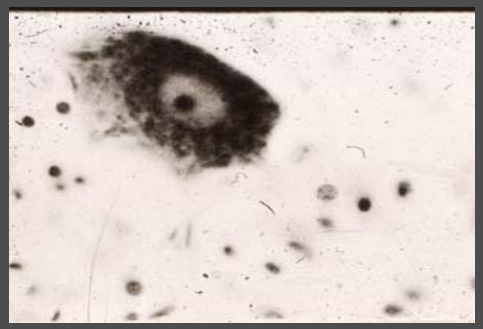
Improving the quality of preclinical research through more rigorous study design and transparent reporting

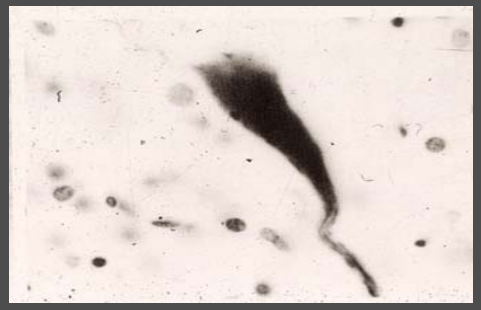
Story Landis, PhD
Advisory Committee to the Director (ACD)
National Institutes of Health (NIH)
June 13-14, 2013

Amyotrophic lateral sclerosis

Death within 5 years of diagnosis
Central pathological finding is motor neuron death

