GOOD SCIENCE:

Design

Welfare

Education

Manuel Berdoy
MPhil DPhil PGDipLATHE

Course Director and Lead NTCO at Oxford University.
ICLAS Governing Board Member, and European Committee
FELASA Working Group on Exp. Design
LASA Council Member,
Co-chair of the LASA Education and Training Ethics committee

GOOD SCIENCE:

Design of experiments

Welfare

Education

Manuel Berdoy
MPhil DPhil PGDipLATHE

Course Director and Lead NTCO at Oxford University.
ICLAS Governing Board Member, and European Committee
FELASA Working Group on Exp. Design
LASA Council Member,
Co-chair of the LASA Education and Training Ethics committee

UNIVERSITY OF OXFORD

GOOD SCIENCE:

Design of experiments

Welfare of animals

Education

Manuel Berdoy
MPhil DPhil PGDipLATHE

Course Director and Lead NTCO at Oxford University.
ICLAS Governing Board Member, and European Committee
FELASA Working Group on Exp. Design
LASA Council Member,
Co-chair of the LASA Education and Training Ethics committee

UNIVERSITY OF OXFORD

GOOD SCIENCE:

Design of experiments

Welfare of animals

Education of researchers / animal users.

Manuel Berdoy
MPhil DPhil PGDipLATHE

Course Director and Lead NTCO at Oxford University.
ICLAS Governing Board Member, and European Committee
FELASA Working Group on Exp. Design
LASA Council Member,
Co-chair of the LASA Education and Training Ethics committee

UNIVERSITY OF OXFORD

GOOD SCIENCE: (Reproducible, Efficient and Ethical)

Design of experiments

Welfare of animals

Education of researchers / animal users.

Manuel Berdoy
MPhil DPhil PGDipLATHE

Course Director and Lead NTCO at Oxford University.
ICLAS Governing Board Member, and European Committee
FELASA Working Group on Exp. Design
LASA Council Member,
Co-chair of the LASA Education and Training Ethics committee

UNIVERSITY OF OXFORD

GOOD SCIENCE: (Reproducible, Efficient and Ethical)

Design of experiments We are not doing a great job !

Welfare of animals

Education of researchers / animal users.

UNIVERSITY OF OXFORD

GOOD SCIENCE: (Reproducible, Efficient and Ethical)

Design of experiments We are not doing a great job !

Welfare of animals Ethical aspects integral to good science

Education of researchers / animal users.

UNIVERSITY OF OXFORD

GOOD SCIENCE: (Reproducible, Efficient and Ethical)

Design of experiments

We are not doing a great job !

