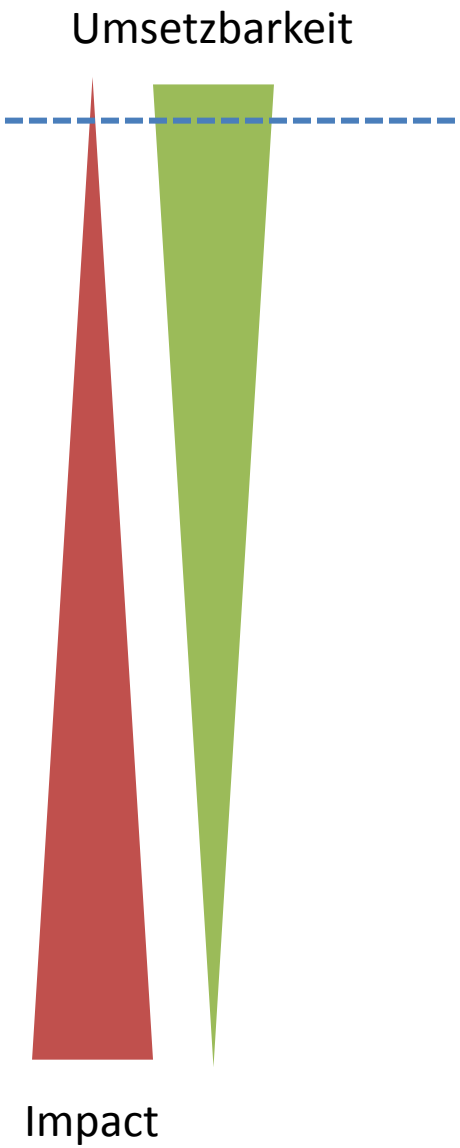
Outline

What is the problem?

Why do we have a problem?

How can we solve the problem?

How can we solve the problem?

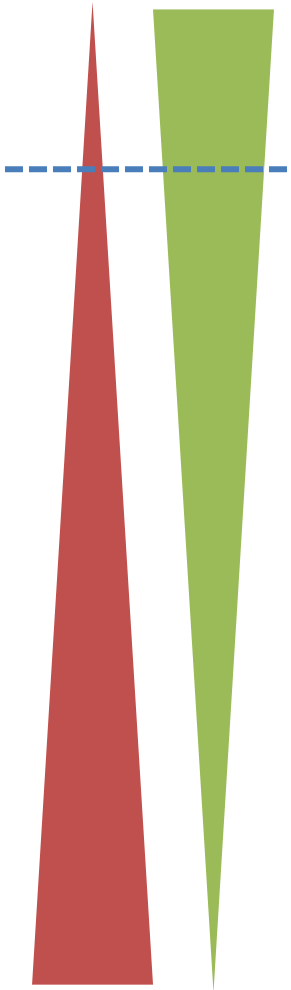


Open access Publikation

OPEN ACCESS

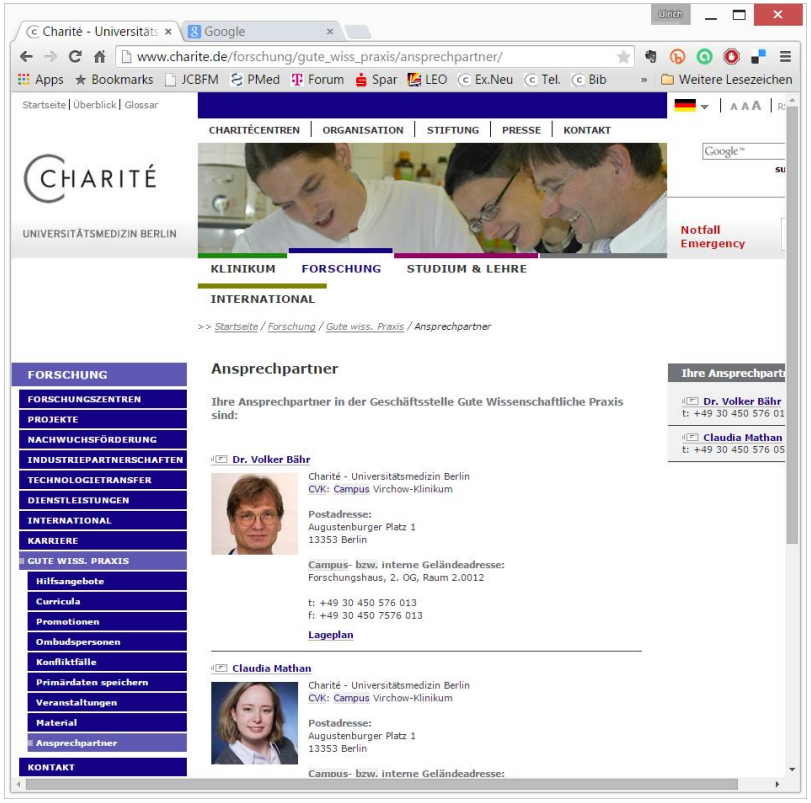
How can we solve the problem?