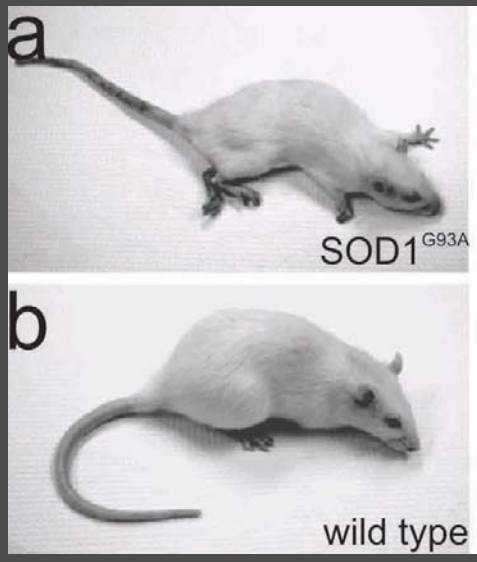
Normal

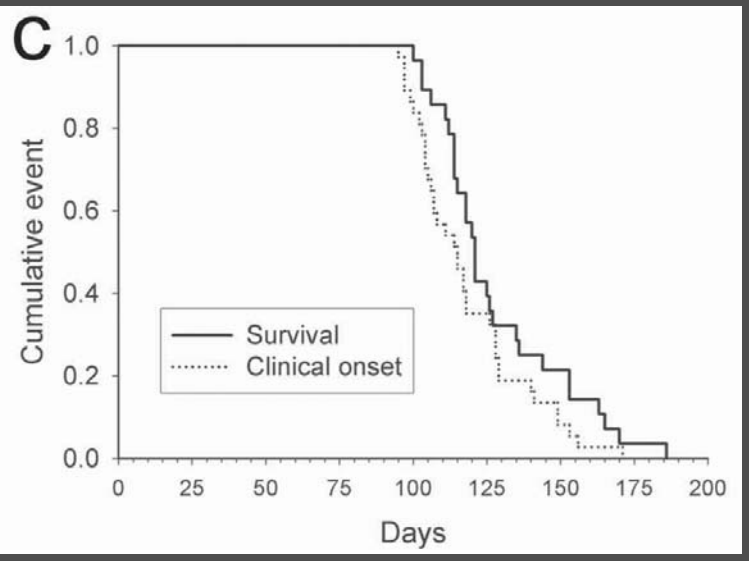




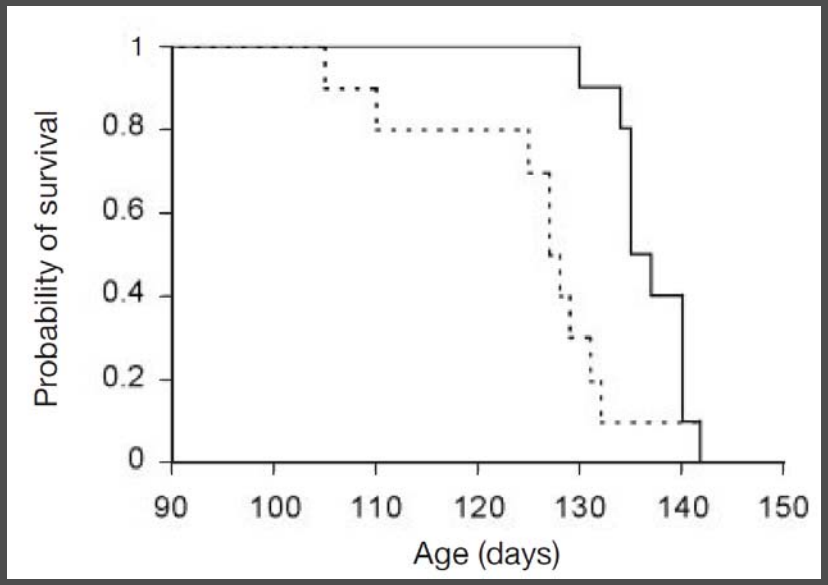
ALS

3% of cases from gain of function mutations in SOD1
Rodents over-expressing mSOD1 recapitulate ALS





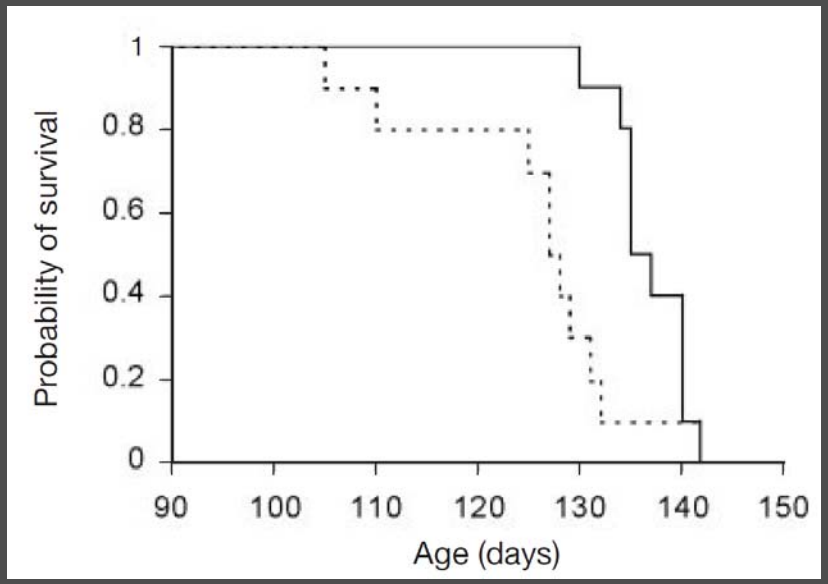
Enhanced survival of SOD1 transgenic mice with minocycline led to a Phase III clinical trial for ALS patients



- $SOD1^{G93A}$ transgenic mice
- Treatment started at 5 weeks of age
- i.p. 10mg/kg/day
- Nature 2002

Trial initiated 2003, completed 2007.
412 patients treated for 9 months
Randomized placebo controlled
Patients treated with minocycline failed more rapidly than those on placebo

Enhanced survival of SOD1 transgenic mice with minocycline led to a Phase III clinical trial for ALS patients



Trial initiated 2003, completed 2007.
412 patients treated for 9 months
Randomized placebo controlled
Patients treated with minocycline failed more rapidly than those on placebo

- $SOD1^{G93A}$ transgenic mice
- Treatment started at 5 weeks of age
- i.p. 10mg/kg/day
- **10** animals / group
- Not randomized
- Not blinded

Could the enhanced survival benefit have been due to small sample size and/or bias?