Welfare of animals

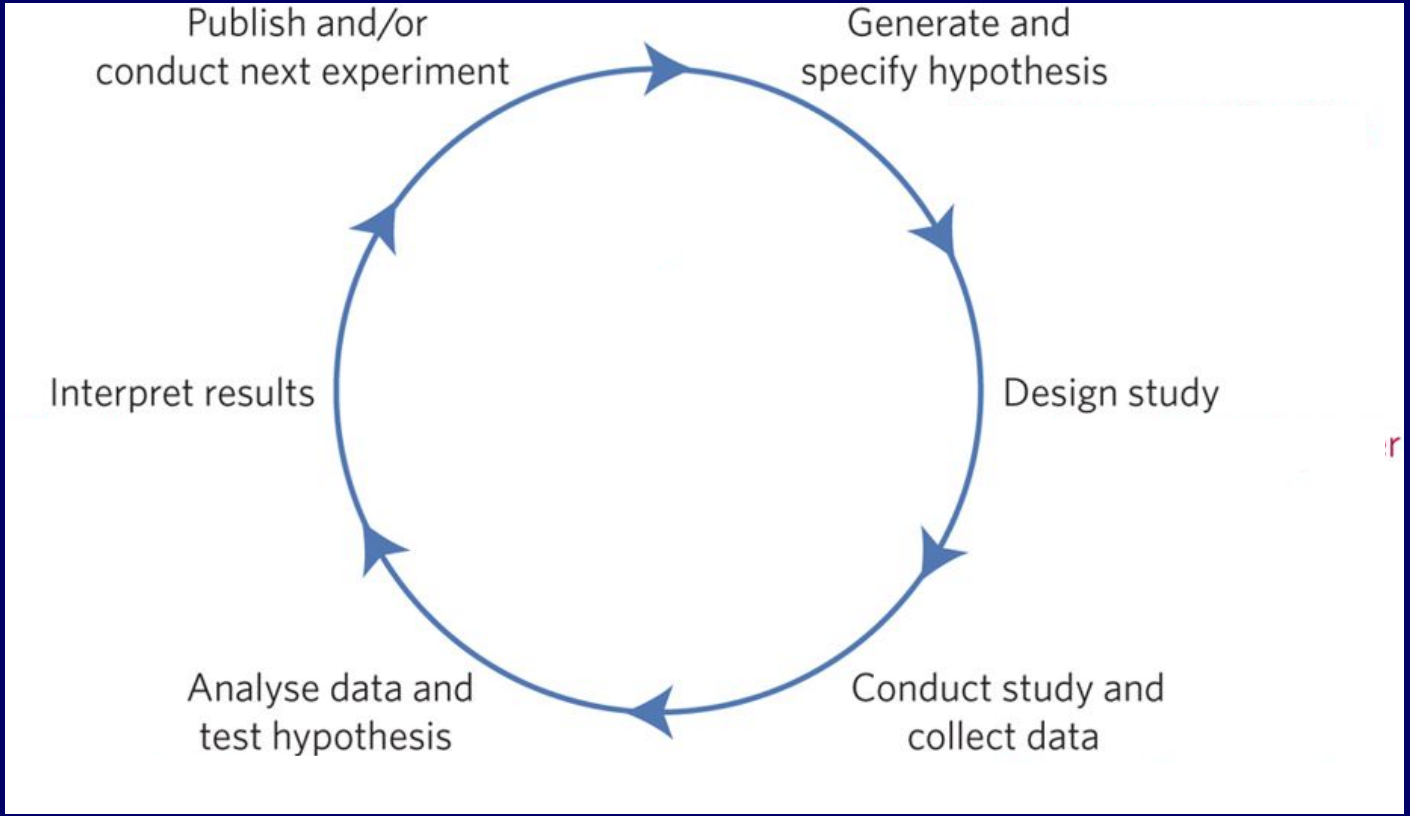
Ethical aspects integral to good science

Education of researchers / animal users.

