The importance of study design: statistical approach

Maia Lesosky
Division of Epidemiology & Biostatistics
School of Public Health & Family Medicine
University of Cape Town

Ethical animal research

- **Scientific validity**

- **Three ‘R’s: Replace, Refine, Reduce**
- Replace: New models, mathematical/statistical models (?) - more could be done here
- Refine: minimize harm, use trained personnel, better housing etc – all of these things are improving and improved
- Reduce: the topic of my talk (mainly)

A necessary condition for transferability is sound science

Is animal research sufficiently evidence based to be a cornerstone of biomedical research?

Public acceptance of the use of animals in biomedical research is conditional on it producing benefits for humans. **Pandora Pound** and **Michael Bracken** argue that the benefits remain unproved and may divert funds from research that is more relevant to doctors and their patients

Pandora Pound *medical sociologist*[1], Michael B Bracken *Susan Dwight Bliss professor of epidemiology*[2]

[1]Bath, UK; [2]Yale University Schools of Public Health and Medicine, New Haven CT, USA

Research in Translation

Can Animal Models of Disease Reliably Inform Human Studies?

H. Bart van der Worp[1]*, David W. Howells[2], Emily S. Sena[2,3], Michelle J. Porritt[2], Sarah Rewell[2], Victoria O'Collins[2], Malcolm R. Macleod[3]

1 Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Centre Utrecht, Utrecht, The Netherlands, **2** National Stroke Research Institute & University of Melbourne Department of Medicine, Austin Health, Melbourne, Australia, **3** Department of Clinical Neurosciences, University of Edinburgh, Edinburgh, United Kingdom

Can you trust your animal study data?

Ian S. Peers, Marie C. South, Peter R. Ceuppens, Jonathan D. Bright and Elizabeth Pilling

Scientific **validity:**

- Validity is a hard construct to measure.
- Research may be ‘internally’ valid, but externally invalid.

Internal validity is controllable and mainly a reflection on study design and analytic principles

Scientifically unsound research is unethical by definition.

Sound study design begins with the research question

- Pilot studies: “Does this exist/happen/possible?”
- Exploratory studies: “What happens when….?”
- Confirmatory studies: “Is A better then B?”

Each of these *should* have different approaches to design, power, sample size and analysis.

Sound study design

A design appropriate for the research question

Enough numbers (n) to resolve the hypothesis without ambiguity.

Reduction/removal of known sources of bias:

- Randomisation
- Blinding
- Intention-to-treat analysis
- Publishing all (not just significant) results
- Pre-registration of protocols & analysis plans

The vast majority of studies (even human)…

- Are underpowered
- Do not replicate

Why Most Clinical Research Is Not Useful

John P. A. Ioannidis

Published: June 21, 2016 • https://doi.org/10.1371/journal.pmed.1002049

Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature

Denes Szucs, John P. A. Ioannidis

Published: March 2, 2017 • https://doi.org/10.1371/journal.pbio.2000797

OPINION ARTICLE

Front. Psychol., 13 July 2017 | https://doi.org/10.3389/fpsyg.2017.01184

Targeting Next Generations to Change the Common Practice of Underpowered Research

Rik Crutzen[1]* and Gjalt-Jorn Y. Peters[2,3]

[1]Department of Health Promotion, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, Netherlands
[2]Faculty of Psychology and Education Science, Open University of the Netherlands, Heerlen, Netherlands
[3]Department of Work and Social Psychology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands

Underpowered studies remain ubiquitous (Maxwell, 2004; Bakker et al., 2012; Button et al., 2013; Turner et al., 2013; Szucs and Ioannidis, 2017) despite strong pleas to change this practice (Cohen, 1988, 1990, 1992). As with any complex problem, multiple factors contribute to the ubiquity of conducting underpowered studies, and a wide range of efforts is

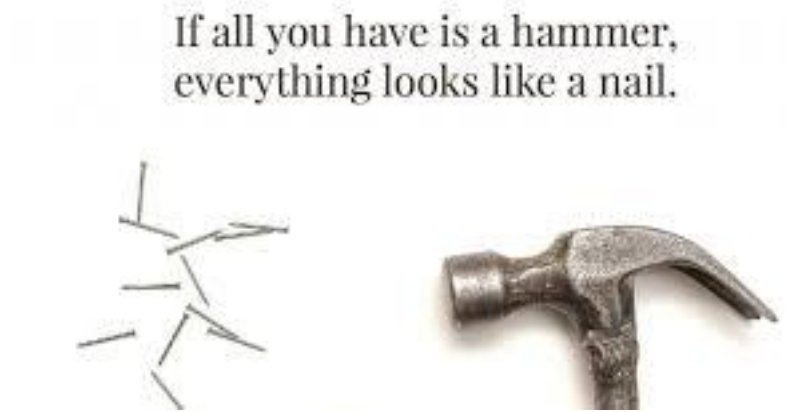
Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions

Andrew D. Higginson, Marcus R. Munafò

Published: November 10, 2016 • https://doi.org/10.1371/journal.pbio.2000995

- More statistical attention (training, material) is paid to confirmatory studies (both in design and in analysis) then to other types of studies

- Sometimes this results in a mismatch between researcher needs and researcher knowledge



- Plus: study design has moved on..

