

Warum scheitern so viele im Labor hoffnungsvolle Therapien in der klinischen Prüfung?



Berlin, 24.9.2015



Center for Stroke
Research Berlin

What is the problem?

Why do we have a problem?

How can we solve the problem?

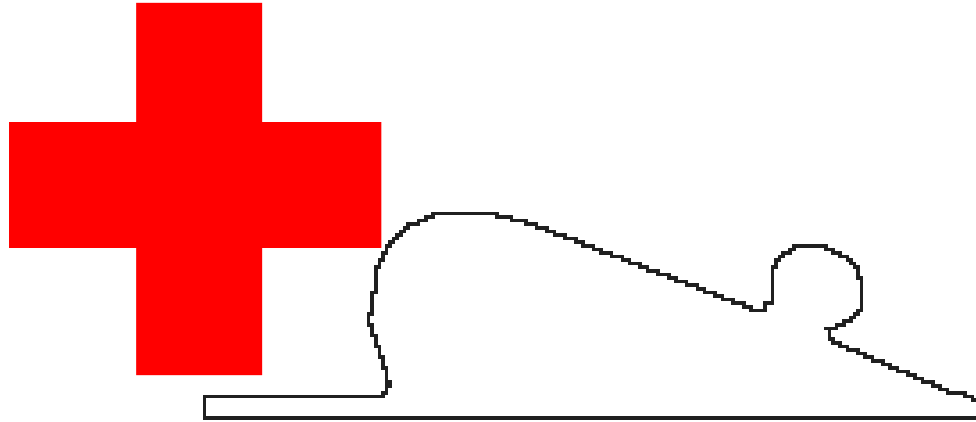
What is the problem?

Why do we have a problem?

How can we solve the problem?

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

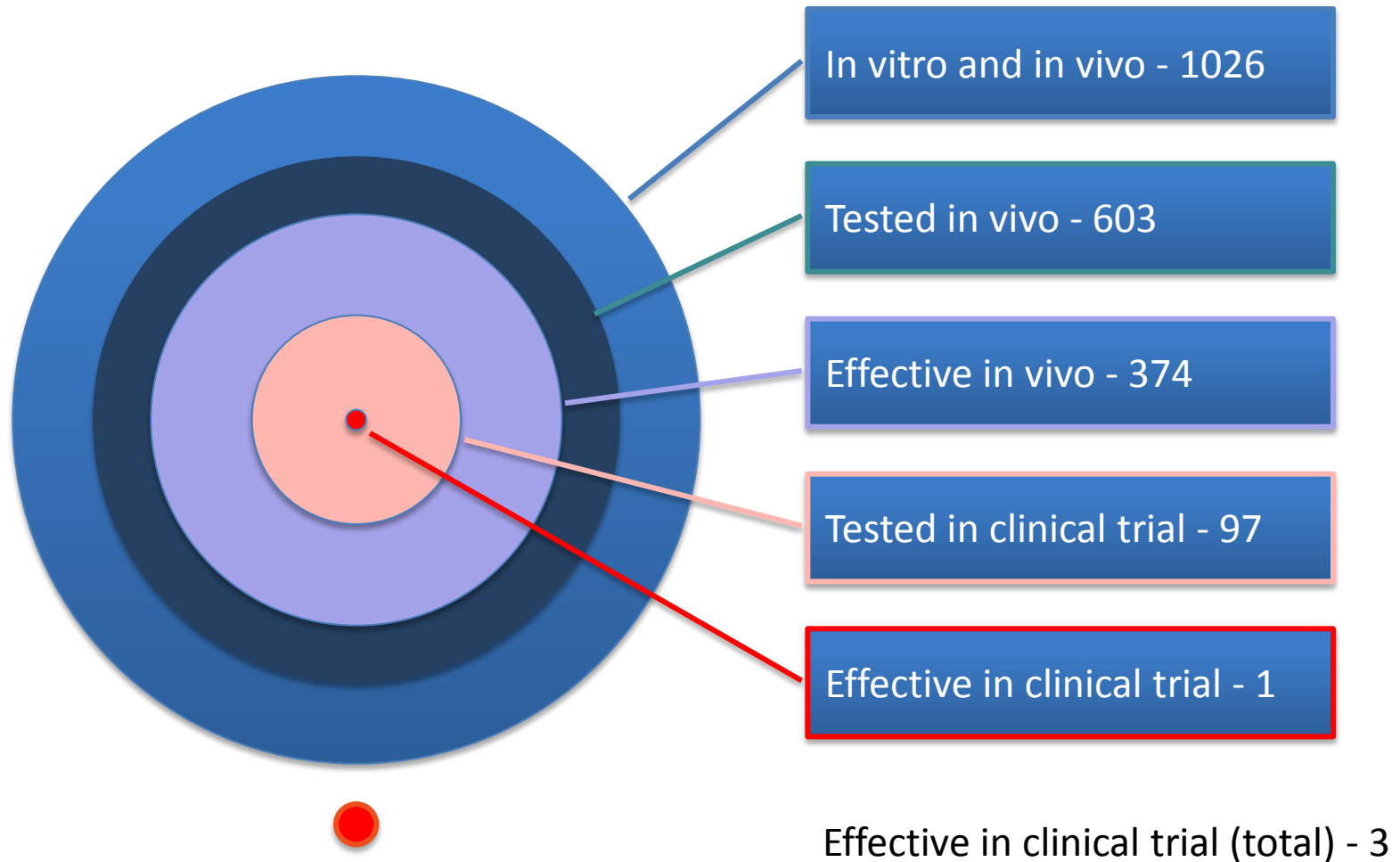
Lindsay Symon, Neurosurgeon



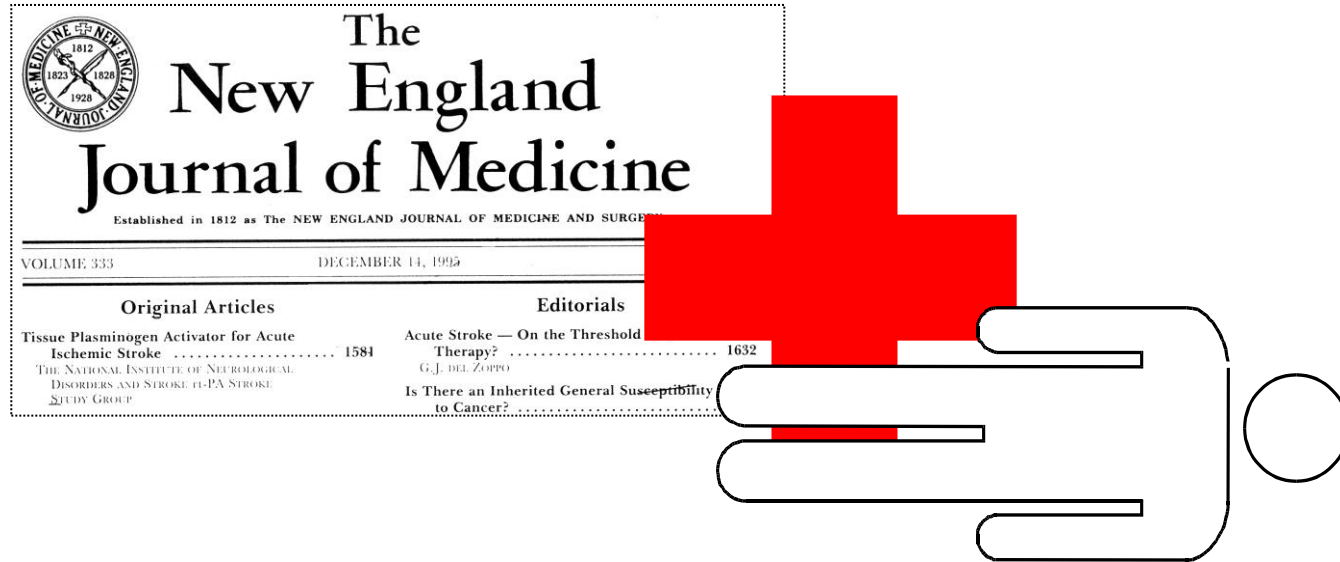
A typical intervention in exp. stroke studies reduces infarct sizes by 30-50 %.

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

1026 interventions in experimental stroke



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.