Some solutions for the road ahead

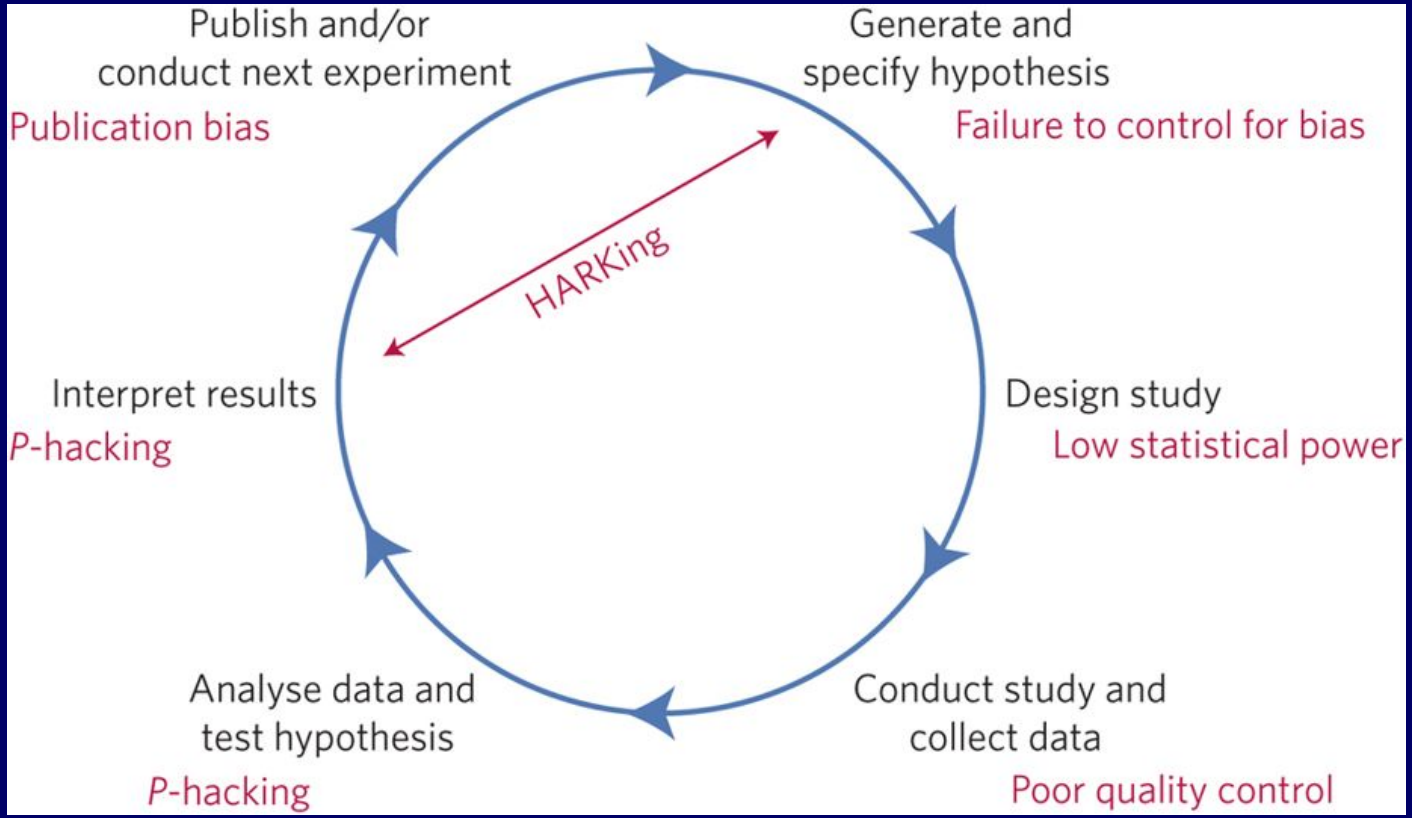
UNIVERSITY OF OXFORD

Searching for Truth = Experiment



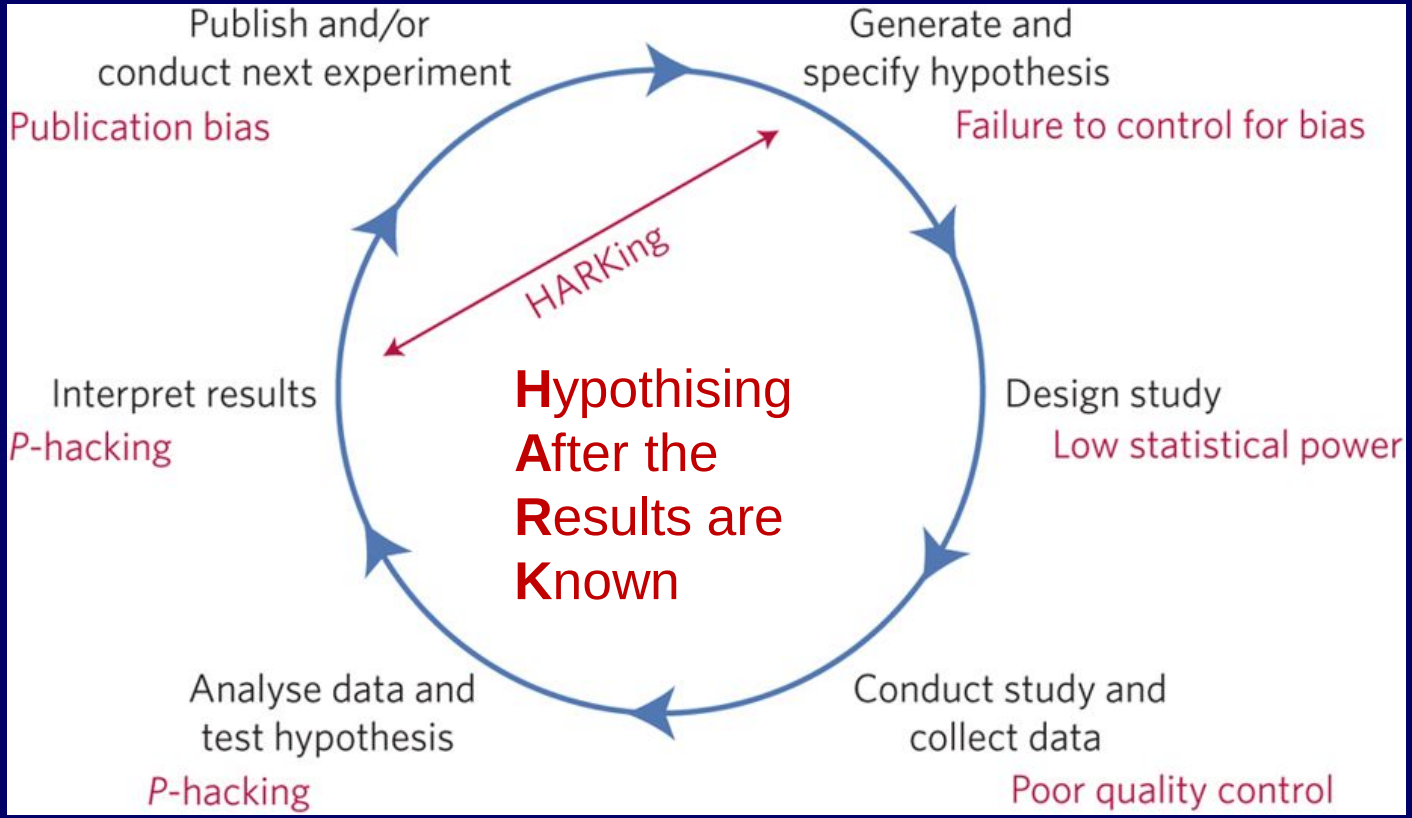
Munafo et al., Nature Human Behaviour, 2017, 1:0021