Umsetzbarkeit

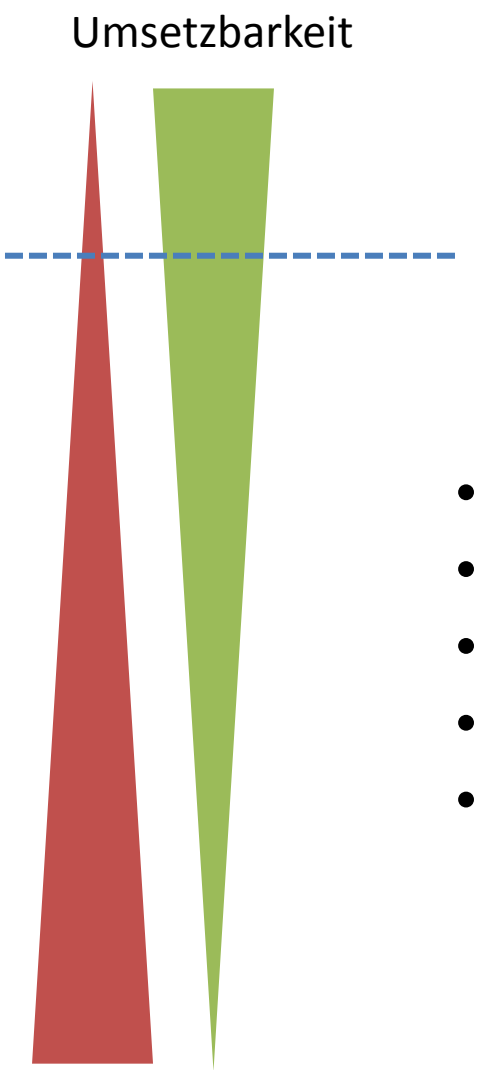


Impact

Good Scientific Practice Offices



How can we solve the problem?



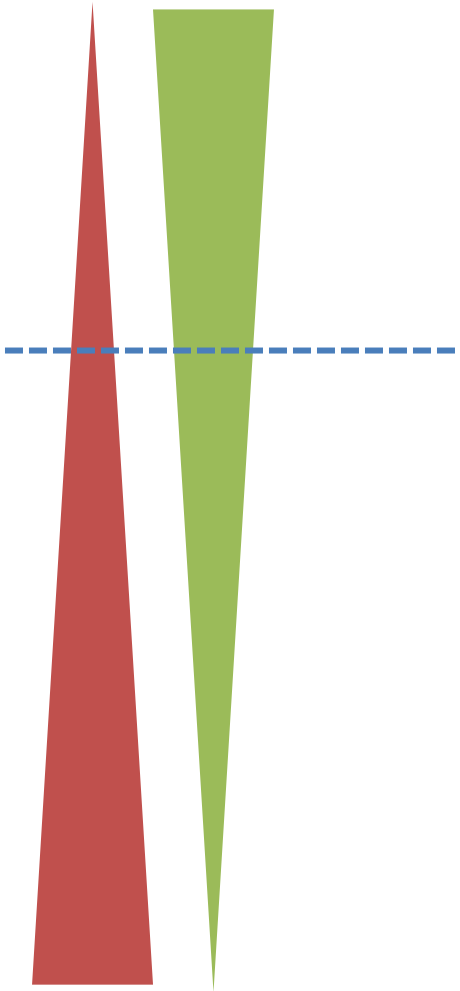
Ausbildung in Methodenkompetenz

- Biostatistics
- Clinical epidemiology
- Evidence based medicine
- Experimental design
- ....