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





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

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

JAMA

The Journal of the American Medical Association

Commentary

Lost in Translation

Bumps in the Road Between Bench and Bedside

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.

© 2006 Elsevier Inc. All rights reserved.

of
of
vas
of
cal
tic
ted
art-
t),
in
Ds
up
ors

OPEN ACCESS Freely available online

PLOS CLINICAL TRIALS

Essay

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

John P. A. Ioannidis

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



RESEARCH

Open Access

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

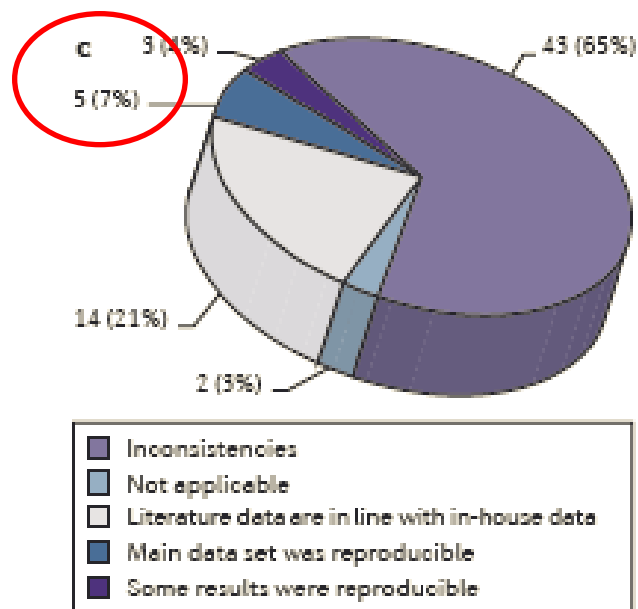
Jeffrey L Cummings^{1*}, Travis Morstorf² and Kate Zhong¹

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.

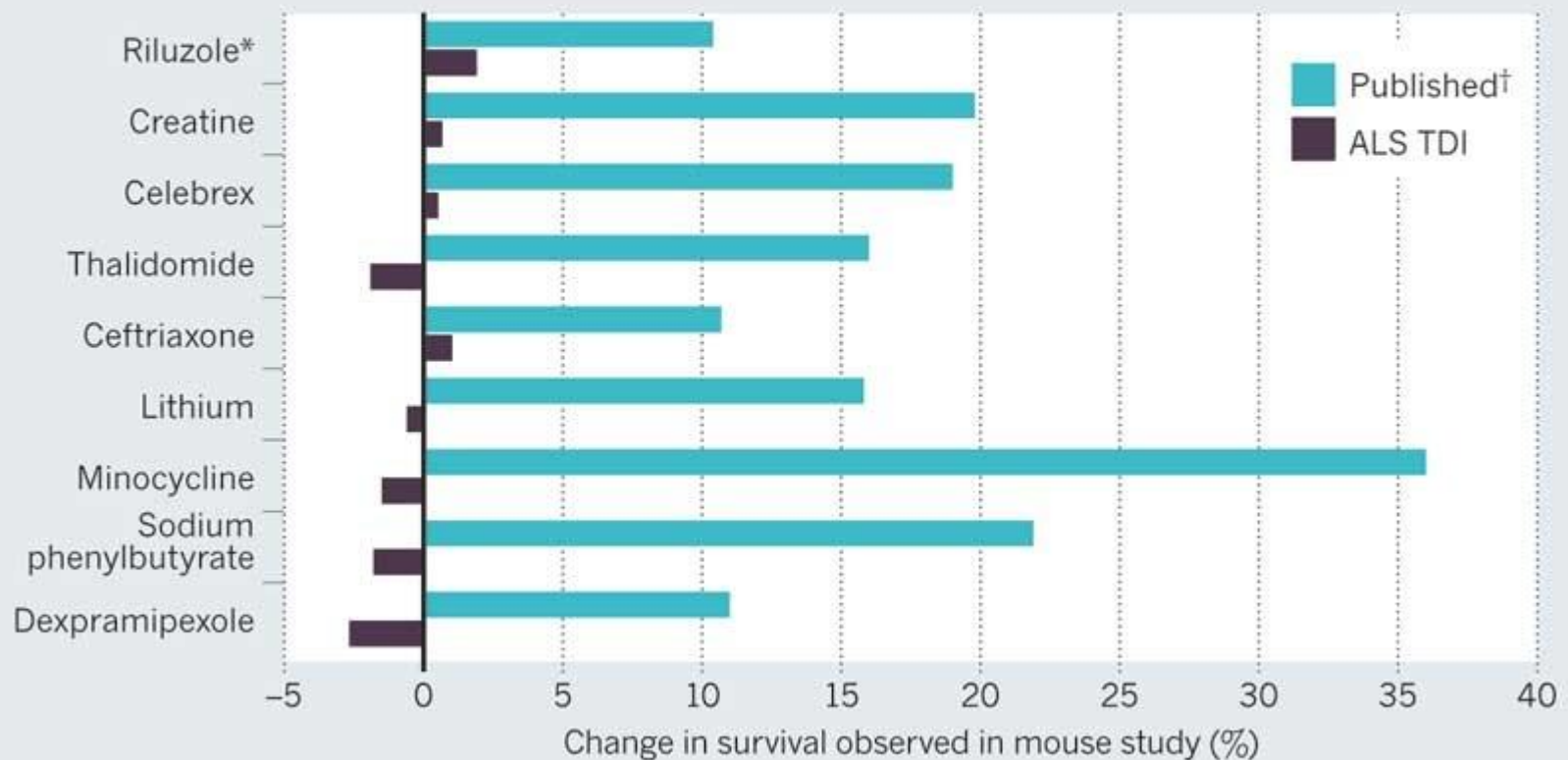
REPRODUCIBILITY OF RESEARCH FINDINGS

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)

DUE DILIGENCE, OVERDUE

Results of rigorous animal tests by the Amyotrophic Lateral Sclerosis Therapy Development Institute (ALS TDI) are less promising than those published. All these compounds have disappointed in human testing.



*Although riluzole is the only drug currently approved by the US Food and Drug Administration for ALS, our work showed no survival benefit.

†References for published studies can be found in supplementary information at go.nature.com/hf4jf6.

Special Issue: NIH Replication Studies

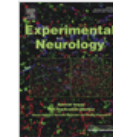
Edited By Oswald Steward and Phillip Popovich



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 Kleitmanⁱ

[+ 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

PERSPECTIVE

The Economics of Reproducibility in Preclinical Research

Leonard P. Freedman^{1*}, Iain M. Cockburn², 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

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.



 OPEN ACCESS

Citation: Freedman LP, Cockburn IM, Simcoe TS (2015) The Economics of Reproducibility in Preclinical Research. *PLOS ONE* 10(12): e0143317. doi:10.1371/journal.pone.0143317

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

What is the problem?

Why do we have a problem?

How can we solve the problem?

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 false findings in the field. In this paper, we use a mathematical model to estimate the probability that a research claim is true given the following factors: some

factor
some

Mod
Posit

Sever
point
rate c

OPEN ACCESS Freely available online

PLOS BIOLOGY

Evaluation of Excess Significance Bias in Animal Studies of Neurological Diseases

Konstantinos K. Tsilidis^{1,9}, Orestis A. Panagiotou^{1,9}, Emily S. Sena^{2,3}, Eleni Aretouli^{4,5}, Evangelos Evangelou¹, David W. Howells³, Rustam Al-Shahi Salman², Malcolm R. Macleod², 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,

Department of Methods
University of Thessaloniki,
University School of

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

Katherine S. Button^{1,2}, John P. A. Ioannidis³, Jonathan Flint⁵, Emma S. J. Robinson⁶