Searching for Truth = Experiment

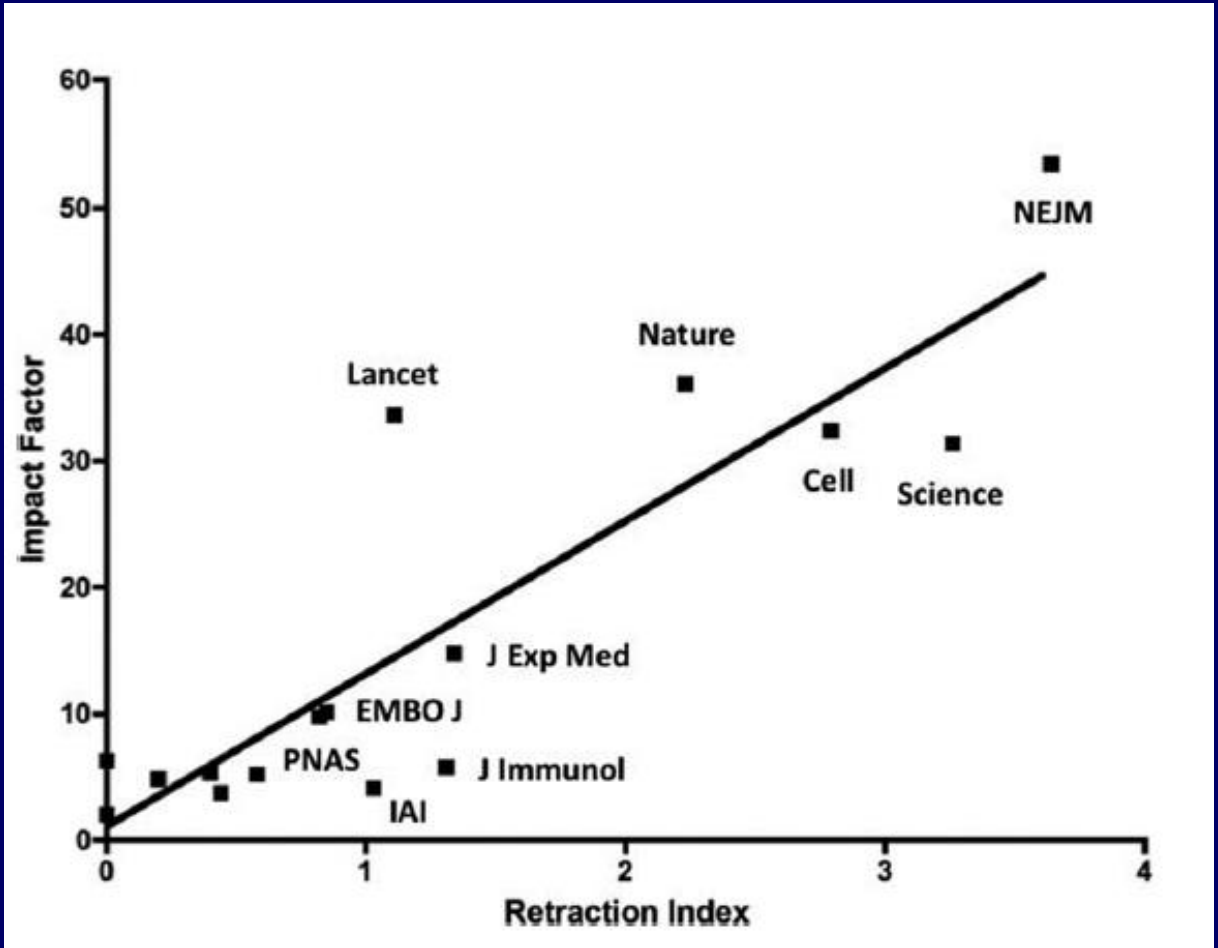


Munafo et al., Nature Human Behaviour, 2017, 1:0021

Searching for Truth = Experiment



Munafo et al., Nature Human Behaviour, 2017, 1:0021



Fang & Casadevall 2011

Retraction Watch:
Tracking retractions as a window into the scientific process
https://retractionwatch.com/

Reproducibility issues

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

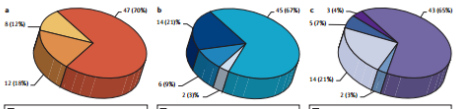
Florian Prinz, Thomas Schlange and Khusru Asadullah

A recent report by Arrowsmith noted that the success rates for new development projects in Phase II trials have fallen from 28% to 18% in recent years, with insufficient efficacy being the most frequent reason for failure (Phase II failures: 2008–2010. *Nature Rev. Drug Discov.* **10**, 328–329 (2011))[1]. This indicates the limitations of the predictivity of disease models and also that the validity of the targets being investigated is frequently questionable, which is a crucial issue to address if success rates in clinical trials are to be improved.

Candidate drug targets in industry are derived from various sources, including in-house target identification campaigns, in-licensing and public sourcing, in particular based on reports published in the literature and presented at conferences. During the transfer of projects from an academic to a company setting, the focus changes from 'interesting' to 'feasible/marketable', and the financial costs of pursuing a full-blown drug discovery and development programme for a particular target could ultimately be hundreds of millions of Euros. Even in the earlier stages, investments in activities such as high-throughput screening programmes are substantial, and thus the validity of published data on potential targets is crucial for companies when deciding to start novel projects.

To mitigate some of the risks of such investments ultimately being wasted, most pharmaceutical companies run in-house target validation programmes. However, validation projects that were started in our company based on exciting published data have often resulted in disillusionment when key data could not be reproduced. Talking to scientists, both in academia and in industry, there seems to be a general impression that many results that are published are hard to reproduce. However, there is an imbalance between this apparently widespread impression and its public recognition (for example, see REFS 2,3), and the surprisingly few scientific publications dealing with this topic. Indeed, to our knowledge, so far there has been no published in-depth, systematic analysis that compares reproduced research results with published results for wet-lab experiments related to target identification and validation.