Student .... Postdoc .... PI.... Director

How can we solve the problem?

Umsetzbarkeit



Impact

Compliance mit exist. Guidelines

BLINDING ✓

POWER ✓

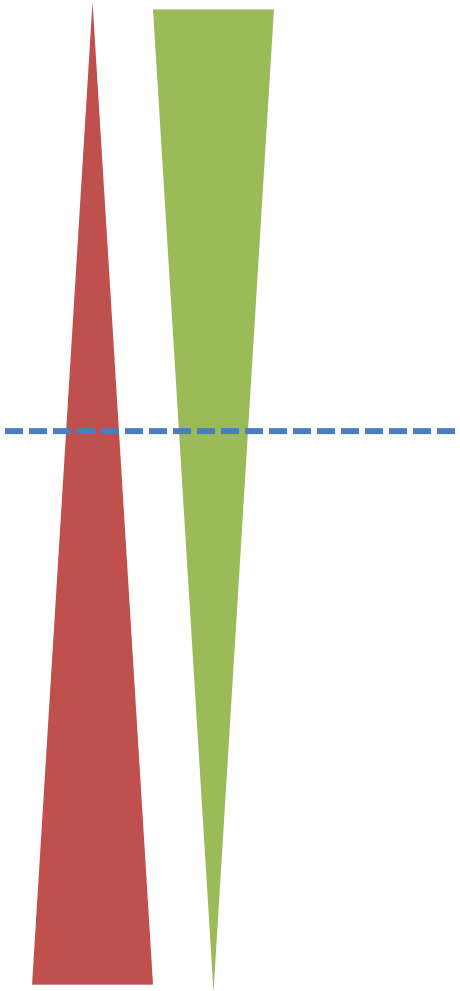
RANDOMISATION ✓

CONSORT

etc.

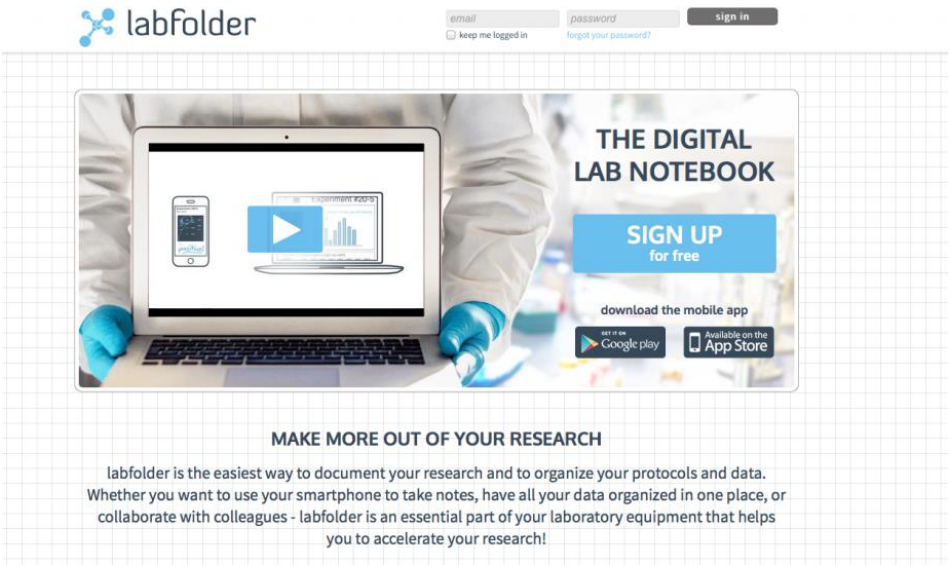
How can we solve the problem?