Nature Reviews Neuroscience | AOP, published online 12 October 2014

trials. We
l disorders,
eta-Analysis
statistically
tudy under

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

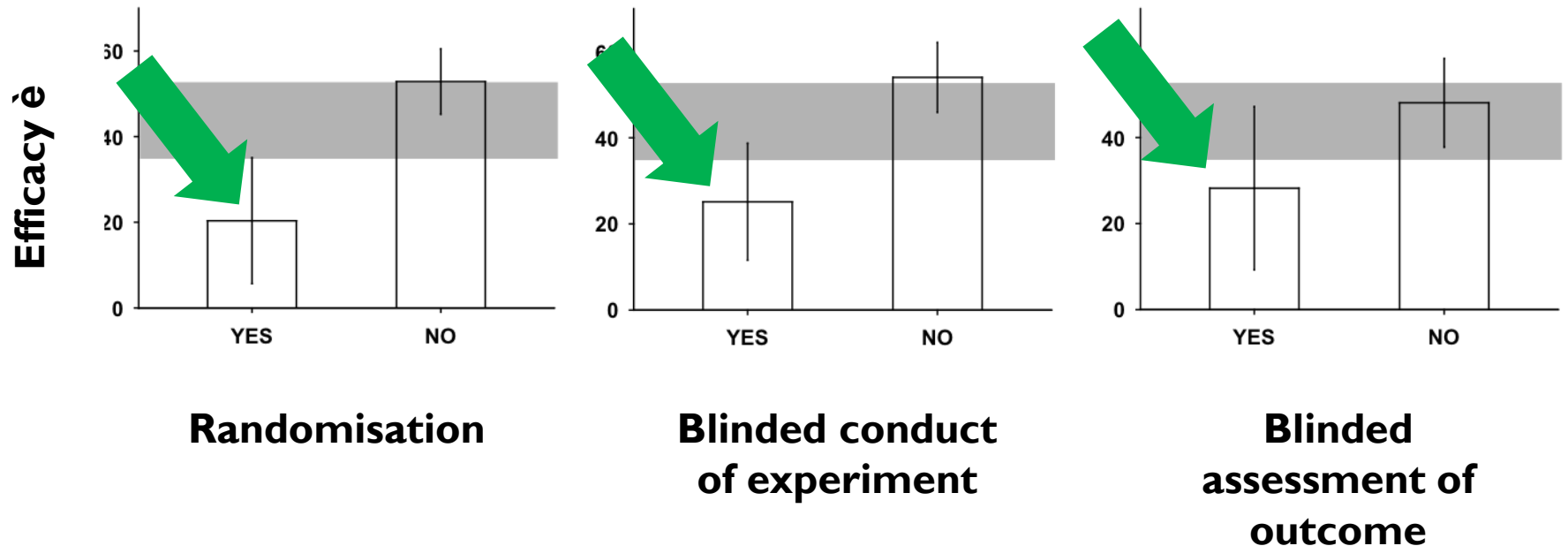
Background and Purpose—At least part of the failure in the transition from experimental to clinical studies in stroke has been attributed to the imprecision introduced by problems in the design of experimental stroke studies. Using a metaepidemiologic approach, we addressed the effect of randomization, 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

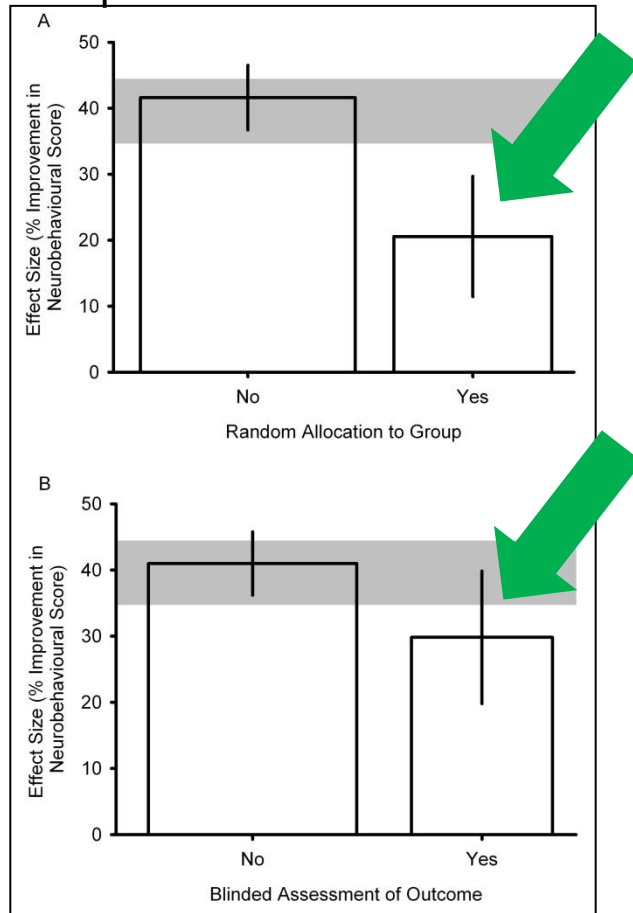
Effect size inversely correlates with study quality