Modern study design is far more complex than most clinical/pre-clinical researchers have had exposure to.

Eg. adaptive and group sequential designs, Bayesian frameworks for analysis

Guidelines for the Design and Statistical Analysis of Experiments Using Laboratory Animals

Michael F. W. Festing a

Int J Mol Sci. 2015 Oct; 16(10): 24048–24058.
Published online 2015 Oct 12. doi: 10.3390/ijms161024048
PMCID: PMC4632737

The Potential of Adaptive Design in Animal Studies

Arshad Majid,[1,*] Ok-Nam Bae,[2] Jessica Redgrave,[1] Dawn Teare,[3] Ali Ali,[1] and Daniel Zemke[1]

Katalin Prokai-Tatrai, Academic Editor

Increasing efficiency of preclinical research by group sequential designs

Konrad Neumann, Ulrike Grittner, Sophie K. Piper, Andre Rex, Oscar Florez-Vargas, George Karystianis, Alice Schneider, Ian Wellwood, Bob Siegerink, John P. A. Ioannidis, Jonathan Kimmelman, Ulrich Dirnagl

Published: March 10, 2017 • https://doi.org/10.1371/journal.pbio.2001307

Ethical benefits to adaptive design

- Adaptive dose finding decreases the number of subjects exposed to ineffective or toxic doses and allows a faster transition to safe and effective doses.
- Dropping inferior treatment groups allows subjects to be reassigned to ones that are more successful.
- Adaptive treatment switching, biomarker adaptive strategies, and target population enrichment allow subjects to receive better, more individualized care than by random group assignment.
- Adaptive design allows the required number of animals to be reduced if a significant effect is detected early or potentially painful treatments to be dropped if no effect is seen.

Int. J. Mol. Sci. 2015, 16, 24048-24058; doi:10.3390/ijms161024048

Unrealistic choice of effect sizes

- “For animal studies, effects of realistic treatment doses might be small, and therefore appropriately powered studies will have to be large. To increase the generalisability of findings, investigators should plan for heterogeneity in the circumstances of testing. For these large studies to be feasible, consideration should be given to the development of multicentre animal studies.”
- Creating largely homogeneous experiments aids reproducibility and boosts statistical power, but has a cost of generalizability: the few drugs that have translated successfully from animals are effective across a broad range of circumstances (see, for example, E. S. Sena *et al*. *J. Cereb. Blood Flow Metab.* **30**, 1905–1913; 2010).


Research: increasing value, reducing waste 2

Increasing value and reducing waste in research design, conduct, and analysis

John P A Ioannidis, Sander Greenland, Mark A Hlatky, Muin J Khoury, Malcolm R Macleod, David Moher, Kenneth F Schulz, Robert Tibshirani

Lancet 2014; 383: 166–75 Correctable weaknesses in the design, conduct, and analysis of biomedical and public health research studies can

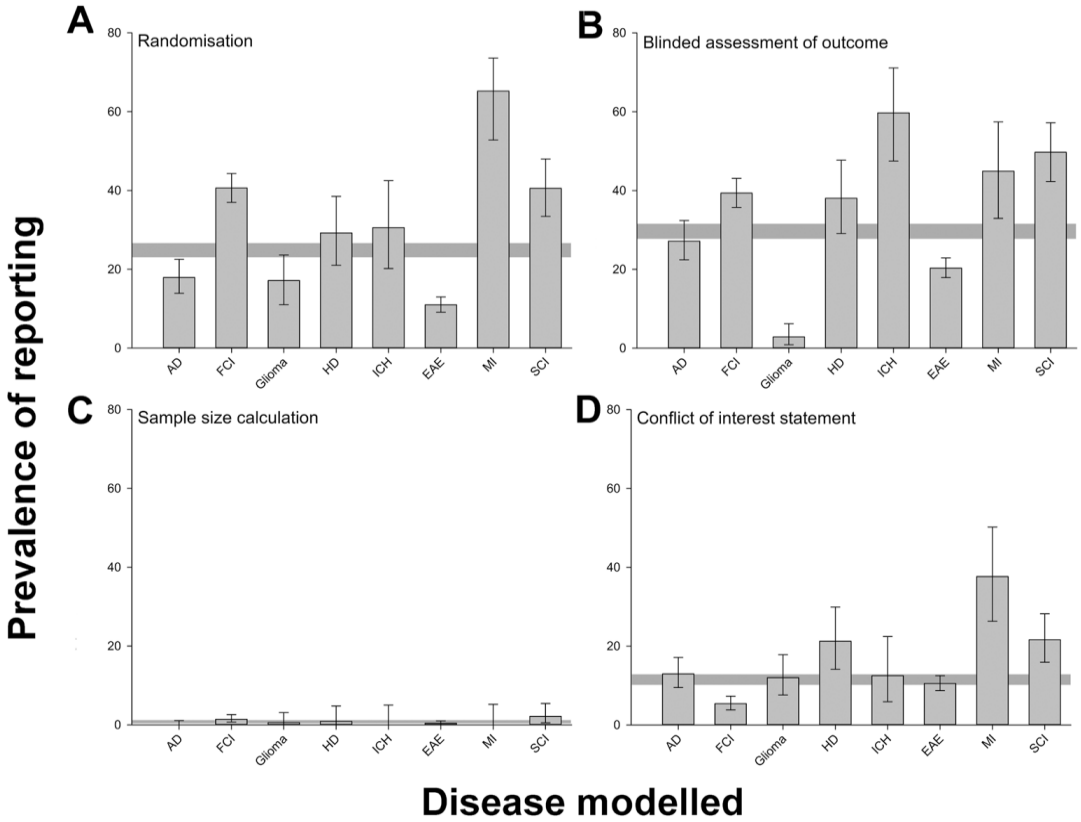
Just…issues...