Umsetzbarkeit

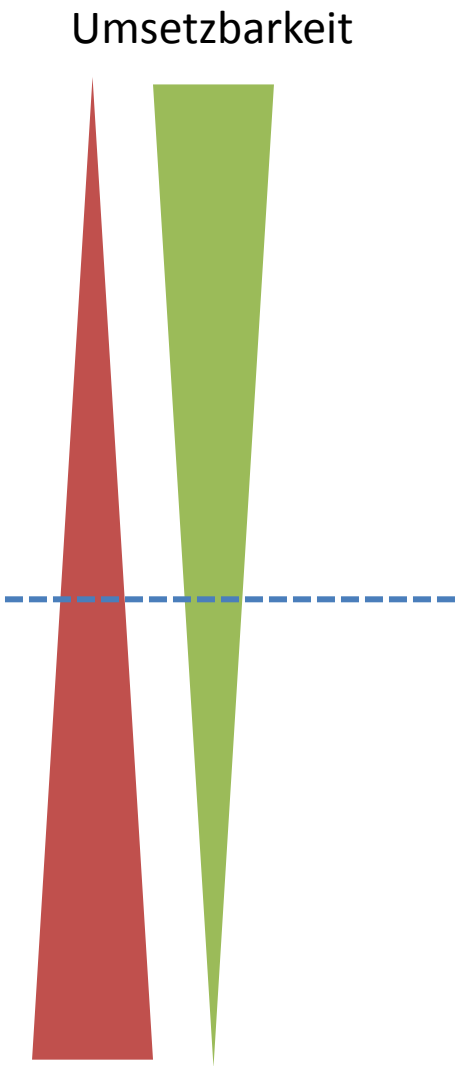


Impact

Elektronisches Laborbuch



How can we solve the problem?

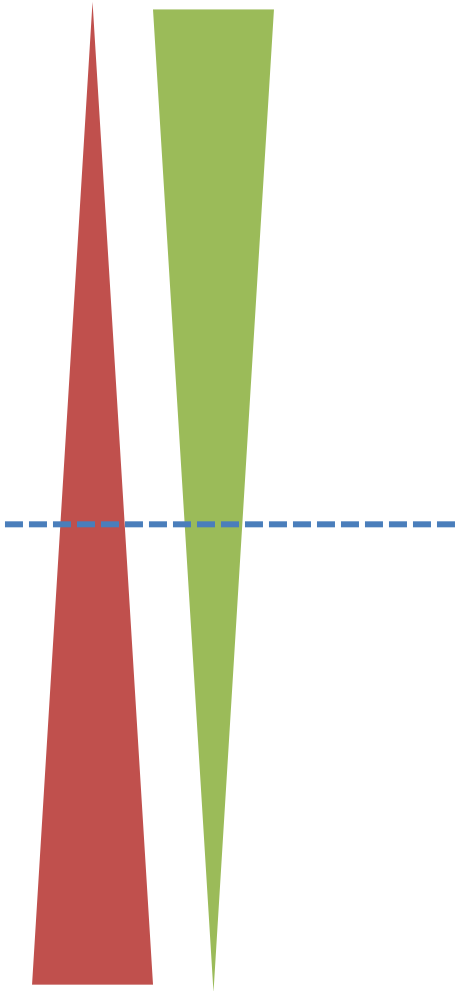


Open data / Repositorien / 'Negative' Studien

figshare
credit for all your research

How can we solve the problem?