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

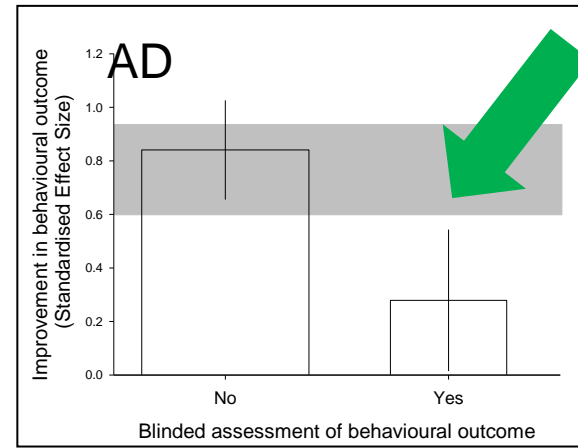


Low internal validity: Meta-research exposes selection and performance bias

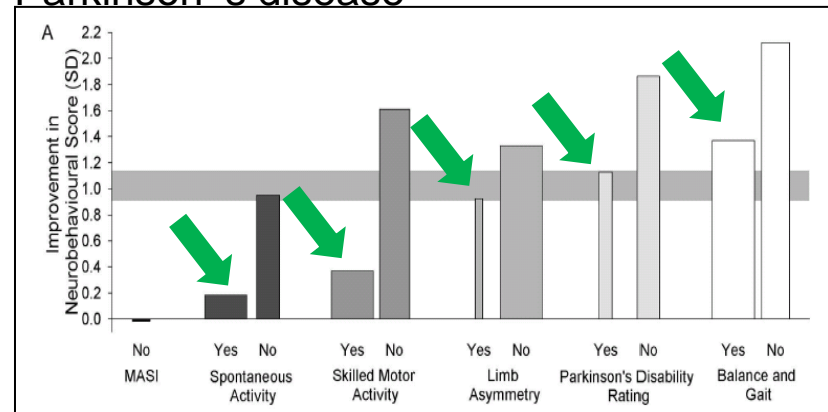
Multiple Sclerosis



Alzheimer's disease



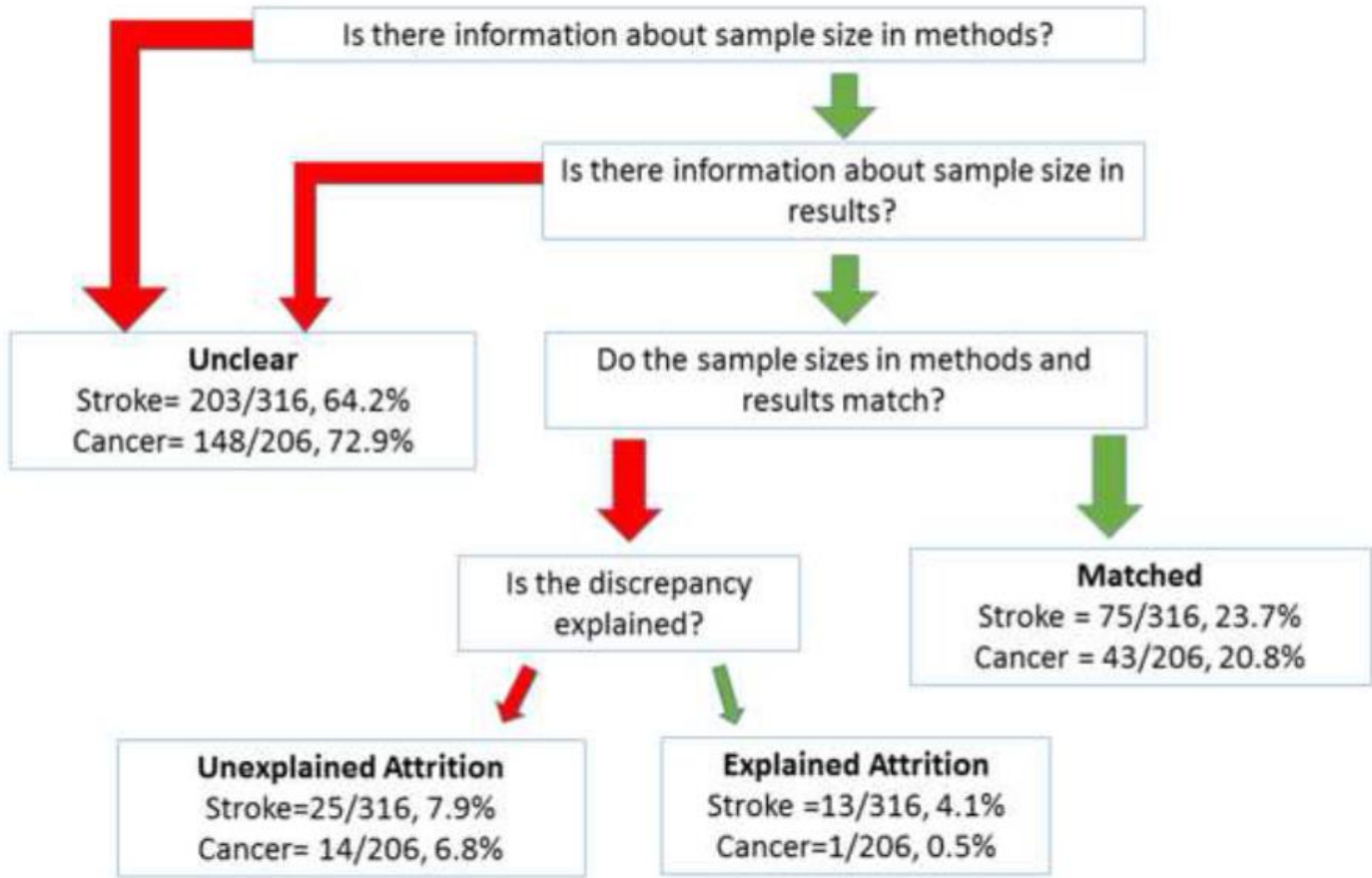
Parkinson's disease



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

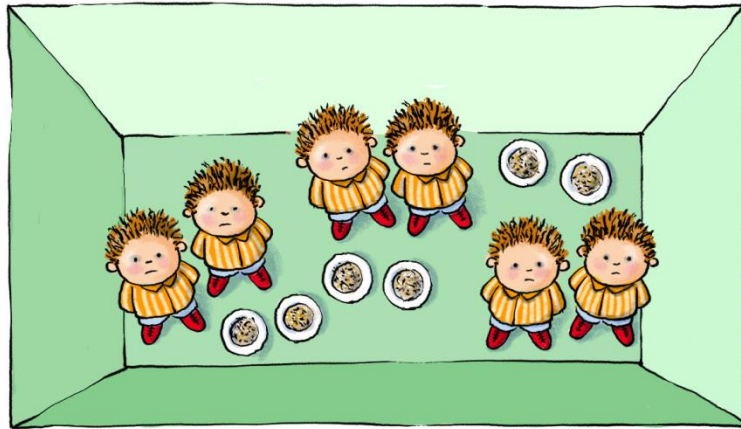
Disease modelled		Number of Publications	Sample Size Calculation (%)	Random Allocation to Group (%)	Blinded conduct of experiment (%)	Blinded Assessment of Outcome (%)
Alzheimer's Disease		428	0	16	n/a	22
Multiple Sclerosis		1117	<1	9	n/a	16
Parkinson's Disease		252	<1	16	n/a	15
Intracerebral Haemorrhage		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)



- 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 m² isolator tents on an enriched granola diet

vs.

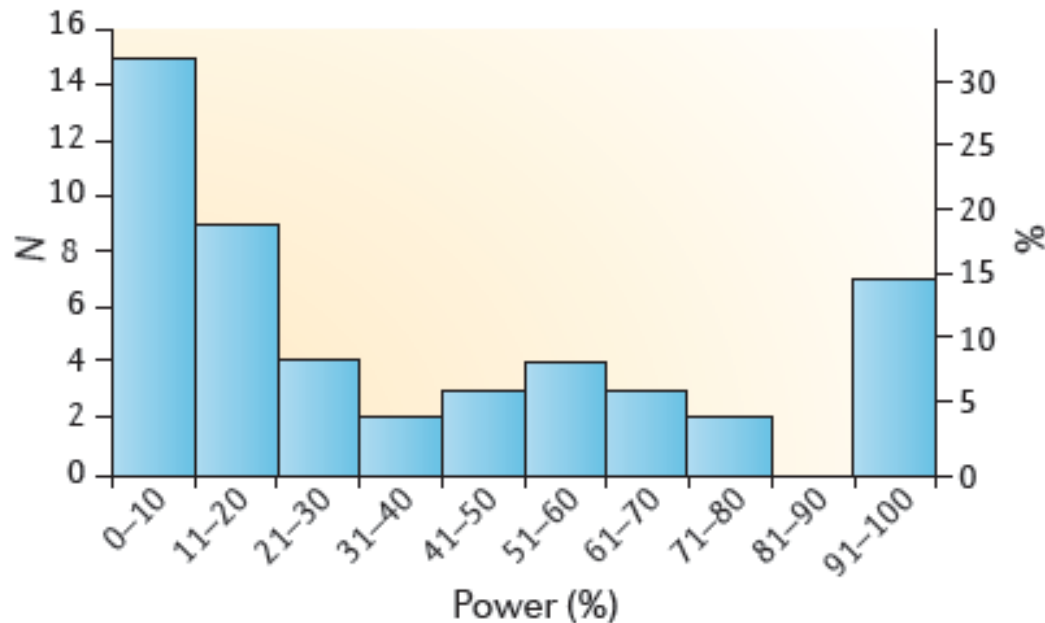


Patients of both sexes, elderly, comorbid, multiple medications, exposed to multiple pathogens and antigens throughout life

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

Katherine S. Button^{1,2}, John P. A. Ioannidis³, Claire Mokrysz¹, Brian A. Nosek⁴, Jonathan Flint⁵, Emma S. J. Robinson⁶ and Marcus R. Munafò¹