Early research in the pharmaceutical industry, with a dedicated budget and scientists who mainly work on target validation to increase the confidence in a project, provides a unique opportunity to generate a broad data set on the reproducibility of published data. To substantiate our incidental observations that published reports are frequently not reproducible with quantitative data, we performed an analysis of our early (target identification and validation) in-house projects in our strategic research fields of oncology, women's health and cardiovascular diseases that were performed over the past 4 years (FIG. 1a). We distributed a questionnaire to all involved scientists from target discovery, and queried names, main relevant published data (including citations), in-house data obtained and their relationship to the published data, the impact of the results obtained for the outcome of the projects, and the models



d

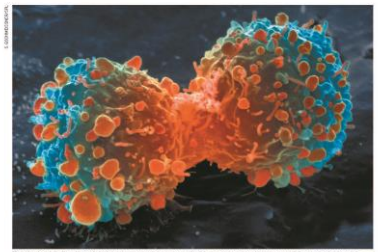
	Model reproduced 1:1	Model adapted to internal needs (cell line, assays)	Literature data transferred to another indication	Not applicable
In-house data in line with published results	1 (7%)	12 (86%)	0	1 (7%)
Inconsistencies that led to project termination	11 (16%)	26 (60%)	2 (5%)	4 (9%)

Figure 1 | Analysis of the reproducibility of published data in 67 in-house projects. a | This figure illustrates the distribution of projects within the oncology, women's health and cardiovascular indications that were analysed in this study. **b** | Several approaches were used to reproduce the published data. Models were either exactly copied, adapted to internal needs (for example, using other cell lines than those published, other assays and so on) or the published data was transferred to models for another indication. 'Not applicable' refers to projects in which general hypotheses could not be verified. **c** | Relationship of published data to in-house data. The proportion of each of the following outcomes is shown: data were completely in line with published data; the main set was reproducible; some results (including the most relevant hypothesis) were reproducible; or the data showed inconsistencies that led to project termination. 'Not applicable' refers to projects that were almost exclusively based on in-house data, such as gene expression analysis. The number of projects and the percentage of projects within this study (a–c) are indicated. **d** | A comparison of model usage and reproducible and irreproducible projects is shown. The respective numbers of projects and the percentages of the groups are indicated.