Umsetzbarkeit



Impact

Pregregistration

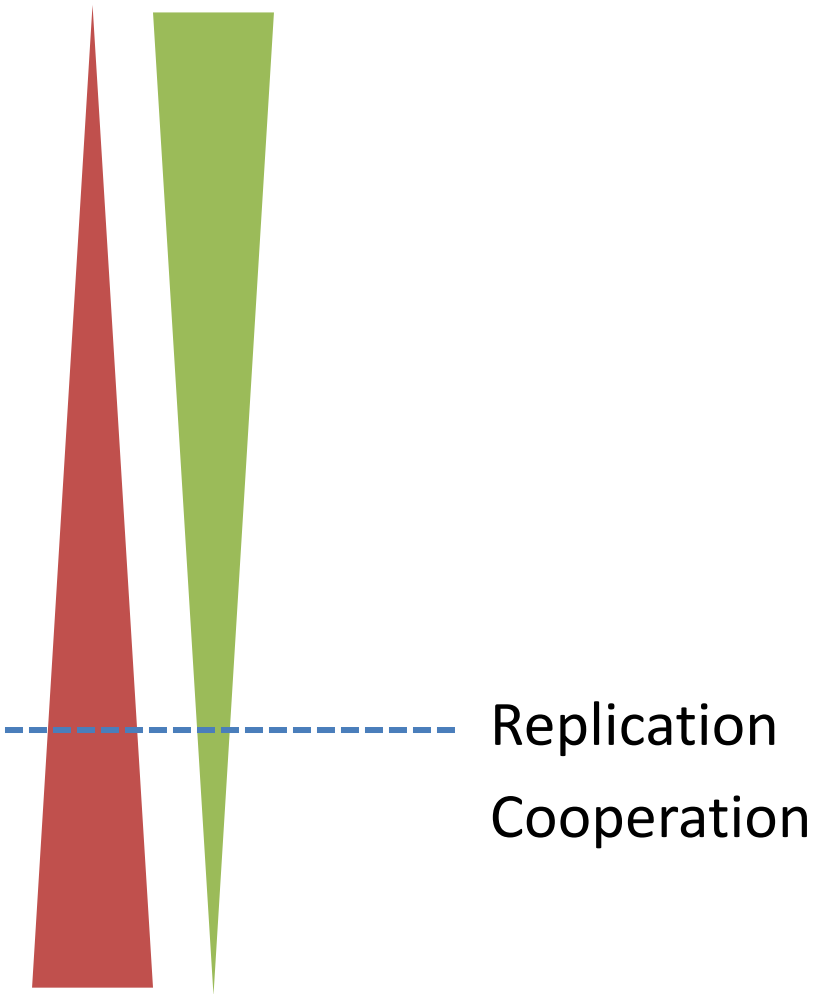
"Prior to inspection of the data, a preregistration protocol was published online http://confrepneurosci.blogspot.nl/2012/06/advanced-methods-and-analyses_26.html)."

Guidelines for reviewers

Registered Reports are a form of empirical article in which the methods and proposed analyses are pre-registered and reviewed prior to research being conducted. High quality protocols are then provisionally accepted for publication before data collection commences. This format of article is designed to reward best practice in adhering to the hypothetico-deductive model of the scientific method. It neutralises a number of questionable research practices, including low statistical power, selective reporting of results, and publication bias, while also allowing complete flexibility to conduct exploratory (unregistered) analyses and report serendipitous findings. (Chambers, 2013).

General reviewer guidelines can be found here: http://www.elsevier.com/reviewers/reviewer-guidelines

How can we solve the problem?



Replication

Cooperation

RESEARCH ARTICLE

STROKE

Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia

Science Translational Medicine
AAAS

SCIENTIFIC REPORTS

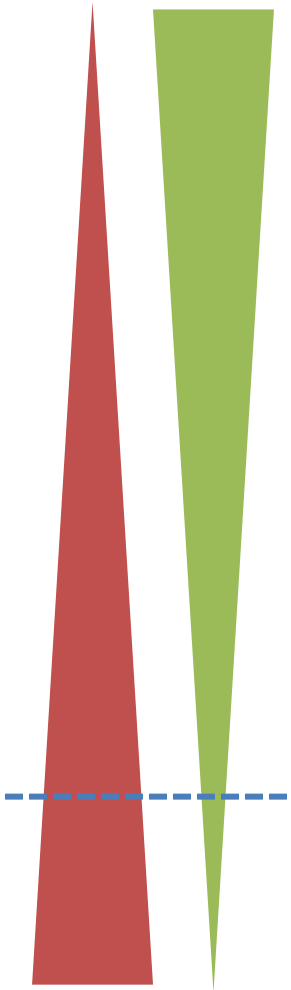
OPEN

A combined pre-clinical meta-analysis and randomized confirmatory trial approach to improve data validity for therapeutic target validation