Design, power, and interpretation of studies in the standard murine model of ALS

SEAN SCOTT[1], JANICE E. KRANZ[1], JEFF COLE[1], JOHN M. LINCECUM[1], KENNETH THOMPSON[1], NANCY KELLY[1], ALAN BOSTROM[2], JILL THEODOSS[1], BASHAR M. AL-NAKHALA[1], FERNANDO G. VIEIRA[1], JEYANTHI RAMASUBBU[1] & JAMES A. HEYWOOD[1]

ALS Therapy Development Institute (ALS TDI)

"In the past five years we have screened more than 70 drugs in 18000 mice across 221 studies, using rigorous and appropriate statistical methodologies. While we were able to measure a significant difference in survival between males and females with great sensitivity, we observed no statistically significant positive (or negative) effects for any of the 70 compounds tested, including several previously reported as efficacious. "

"….We retested several compounds reported in major animal studies (minocycline, creatine, celecoxib, sodium phenylbutyrate, ceftriaxone, WHI-P131, thalidomide, and riluzole) …and found no survival benefit in the SOD1(G93A) mouse for any compounds (including riluzole) administered by their previously reported routes and doses. ……the majority of published effects are most likely measurements of noise in the distribution of survival means as opposed to actual drug effect."

Amyotroph Lateral Scler 2008; 9: 4-15

Beware the creeping cracks of bias

Evidence is mounting that research is riddled with systematic errors. Left unchecked, this could erode public trust, warns **Daniel Sarewitz**.

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

Florian Prinz, Thomas Schlange and Khusru Asadullah

Statistical Design Considerations in Animal Studies Published Recently in *Cancer Research*

Kenneth R. Hess

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.

Why animal research needs to improve

Many of the studies that use animals to model human diseases are too small and too prone to bias to be trusted, says **Malcolm Macleod**.

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Helping editors, peer reviewers and authors improve the clarity, completeness and transparency of reporting health research

David Moher*[1,2], Iveta Simera[3], Kenneth F Schulz[4], John Hoey[5] and Douglas G Altman[3]

Reforming Science: Methodological and Cultural Reforms

Drug targets slip-sliding away

The starting point for many drug discovery programs is a published report on a new drug target. Assessing the reliability of such papers requires a nuanced view of the process of scientific discovery and publication.

Translating animal research into clinical benefit

Poor methodological standards in animal studies mean that positive results may not translate to the clinical domain

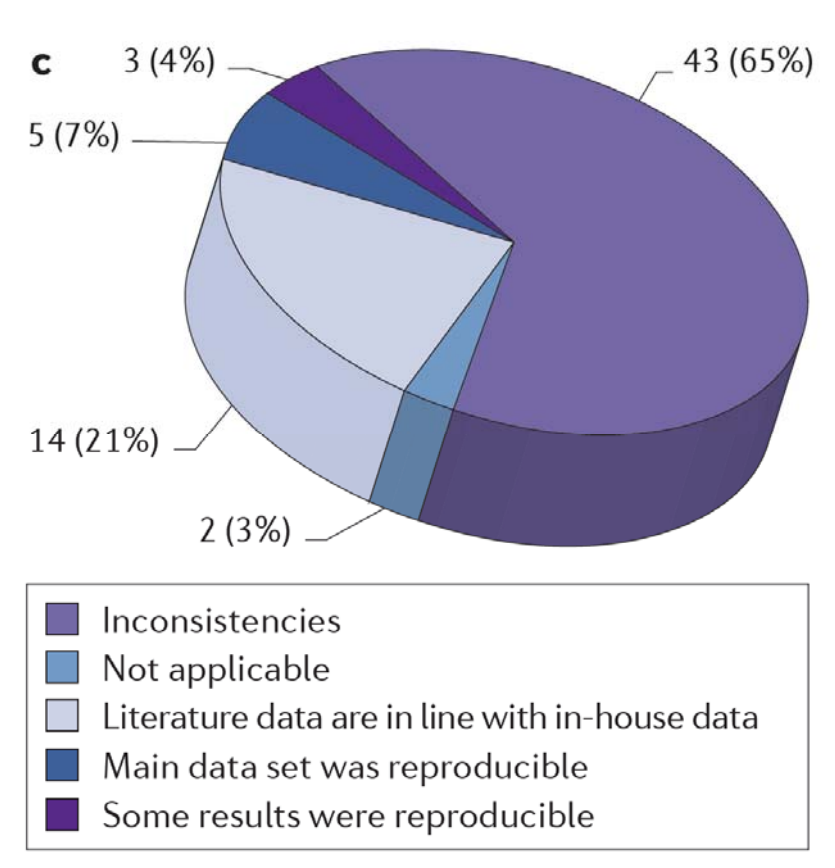
Almost 2/3 of 67 in-house projects could not replicate data published by others