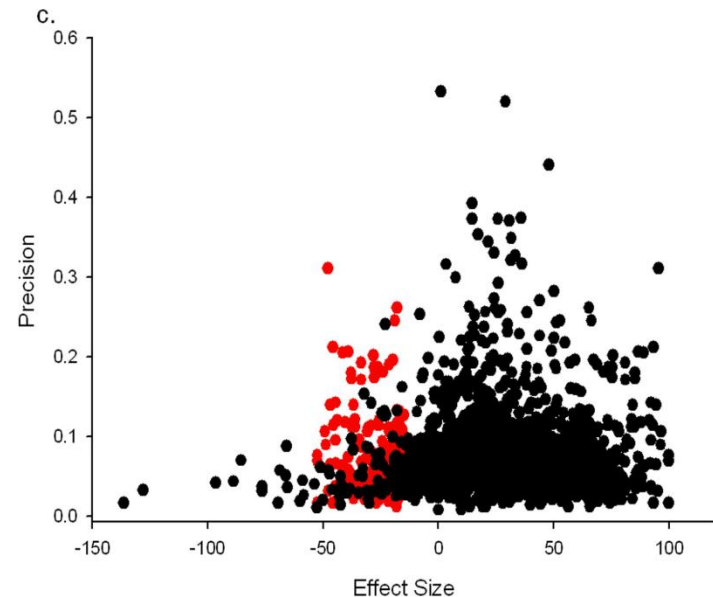
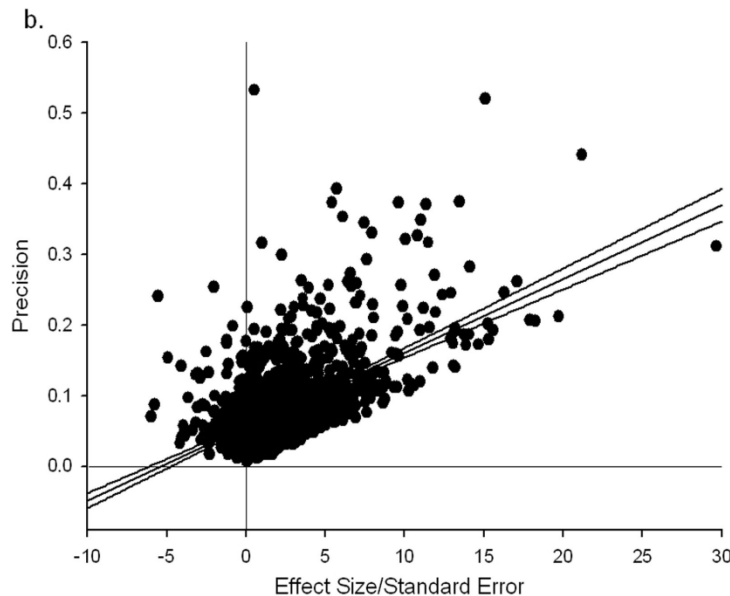
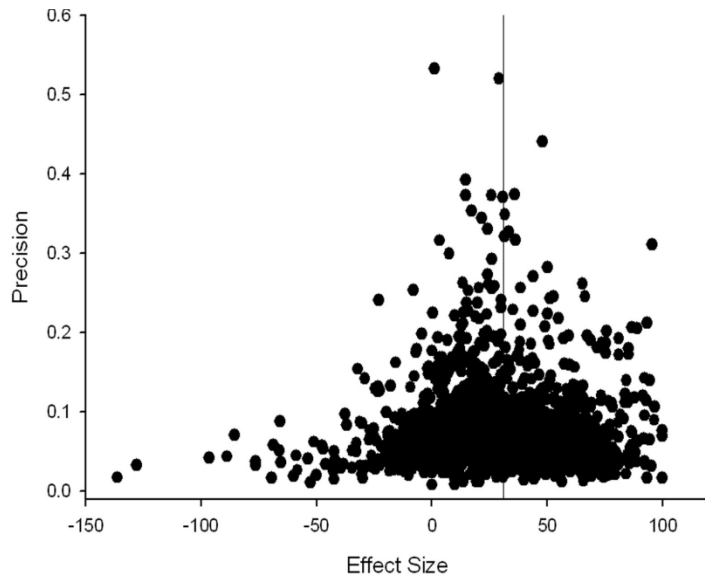
Nature Reviews Neuroscience | AOP, published online 10 April 2013; doi:10.1038/nrn3475



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."

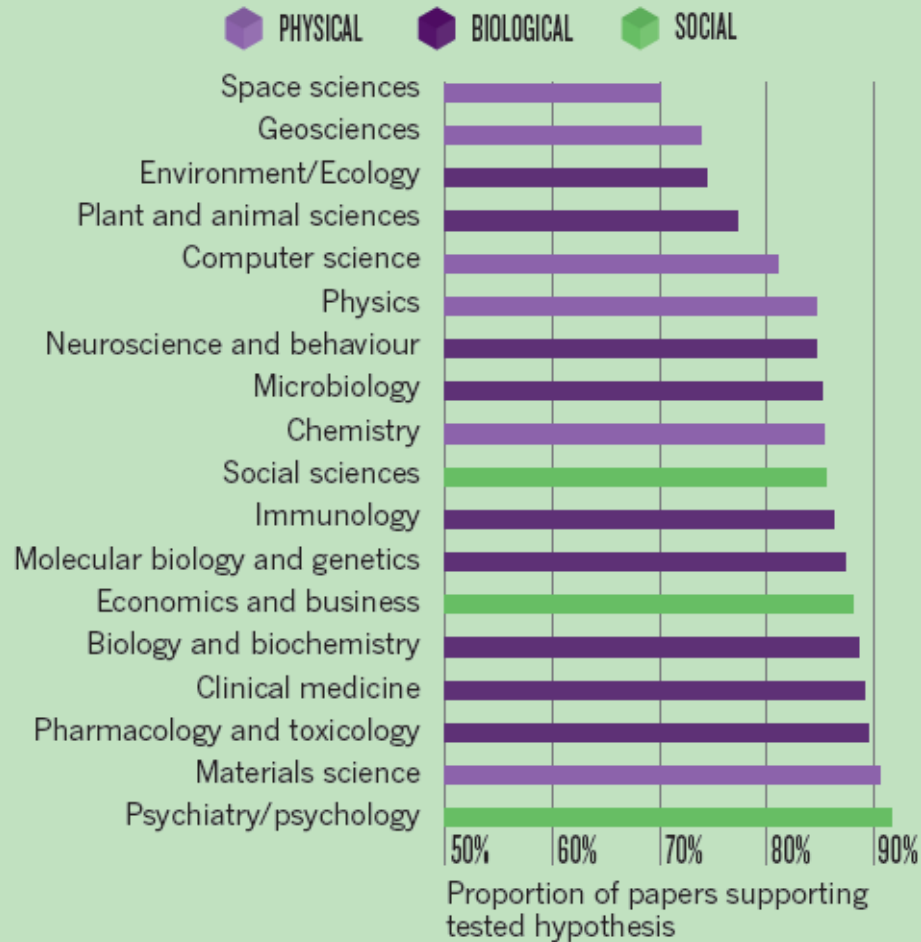


Publication bias in reports of animal stroke studies leads to major overstatement of efficacy

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)

ACCENTUATE THE POSITIVE

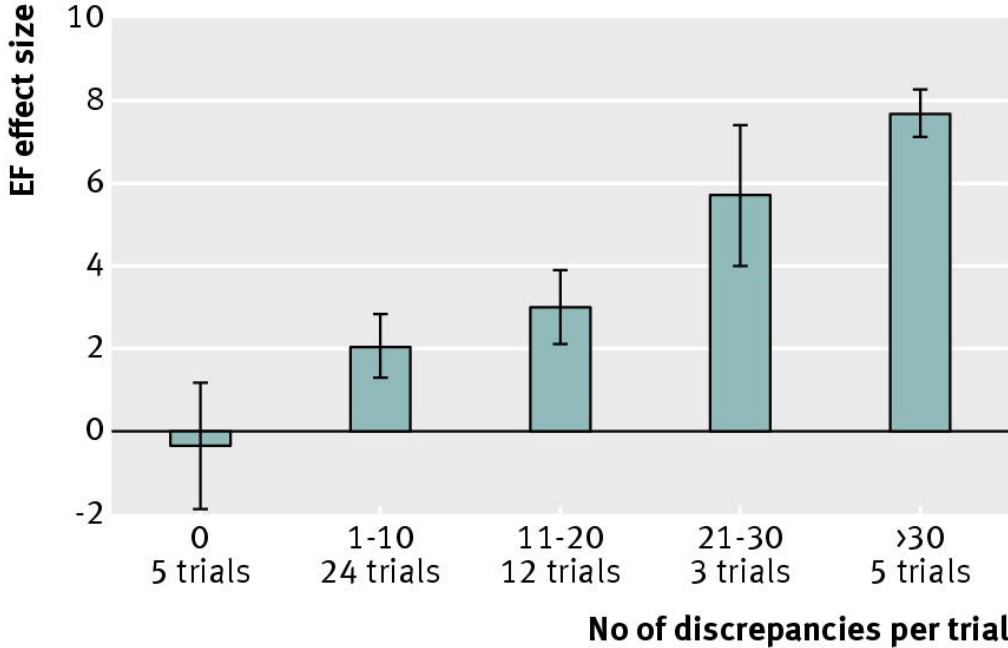
A literature analysis across disciplines reveals a tendency to publish only 'positive' studies — those that support the tested hypothesis. Psychiatry and psychology are the worst offenders.



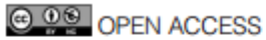


RESEARCH

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



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-Jaoude⁵

¹School of Medical Sciences, Bangor University, Bangor, Wales, UK

²Emory University, Atlanta, Georgia, USA

³Critical and Ethical Mental Health Research Group.

Robir

Adela

Austr

Joan

of He

Adela

Austr

⁵Dep

Hosp

Unive

Ontar

Corre

Jon Ju

Addi

online

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 efficacy and safety of paroxetine and imipramine with placebo

in the treatment of major depressive disorder in adolescents. The study was done to support the marketing of paroxetine in adolescents. The dataset from the study was reanalysed to determine whether clinically significant differences in efficacy and safety were observed.