Findings confirmed in only 14 out of 67 studies (21%)

- Prinz F, Schlange T, Asadullah K (2011). *Nat Rev Drug Discov* **10**(9): 712.

COMMENT



Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

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.

Efforts over the past decade to characterize the genetic alterations in human cancers have led to a better understanding of molecular drivers of this complex set of diseases. Although we in the cancer field hoped that this would lead to more effective drugs, historically, our ability to translate cancer research to clinical success has been remarkably low[1]. Sadly, clinical trials in oncology have the highest failure rate compared with other therapeutic areas. Given the high unmet need in oncology, it is understandable that barriers to clinical development may be lower than for other disease areas, and a larger number of drugs with suboptimal preclinical validation will enter oncology trials. However, this low success rate is not sustainable or acceptable, and investigators must reassess their approach to translating discovery research into greater clinical success and impact.

Many factors are responsible for the high failure rate, notwithstanding the inherently difficult nature of this disease. Certainly, the limitations of preclinical tools such as inadequate cancer cell-line and mouse models[2] make it difficult for even ▶

Findings confirmed in only 6 out of 53 studies (11%)

Begley CG, Ellis LM (2012). Drug development: *Nature* **483**(7391): 531-533.

ANNOUNCEMENT

Reducing our irreproducibility

Over the past year, *Nature* has published a string of articles that highlight failures in the reliability and reproducibility of published research (collected and freely available at go.nature.com/huhbyr). The problems arise in laboratories, but journals such as this one compound them when they fail to exert sufficient scrutiny over the results that they publish, and when they do not publish enough information for other researchers to assess results properly.

From next month, *Nature* and the Nature research journals will introduce editorial measures to address the problem by improving the consistency and quality of reporting in life-sciences articles. 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.

Central to this initiative is a checklist intended to prompt authors to disclose technical and statistical information in their submissions, and to encourage referees to consider aspects important for research reproducibility (go.nature.com/oloeip). It was developed after discussions with researchers on the problems that lead to irreproducibility, including workshops organized last year by US National Institutes of Health (NIH) institutes. It also draws on published concerns about reporting standards (or the lack of them) and the collective experience of editors at Nature journals.

Reproducibility issues

CORRESPONDENCE

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

Florian Prinz, Thomas Schlange and Khusru Asadullah

Findings confirmed in
out of 67 studies (2

- Prinz F, Schlange T, Asadullah K (2011). *Nat Rev Drug Discov* **10**(9): 712.

NATURE|Vol 444|21/28 December 2006 NEWS

EDISON'S BULBS FAIL TO LIGHT UP AUCTION
First all-science collection sells modestly at Christie's.
www.nature.com/news

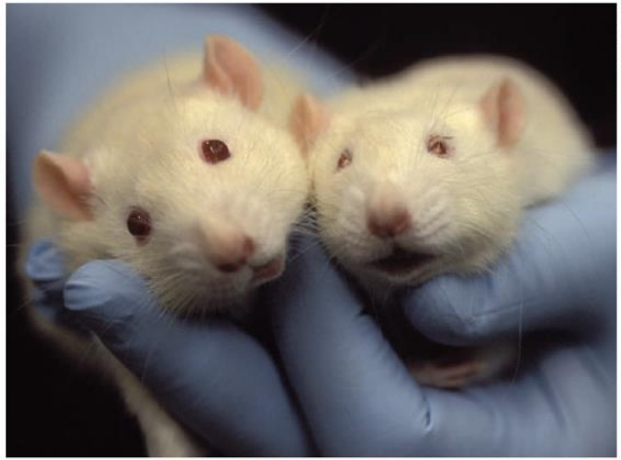
CHRISTIE'S

Animal experiments under fire for poor design

In the contentious world of animal research, one question surfaces time and again: how useful are animal experiments as a way to prepare for trials of medical treatments in humans? The issue is crucial, as public opinion is behind animal research only if it helps develop better drugs. Consequently, scientists defending animal experiments insist they are essential for safe clinical trials, whereas animal-rights activists vehemently maintain that they are useless.