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

Prinz, Schlange and Asadullah

Bayer HealthCare

Nature Reviews Drug Discovery
2011; 10:712-713



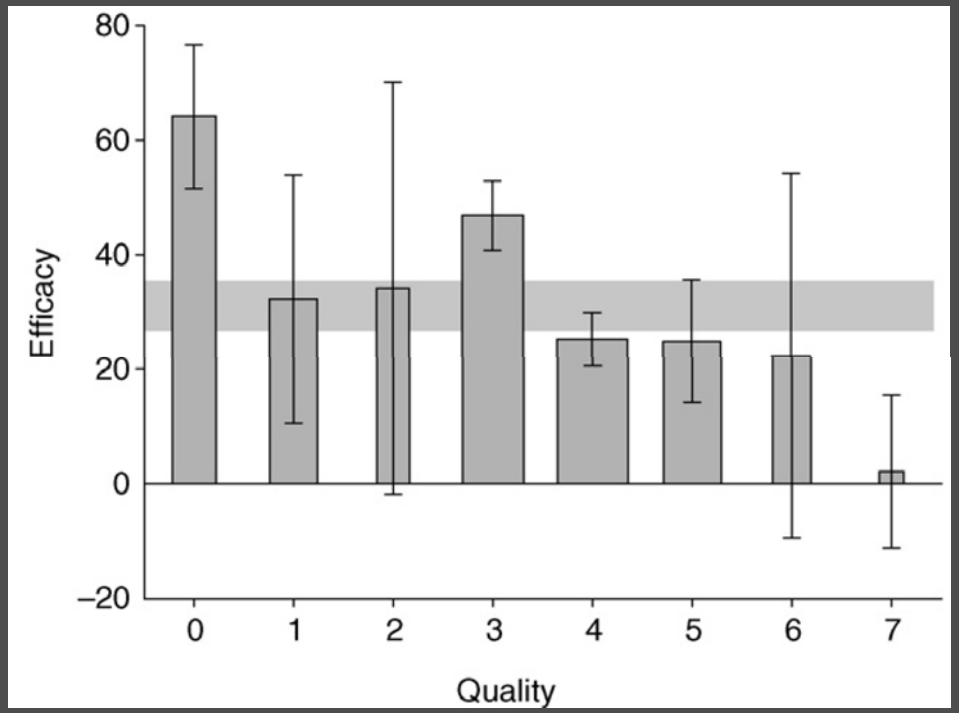
Lack of transparent reporting of methodology is evident for pre-clinical studies

Table 3. Prevelence of selected quality characteristics in other experimental models

	Number of publications	Randomisation (%)	Blinded assessment of outcome (%)	Sample-size calculation (%)
Transgenic stroke studies	157	n/a	3	0
Stroke pathophysiology studies	166	5	18	0
Parkinson's disease	118	12	15	0
Multiple sclerosis	183	2	11	0

Trends Neurosci 2007; 30: 433-439

The fewer methodological parameters are reported, the greater the apparent efficacy!