- This is the largest and most comprehensive survey of this kind carried out to date. We provide evidence that many peer- reviewed, animal research publications fail to report important information regarding experimental and statistical methods.
    - Problems with the transparency and robustness of the statistical analysis in 60%
    - Randomisation reported in only 12%
    - 40% used a less efficient study design then was possible

Survey of the Quality of Experimental Design, Statistical Analysis and Reporting of Research Using Animals

Carol Kilkenny[1]*, Nick Parsons[2], Ed Kadyszewski[3], Michael F. W. Festing[4], Innes C. Cuthill[5], Derek Fry[6], Jane Hutton[7], Douglas G. Altman[8]

Bias is endemic in animal studies



Risk of Bias in Reports of In Vivo Research: A Focus for Improvement Malcolm R. Macleod Aaron Lawson McLean Aikaterini Kyriakopoulou Stylianos Serghiou Arno de Wilde Nicki Sherratt Theo Hirst Rachel Hemblade Zsanett Bahor Cristina Nunes-Fonseca Aparna Potluru Andrew Thomson Julija Baginskitae Kieren Egan Hanna Vesterinen Gillian L. Currie Leonid Churilov David W. Howells Emily S. Sena

Bias reduction

- Humans are biased
- Our own bias is usually invisible to us
- This has been empirically demonstrated over and over
- Bias should be reduced where possible
- Human bias is best reduced by randomisation and blinding
- Typically, when studies are well blinded and concealed and randomised the estimated effect is lower than that of an unblinded equivalent (because our bias is invisible to ourselves)
- Studies should be registered and all results published

- Non-randomised trials had larger effect sizes.
- “Unduly biased animal studies should not be allowed to constitute part of the rationale for human trials.”
- Most animal studies were biased (only 29% reported any randomisation / concealment)

The Need for Randomization in Animal Trials: An Overview of Systematic Reviews

Jennifer A. Hirst[1*], Jeremy Howick[1*], Jeffrey K. Aronson[1], Nia Roberts[2], Rafael Perera[1], Constantinos Koshiaris, Carl Heneghan[1]

1 Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom, **2** Bodleian Libraries, University of Oxford, Oxford, United Kingdom

Insufficient training

- The way that many laboratory studies are reported suggests that scientists are unaware that their methodological approach is without rigour.

- Many laboratory scientists have insufficient training in statistical methods and study design.

- This issue might be a **more important deficiency** than is poor training in clinical researchers, especially for laboratory investigation done by one scientist in an isolated laboratory—by contrast, many people would examine a clinical study protocol and report.


Research: increasing value, reducing waste 2

Increasing value and reducing waste in research design, conduct, and analysis

John P A Ioannidis, Sander Greenland, Mark A Hlatky, Muin J Khoury, Malcolm R Macleod, David Moher, Kenneth F Schulz, Robert Tibshirani

Lancet 2014; 383: 166–75

Correctable weaknesses in the design, conduct, and analysis of biomedical and public health research studies can

Ethics committees

- Request (require?) better scientific practice
  - Protocol pre-registration
  - Publication of all results
  - Randomisation, blinding, outcome concealment
- Leverage role to motivate for better training opportunities for researchers
  - Statistical methods
  - Study design

***Panel 2*: Ten options to improve the quality of animal research**

Protocols and optimum design

1 Creation of a publicly accessible date-stamped protocol preceding data collection and analysis, or clear documentation that research was entirely exploratory
2 Use of realistic sample size calculations
3 Focus on relevance, not only statistical efficiency

Effect-to-bias ratio

4 Random assignment of groups
5 Incorporation of blind observers
6 Incorporation of heterogeneity into the design, whenever appropriate, to enhance generalisability
7 Increase in multicentre studies
8 Publishers should adopt and implement the ARRIVE (Animal Research: Reporting In Vivo Experiments) guidelines

Workforce and stakeholders

9 Programmes for continuing professional development for researchers

Reproducibility and reward systems

10 Funders should increase attention towards quality and enforce public availability of raw data and analyses