Now a British team has made the first attempt to answer the question in a scientific way, and the result suggests that animal researchers need to raise their game. The team claims that animal experiments are often poorly designed, and so fail to lay the ground properly for subsequent human studies.

The study looked at six treatments that have been evaluated in detail in human trials. The researchers assessed whether animal studies had accurately predicted the outcome of the human work, a task that involved reviewing more than 200 papers. In three of the six cases, the answer was no (P. Perel *et al. Br. Med. J.* doi:10.1136/bmj.39048.407928.BE; 2006).



K. KASMAUSKI/CORBIS

Are animals being wasted in badly thought through experiments?

out of 53 studies (11%)

Begley CG, Ellis LM (2012). Drug development: *Nature* **483**(7391): 531-533.

VOL 496 | 25 APRIL 2013

ng our lucibility

ature has published a string of articles that the reliability and reproducibility of pub- ed and freely available at go.nature.com/ arise in laboratories, but journals such as when they fail to exert sufficient scrutiny y publish, and when they do not publish other researchers to assess results properly. *ture* and the Nature research journals will ures to address the problem by improving ality of reporting in life-sciences articles. n and improve the reliability of published tematically ensure that key methodologi- and we will give more space to methods ne statistics more closely and encourage t, for example by including their raw data. e is a checklist intended to prompt authors d statistical information in their submis- referees to consider aspects important for (go.nature.com/oloeip). It was developed researchers on the problems that lead to ing workshops organized last year by US National Institutes of Health (NIH) institutes. It also draws on published concerns about reporting standards (or the lack of them) and the collective experience of editors at Nature journals.

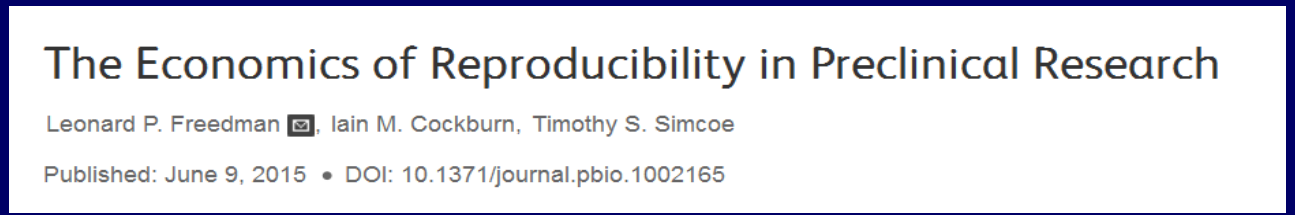
It is enormously expensive

The cost of of low reproducibility in preclinical research (i.e. not just animal research) in the US alone, is estimated at:…

It is enormously expensive

The cost of of low reproducibility in preclinical research (i.e. not just animal research) in the US alone, is estimated at:…

$ 28 000 000 000 !


The Economics of Reproducibility in Preclinical Research

Leonard P. Freedman, Iain M. Cockburn, Timothy S. Simcoe

Published: June 9, 2015 • DOI: 10.1371/journal.pbio.1002165

(Freedman, Cockburn & Simcoe PLOS Biology | DOI:10.1371/, 2015)

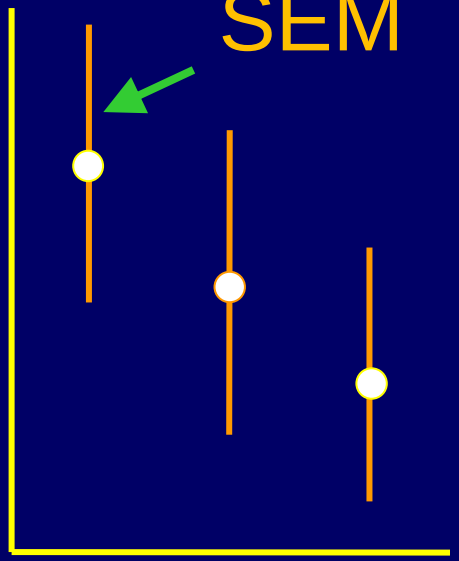
Why we are here: an inescapable LAW in research practice