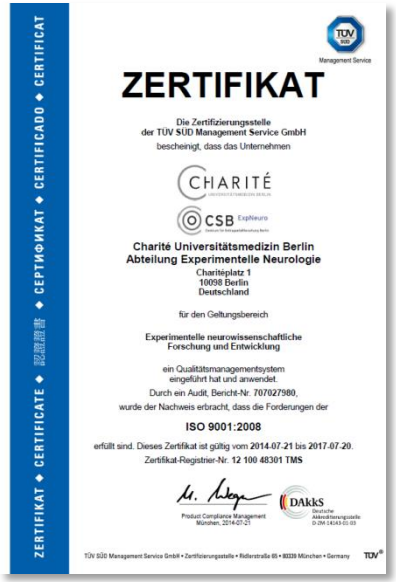
Received: 18 February 2015
Accepted: 27 July 2015
Published: 27 August 2015

How can we solve the problem?

Umsetzbarkeit



Impact



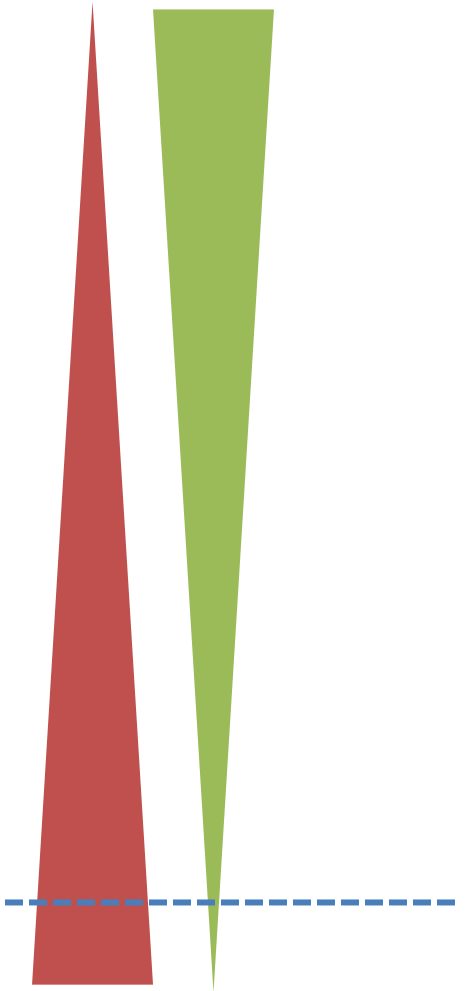
Strukturiertes Qualitätsmanagement

Critical incidence reporting (*Lab* CIRS / 'Morbidity & Mortality conferences')

(Peer-) Auditing ('Trust but verify')

How can we solve the problem?

Umsetzbarkeit



Impact

VIEWPOINT

Assessing Value in Biomedical Research
The PQRST of Appraisal and Reward

JAMA®
The Journal of the American Medical Association

Table. PQRST Index for Appraising and Rewarding Research

	Operationalization	
Item in PQRST Index	Example	Data Source
P (productivity)	Number of publications in the top tier % of citations for the scientific field and year	ISI Essential Science Indicators (automated)
	Proportion of funded proposals that have resulted in ≥1 published reports of the main results	Funding agency records and automated recording of acknowledged grants (eg, PubMed)
	Proportion of registered protocols that have been published 2 y after the completion of the studies	Study registries such as ClinicalTrials.gov for trials
Q (quality of scientific work)	Proportion of publications that fulfill ≥1 quality standards	Need to select standards (different per field/design) and may then automate to some extent; may limit to top-cited articles, if cumbersome
R (reproducibility of scientific work)	Proportion of publications that are reproducible	No wide-coverage automated database currently, but may be easy to build, especially if limited to the top-cited pivotal papers in each field
S (sharing of data and other resources)	Proportion of publications that share their data, materials, and/or protocols (whichever items are relevant)	No wide-coverage automated database currently, but may be easy to build, eg, embed in PubMed at the time of creation of PubMed record and update if more is shared later
T (translational influence of research)	Proportion of publications that have resulted in successful accomplishment of a distal translational milestone, eg, getting promising results in human trials for intervention tested in animals or cell cultures, or licensing of intervention for clinical trials	No wide-coverage automated database currently, would need to be curated by appraiser (eg, funding agency) and may need to be limited to top-cited papers, if cumbersome