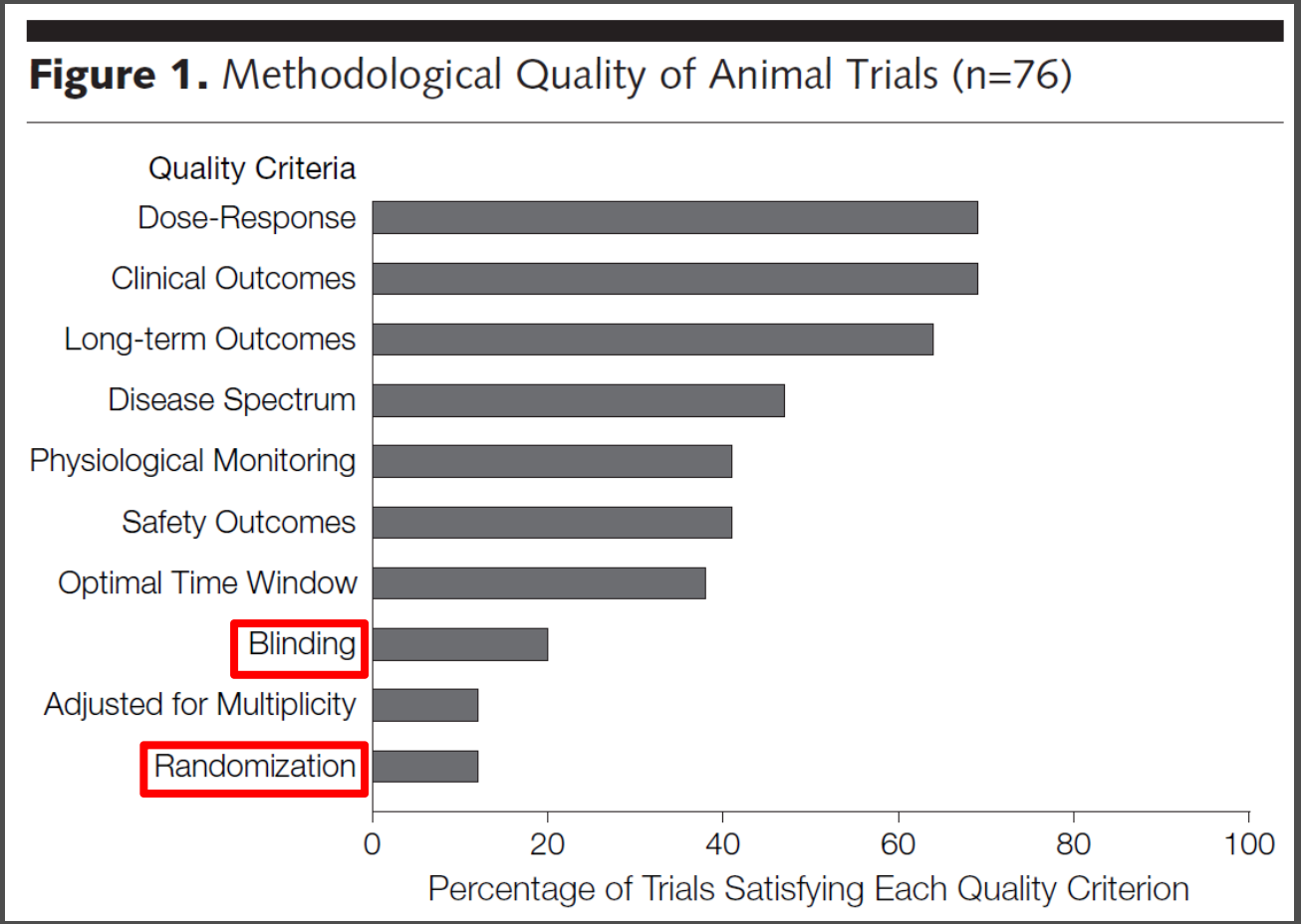
Effect size for studies of **FK506** (Tacrolimus) in experimental stroke.