The study was a randomised, double-blind, placebo-controlled trial.

The study was conducted at several sites, including centres, from

at least

ia included

al disorde

of at least

of at least

of at least

of at least

The screenshot shows the homepage of the Restoring Study 329 website. The header features the title "Restoring Study 329" in large red letters, with the tagline "SCIENTIFIC INTEGRITY THROUGH DATA BASED MEDICINE" below it. A navigation menu includes "HOME", "BACKGROUND", "MEDIA", "TEAM", and "WHAT CAN I DO?". The main content area has a headline: "BMJ Publishes Study Revealing How Flawed Drug Research Fails a Trusting Public". Below this is a sub-headline: "Band of Intrepid Researchers 'Sets the Record Straight' on Ghostwritten Study". A paragraph of text follows, starting with "Toronto (September 16, 2015) — Today the BMJ published Restoring Study 329, a decade-long effort by researchers to uncover the truth about the safety of an antidepressant approved for use by adolescents. Restoring Study 329 is a reanalysis and rebuttal of the original Study 329 ... more". A link "See news coverage" is provided. At the bottom, another headline reads "Medicine's Most Infamous Clinical Trial" with a sub-headline "Study 329 had all the makings of a scandal." and a partial sentence "Thousands of North American children and".

The screenshot shows a web browser displaying an article on the Spiegel Online website. The browser's address bar shows the URL "www.spiegel.de/gesundheits/diagnose/antidepressiva-paroxetin-imipramin-kein-nutzen". The page header includes the "SPIEGEL ONLINE GESUNDHEIT" logo and navigation links for "Home", "Video", "Themen", "Forum", "English", "DER SPIEGEL", "SPIEGEL TV", "Abo", and "Shop". Below the header is a search bar and a list of categories: "Politik | Wirtschaft | Panorama | Sport | Kultur | Netzwerk | Wissenschaft | Gesundheit | einestages | Karriere | Uni | Reise | Auto | Stil". The main article title is "Depressive Jugendliche: Antidepressiva verschlimmern Gesundheitszustand". Below the title is a photograph of a young man covering his face with his hand, looking distressed. The text below the photo reads: "Deprimierende Wirkung: Antidepressiva sind bei Jugendlichen weder wirksam noch sicher". At the bottom, a sub-headline states: "Die neue Auswertung von Studiendaten bringt Hersteller von Antidepressiva in Erklärungsnot: Zwei weitverbreitete Medikamente wirken bei Jugendlichen nicht. Stattdessen drohen starke Nebenwirkungen."

- 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

What is the problem?

Why do we have a problem?

How can we solve the problem?

Umsetzbarkeit

Open access Publikation



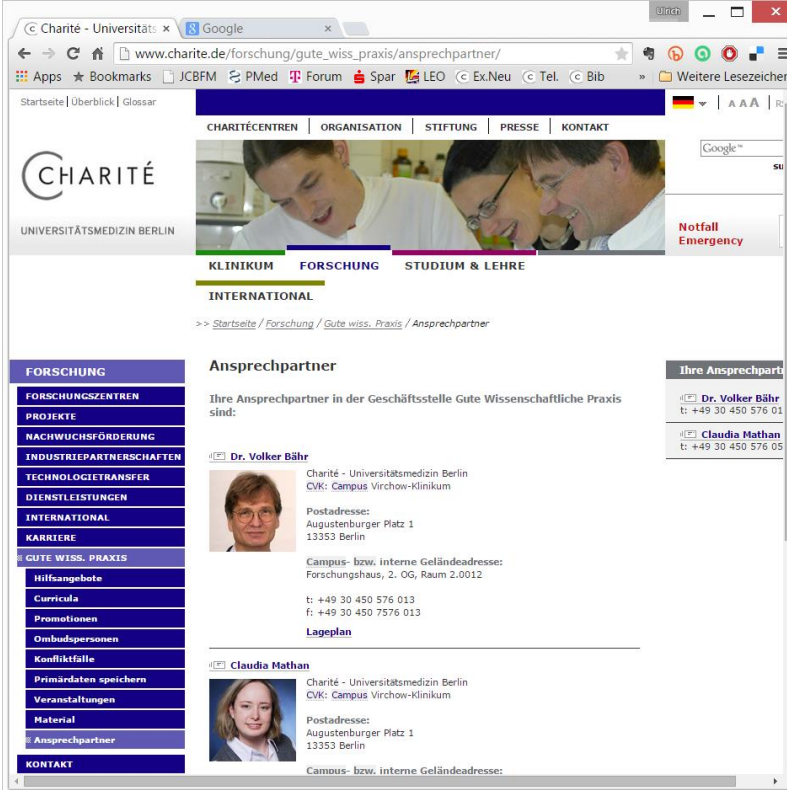
OPEN  **ACCESS**

Impact

How can we solve the problem?

Umsetzbarkeit

Good Scientific Practice Offices



The screenshot shows a web browser window displaying the Charité website. The URL is www.charite.de/forschung/gute_wiss_praxis/ansprechpartner/. The page is titled "Anspruchspartner" and provides contact information for two individuals: Dr. Volker Bähr and Claudia Mathan. The website layout includes a navigation menu at the top with links for "CHARITÉCENTREN", "ORGANISATION", "STIFTUNG", "PRESSE", and "KONTAKT". Below the navigation is a search bar and a "Notfall Emergency" button. The main content area features a "FORSCHUNG" sidebar with various sub-links, a central "Anspruchspartner" section with contact details, and a "Ihre Ansprechpartner" sidebar listing the contact persons.