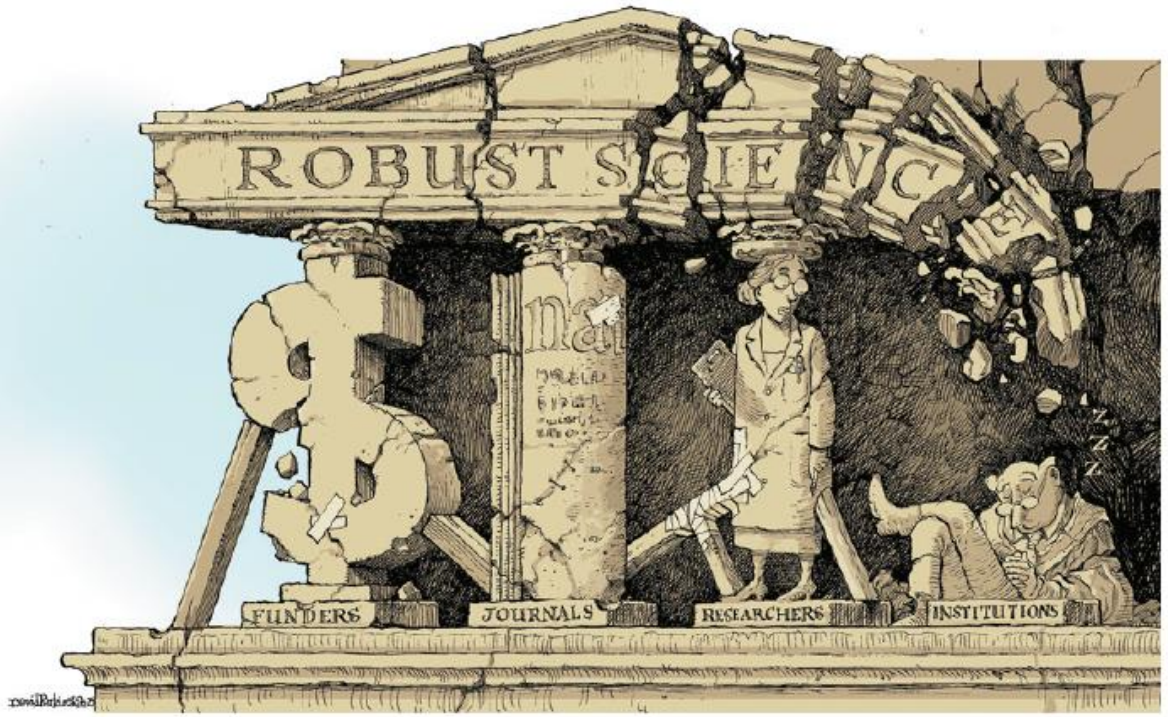
Ioannidis, Khoudri

Entwicklung und Implementierung neuer Indikatoren, Incentivierung (bzw. Disincentivierung)

Summary III: How can we solve the problem?

- Open access Publikation
- Ausbildung / Training (vom Studenten über PI zum Abteilungsleiter)
- Good Scientific Practice Office
- Compliance mit exist. Guidelines
- Elektronisches Laborbuch
- Open data / Repositorien / 'Neg. results'
- Preregistration
- Critical incidence Reporting (Lab CIRS / 'Morbidity & Mortality conferences')
- Replikation / Kooperation
- Strukturiertes Qualitätsmanagement
- (Peer-) Auditing
- Neue Indikatoren und Incentivierung (bzw. Disincentivierung)

nature
3 SEPTEMBER 2015 | VOL 525 | NATURE | 25



Institutions must do their part for reproducibility

Tie funding to verified good institutional practice, and robust science will shoot up the agenda, say **C. Glenn Begley**, **Alastair M. Buchan** and **Ulrich Dirnagl**.