Sena *et al., Trends Neurosci* 2007; 30: 433-439

Inadequate reporting is widespread

Journals:
- Cell
- Nature
- Science
- Nature Medicine
- Nature Genetics
- Nature Immunology
- Nature Biotechnology

>500 citations



Hackam and Redelmeier, *JAMA* 2006; 14: 1731-1732

Investigators need to be incentivized to improve reporting

Publish or perish

Grant support

Impact factor

Innovation

Significance

Novelty

Actions taken by NINDS: Notice in the Guide

Improving the Quality of NINDS-Supported Preclinical and Clinical Research through Rigorous Study Design and Transparent Reporting

Notice Number: NOT-NS-11-023
Release Date: August 10, 2011
Issued by: National Institute of Neurological Disorders and Stroke (NINDS)

Purpose:

> …..NINDS believes that applications that propose preclinical research, or that are based on previous preclinical data, will be greatly strengthened if the design, execution, and interpretation of the proposed studies and supporting data are adequately described. NINDS encourages investigators, whenever possible, to address these elements directly in their applications.

Inclusion of specific language on rigor and transparency in NINDS solicitations and attention to rigor in NINDS study sections

Guidance on the NINDS website

Experimental design:

- Rationale for the selected models and endpoints (animal and/or cellular)
- Adequacy of the controls
- Route & timing of intervention delivery / dosing
- Justification of sample size, including power calculation
- Statistical methods used in analysis and interpretation of results

Minimizing bias:

- Methods of blinding (allocation concealment and blinded assessment of outcome)
- Strategies for randomization and/or stratification
- Reporting of data missing due to attrition or exclusion
- Reporting of all results (negative and positive)

Results:

- Independent validation/replication, if available
- Robustness and reproducibility of the observed results
- Dose-response results
- Verification that interventional drug or biologic reached and engaged the target

Interpretation of results:

- Alternative interpretations of the experimental data
- Relevant literature in support or in disagreement with the results
- Discussion of effect size in relation to potential clinical impact
- Potential conflicts of interest

http://www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf

Actions taken by NINDS:
Workshop
"Optimizing the Predictive Value of Preclinical Research"

- ✓ *Guidance crafters*
- ✓ *Journal editors*
- ✓ *Reviewers*
- ✓ *End users*

A call for transparent reporting to optimize the predictive value of preclinical research

Story C. Landis[1], Susan G. Amara[2], Khusru Asadullah[3], Chris P. Austin[4], Robi Blumenstein[5], Eileen W. Bradley[6], Ronald G. Crystal[7], Robert B. Darnell[8], Robert J. Ferrante[9], Howard Fillit[10], Robert Finkelstein[1], Marc Fisher[11], Howard E. Gendelman[12], Robert Golub[13], John L. Goudreau[14], Robert A. Gross[15], Amelie K. Gubitz[1], Sharon E. Hesterlee[16], David W. Howells[17], John Huguenard[18], Katrina Kelner[19], Walter Koroshetz[1], Dimitri Krainc[20], Stanley E. Lazic[21], Michael S. Levine[22], Malcolm Macleod[23], John M. McCall[24], Richard T. Moxley III[25], Kalyani Narasimhan[26], Linda J. Noble[27], Steve Perrin[28], John D. Porter[1], Oswald Steward[29], Ellis Unger[30], Ursula Utz[1] & Shai D. Silberberg[1]

Nature 2012; 490: 187-191

Workshop and Publication Recommendations

- All relevant stakeholders share the responsibility of bringing about meaningful improvement in the quality of reporting.

- Grant applications and scientific publications which include *in vivo* animal experiments should, at a minimum, report on:
  - Randomization
  - Blinding
  - Sample size estimation
  - Handling of all data

- Clear guidance (e.g. checklist) to submitters and reviewers

- Education and training

Transparent reporting of all animal projects will permit more accurate assessment of their results

An important note about exploratory experiments

For the most part, these best practices do not apply to early stage observational experiments searching for possible differences among experimental groups. Such exploratory testing is frequently conducted using a small sample size, does not have a primary outcome, and is often unblinded. However, because such experiments are likely to be subject to many of the limitations described above, they should be viewed as hypothesis-generating experiments and interpreted as such. Potential discoveries arising from the exploratory phase of the research should be supported by follow-up, hypothesis-testing experiments that take into consideration and adequately report on the core standards detailed above (Box 1).