CHARITÉ
UNIVERSITÄTSMEDIZIN BERLIN

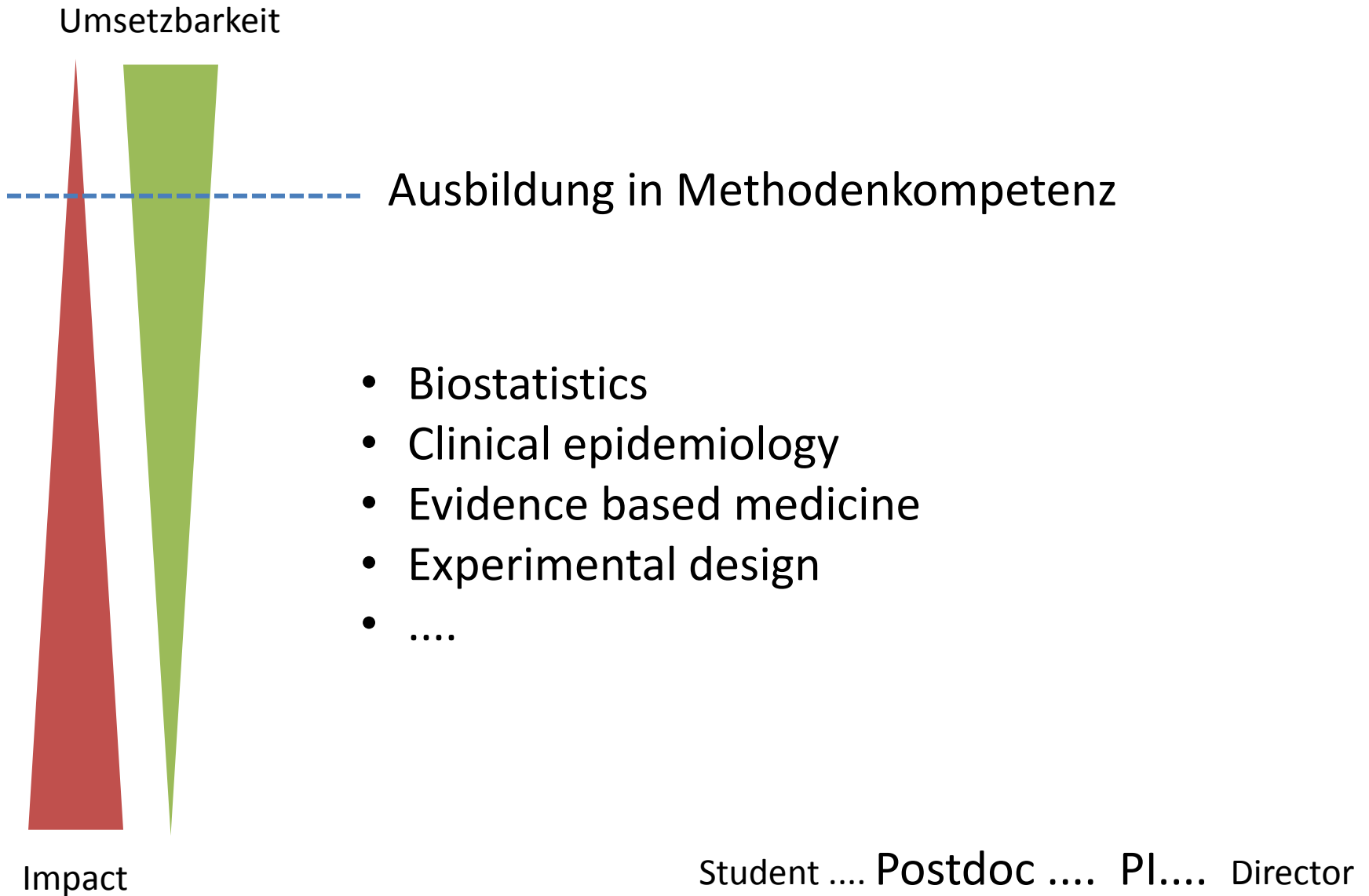
ANSPRUCHSPARTNER

Ihre Ansprechpartner in der Geschäftsstelle Gute Wissenschaftliche Praxis sind:

Dr. Volker Bähr
Charité - Universitätsmedizin Berlin
CVK: Campus Virchow-Klinikum
Postadresse:
Augustenburger Platz 1
13353 Berlin
Campus- bzw. interne Geländeadresse:
Forschungshaus, 2. OG, Raum 2.0012
t: +49 30 450 576 013
f: +49 30 450 7576 013
Lageplan

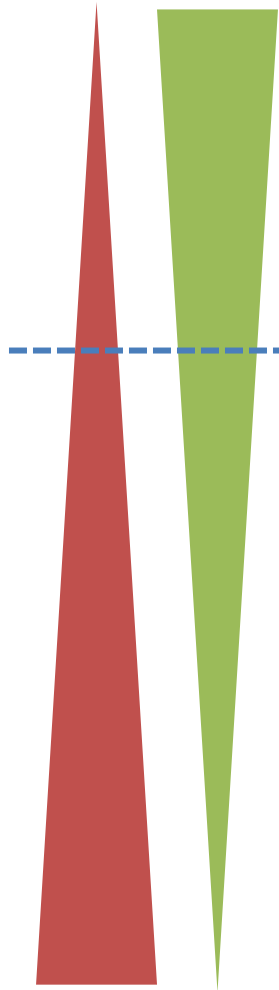
Claudia Mathan
Charité - Universitätsmedizin Berlin
CVK: Campus Virchow-Klinikum
Postadresse:
Augustenburger Platz 1
13353 Berlin
Campus- bzw. interne Geländeadresse:

Impact



How can we solve the problem?

Umsetzbarkeit



Compliance mit exist. Guidelines



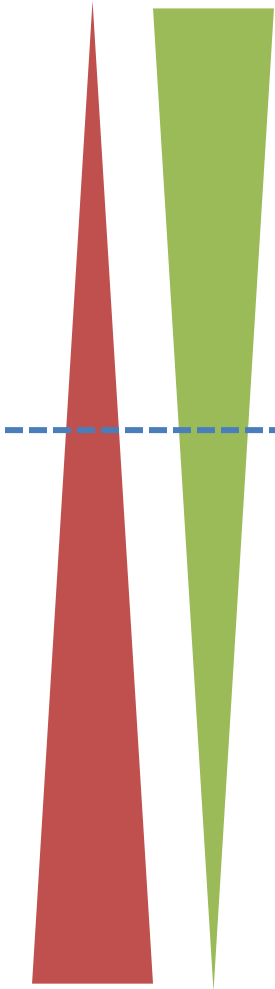
- BLINDING
- POWER
- RANDOMISATION



etc.

Impact

Umsetzbarkeit

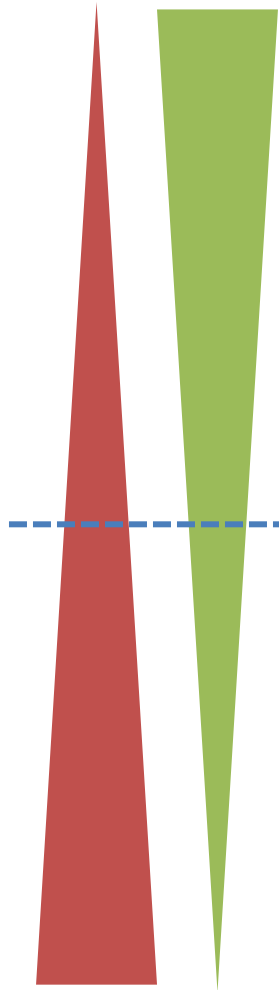


Elektronisches Laborbuch

The screenshot shows the Labfolder website interface. At the top left is the Labfolder logo. To the right are input fields for 'email' and 'password', a 'sign in' button, and a 'keep me logged in' checkbox. Below the login form is a large promotional banner for 'THE DIGITAL LAB NOTEBOOK'. The banner features a central image of a person in a lab coat using a laptop, with a smartphone and a tablet also displaying the app interface. A prominent blue button says 'SIGN UP for free'. Below this, it says 'download the mobile app' with icons for Google Play and the App Store. At the bottom of the banner, the text reads 'MAKE MORE OUT OF YOUR RESEARCH' followed by a paragraph: 'labfolder is the easiest way to document your research and to organize your protocols and data. Whether you want to use your smartphone to take notes, have all your data organized in one place, or collaborate with colleagues - labfolder is an essential part of your laboratory equipment that helps you to accelerate your research!'

Impact

Umsetzbarkeit

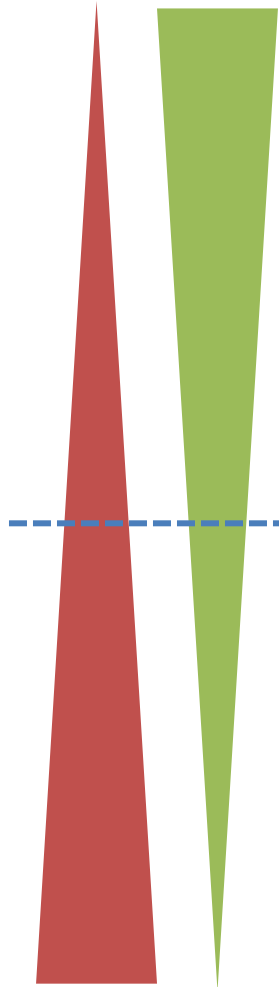


Open data / Repositorien / 'Negative' Studien



Impact

Umsetzbarkeit



Impact

"Prior to inspection of the data, a preregistration protocol was published online

http://confrepneurosci.blogspot.nl/2012/06/advanced-methods-and-analyses_26.html)."