ANNOUNCEMENT

Reducing our irreproducibility

"To ease the interpretation and improve the reliability of published results we will more systematically ensure that key methodological details are reported, and we will give more space to methods sections. We will examine statistics more closely and encourage authors to be transparent, for example by including their raw data."

nature structural & molecular biology

Raising standards

nature cell biology

Raising reporting standards

nature neuroscience

Raising standards

nature immunology

Raising standards

EDITORIAL

NATURE MEDICINE

Raising standards

NICHD PAR-13-195
Preclinical Research on Model Organisms to Predict Treatment Outcomes for Disorders Associated with Intellectual and Developmental Disabilities (R01)

"All projects must adhere to a core set of reporting standards for rigorous study design. The standards are described fully in www.nature.com/nature/journal/v490/n7419/full/nature11556.html"

Department of Defense

Congressionally Directed Medical Research Programs

Duchenne Muscular Dystrophy Research Program

Investigator-Initiated Research Award

Funding Opportunity Number: W81XWH-13-DMDRP-IIRA

Catalog of Federal Domestic Assistance Number: 12.420

All projects should adhere to a core set of reporting standards for rigorous study design. The standards are described fully in www.nature.com/nature/journal/v490/n7419/full/nature11556.html. While these standards are written for preclinical studies, the basic principles of randomization, blinding, sample-size estimation, and data handling derive from well-established best practices in clinical studies and should be applied to those projects as well.

NCI-FDA-NIST Workshop on Standards in Molecular Diagnostics for the Discovery and Validation of Clinically Useful Cancer Biomarkers:

Recommendations from the
National Cancer Institute
U.S. Food and Drug Administration
National Institute of Standards and Technology

Summary of Workshop Recommendations

I. Biomarker Discovery and Development – What Labs Should Do

1. **Enhance analytic accuracy and precision of biomarkers by following CLIA/CAP guidelines**
2. **Enhance the culture of laboratories to improve consistency**
3. **Improve trial design elements needed to test potential biomarkers**
4. **Clearly define the steps from discovery to validation**