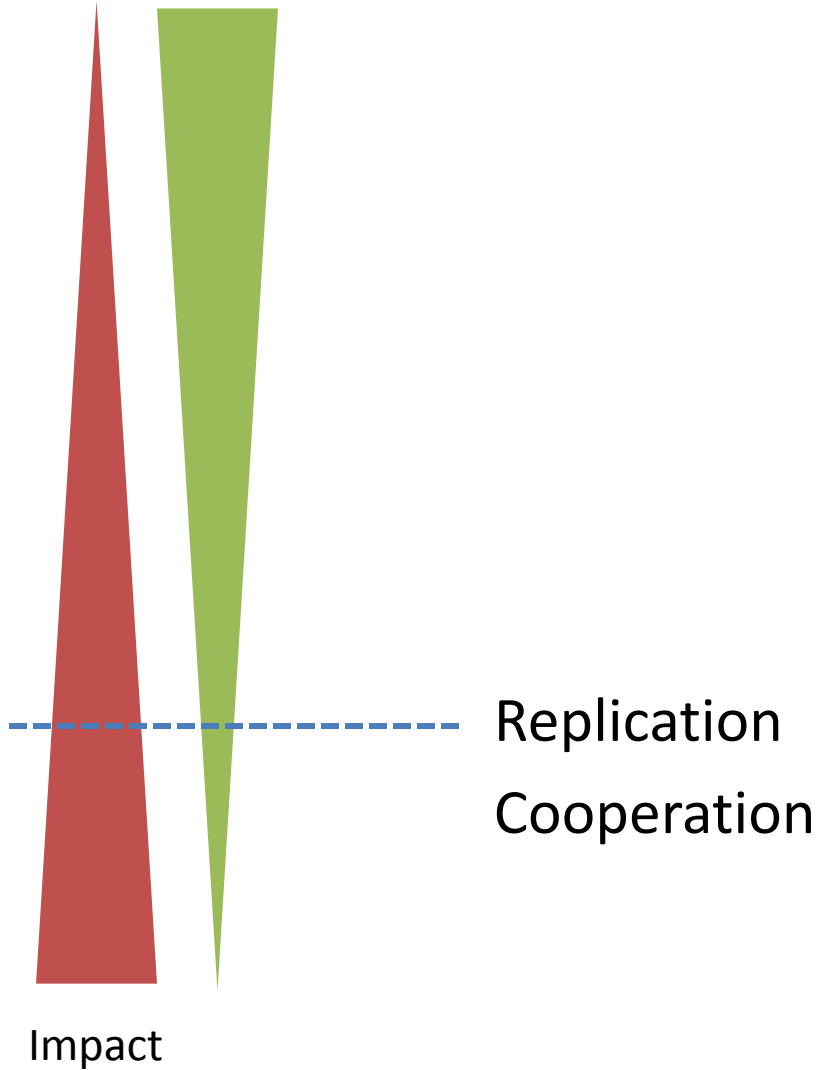
Pregeneration

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>

Umsetzbarkeit



RESEARCH ARTICLE



STROKE

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

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



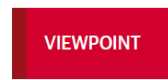
Strukturiertes Qualitätsmanagement

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

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

Impact

Umsetzbarkeit



Assessing Value in Biomedical Research The PQRST of Appraisal and Reward



Table. PQRST Index for Appraising and Rewarding Research

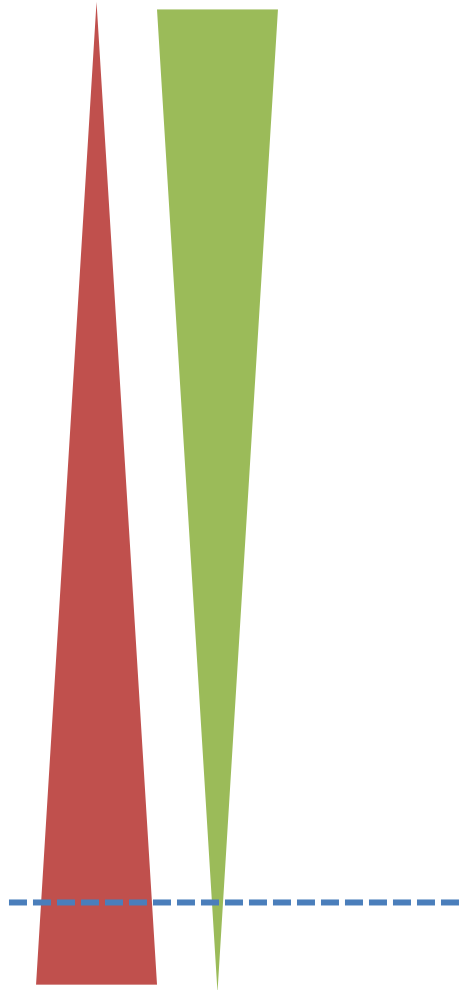
Item in PQRST Index	Example	Operationalization
		Data Source
P (productivity)	Number of publications in the top tier % of citations for the scientific field and year	ISI Essential Science Indicators (automated)
	Proportion of funded proposals that have resulted in ≥1 published reports of the main results	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

jama.com

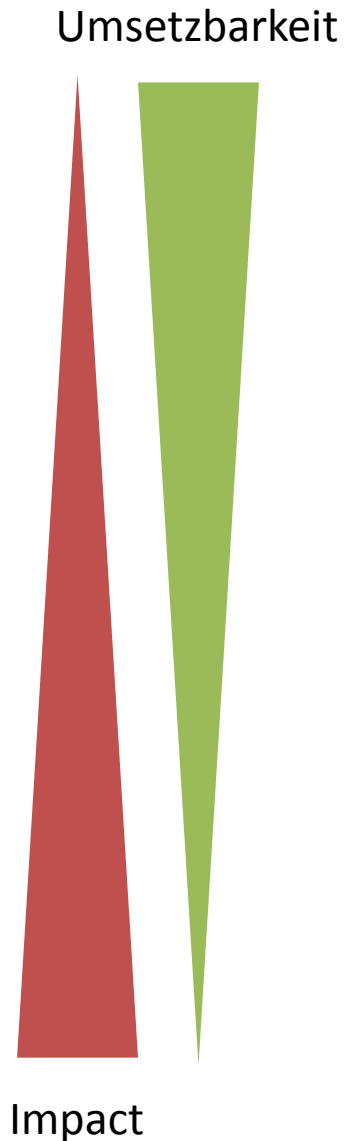
Ioannidis, Khoudri JAMA August 6, 2014 Volume 312, Number 5 483

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

Impact



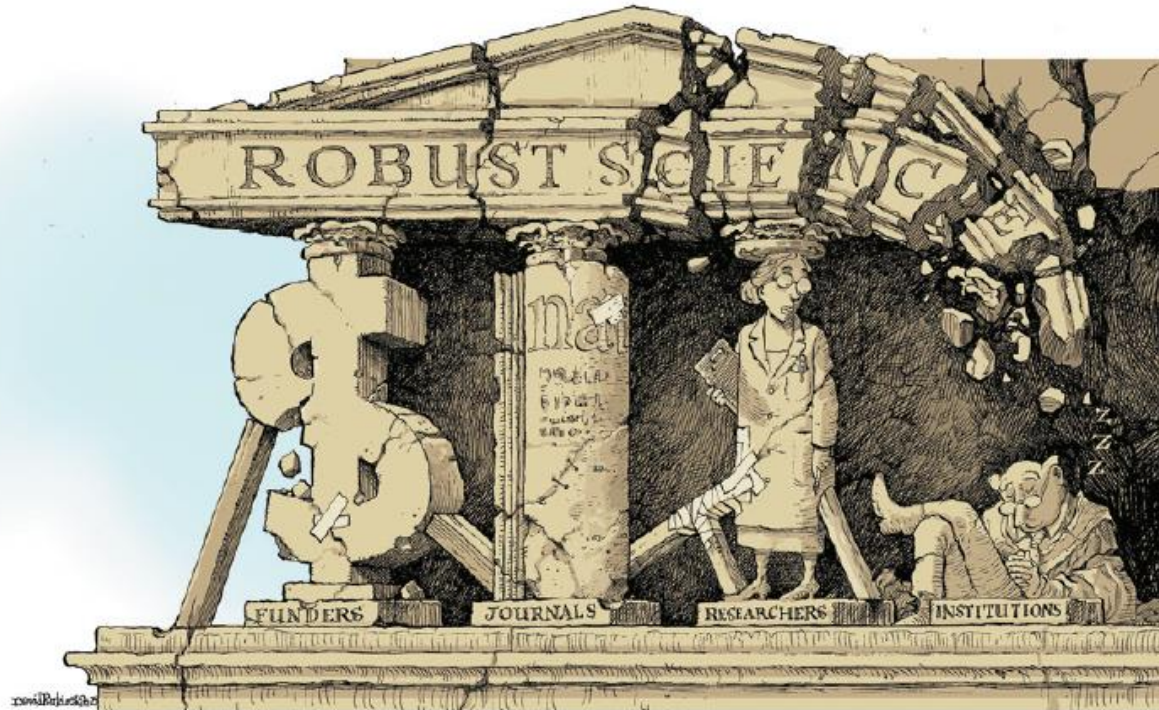
Summary III: How can we solve the problem?



- 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.