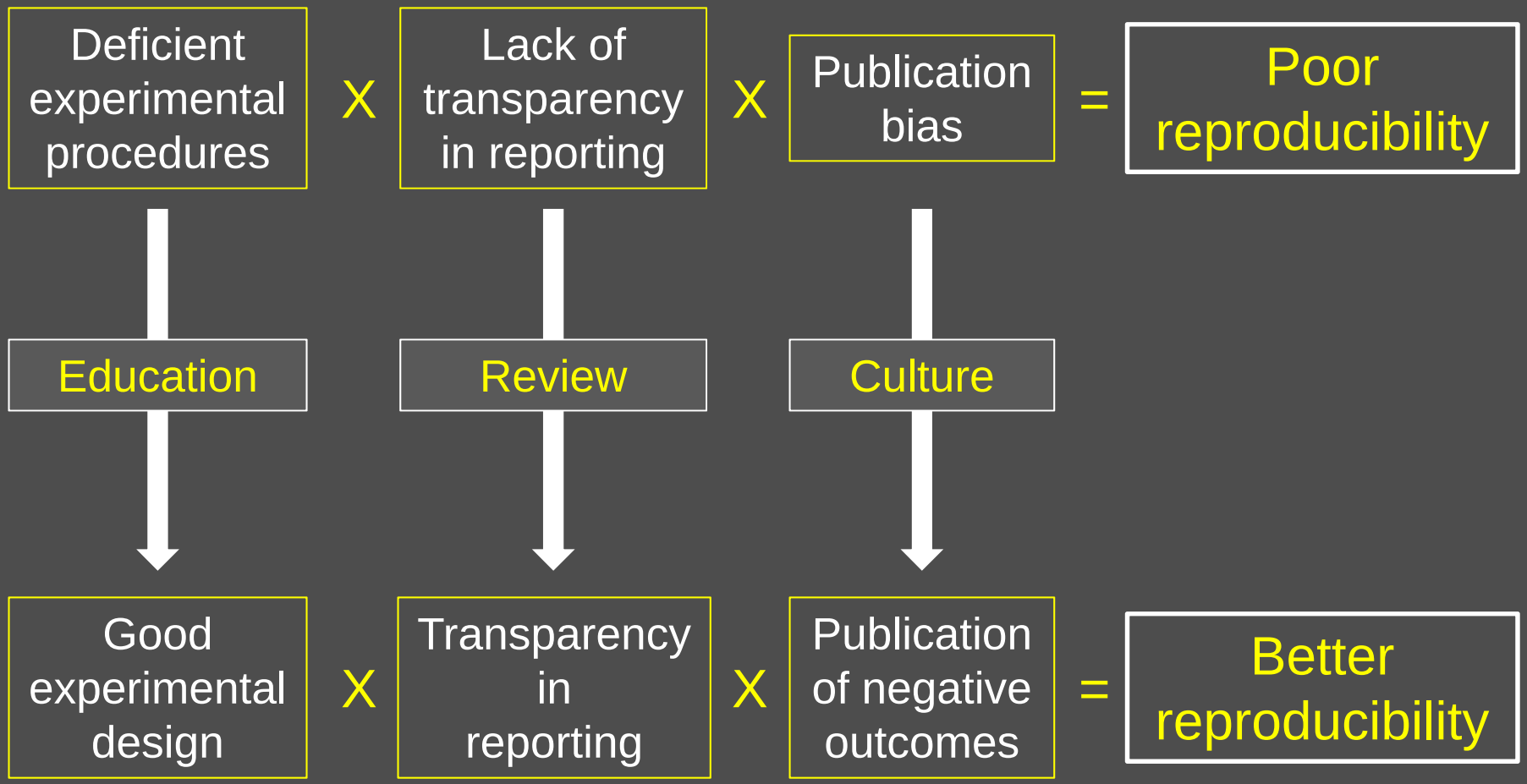
II. Biomarker Evaluation and Approval – What the Agencies Should Do

1. **Develop strong guidelines for biomarker development**
2. **Streamline the biomarker approval process**
3. **Require or encourage greater communication among the Federal agencies**
4. **Develop uniform global laboratory standards**

REporting recommendations for tumor MARKer prognostic studies (REMARK)

Lisa M McShane*, Douglas G Altman, Willi Sauerbrei, Sheila E Taube, Massimo Gion and Gary M Clark
for the Statistics Subcommittee of the NCI–EORTC Working Group on Cancer Diagnostics

- Marker examined, study objectives, prespecified hypotheses
- Characteristics of the study patients, treatments received and how chosen…
- Type of biological material…
- Assay method used and a detailed protocol…
- Method of case selection...
- All clinical endpoints examined
- All candidate variables initially examined or considered for inclusion
- Rationale for sample size…
- All statistical methods…
- How marker values were handled in the analyses
- The flow of patients through the study
- Distributions of basic demographic characteristics
- The relation of the marker to standard prognostic variables
- Univariable analyses of the relation between the marker and outcome
- For key multivariable analyses, report estimated effects
- Estimated effects with confidence intervals…
- Results of further investigations…
- Interpret the results in the context of the pre-specified hypotheses….
- Discuss implications for future research and clinical value.


Deficient experimental procedures
X
Lack of transparency in reporting
X
Publication bias
=
Poor reproducibility
Education
Review
Culture
Good experimental design
X
Transparency in reporting
X
Publication of negative outcomes
=
Better reproducibility

Potential approaches to address lack of reproducibility and transparency of published research findings

1. Raise community awareness.
2. Enhance formal training.
3. Improve the evaluation of scientists and their applications.
4. Increase stability for investigators.
5. Protect the integrity of science by adoption of more systematic review processes.

2. Enhance formal training

- Module on basic training on research integrity in the required Ethics training course for all trainees.
  - Would address research integrity as it relates to experimental biases, and proper study design.
- Incorporation of Experimental Design courses into training awards.
- Similar course materials from currently funded training programs and or universities distributed broadly via the web.