Why we are here: an inescapable LAW in research practice

$$SEM = \frac{SD}{\sqrt{N}}$$

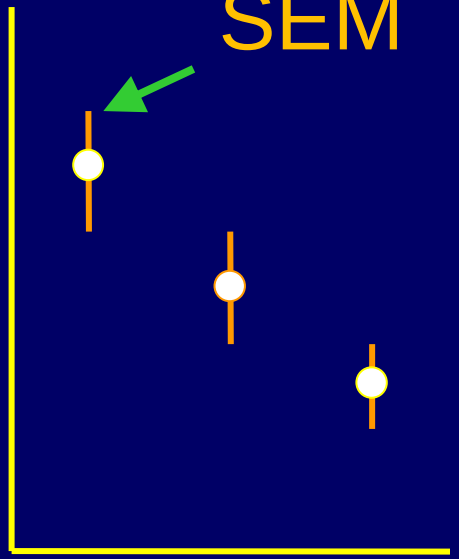
Confidence

$$SEM = \frac{SD}{\sqrt{N}}$$



Confidence

$$SEM = \frac{SD}{\sqrt{N}}$$

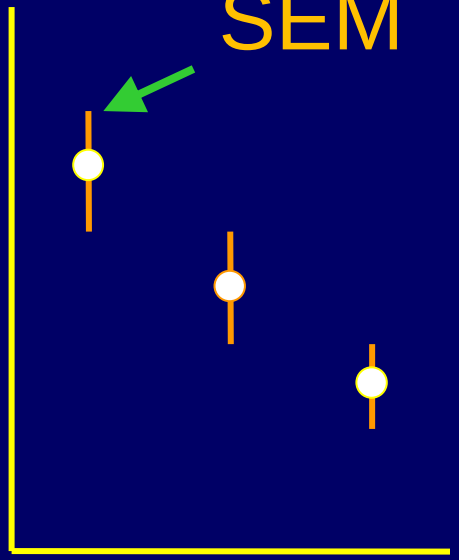


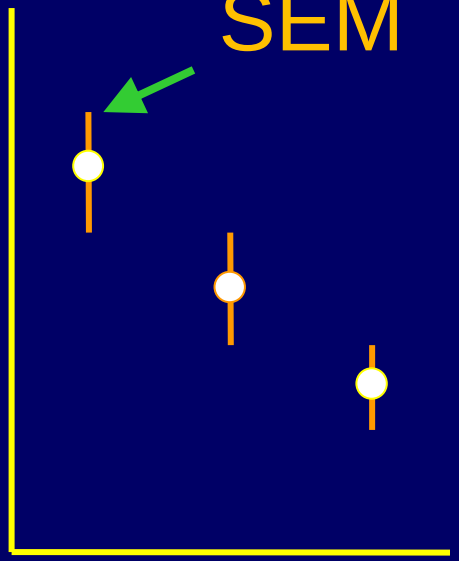
Variation (Noise)

Confidence

$$SEM = \frac{SD}{\sqrt{N}}$$

Sample size
(number of animals)





Confidence

Variation (Noise)

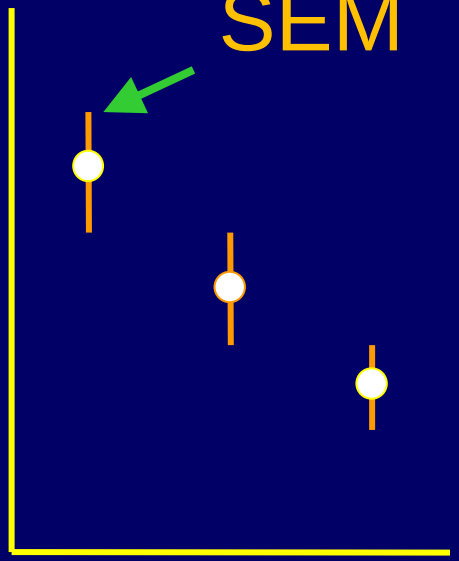
$$SEM = \frac{SD}{\sqrt{N}}$$

Sample size
(number of animals)

REDUCE VARIATION !

- **Increase power** of experiment removing variation through:

- good experimental design (e.g. blocking)
- Good statistical analysis
- Increase effect size (welfare)



Confidence

$$SEM = \frac{SD}{\sqrt{N}}$$

Variation (Noise)

Sample size
(number of animals)

REDUCE VARIATION !

- **Increase power** of experiment removing variation through:

- good experimental design (e.g. blocking)
- Good statistical analysis
- Increase effect size (welfare)

INCREASE WELFARE

- Reduce pain
- Perform procedures competently
- Reduce Stress (incl. husbandry)
- Understand normal behaviour
- Standardisation fallacy

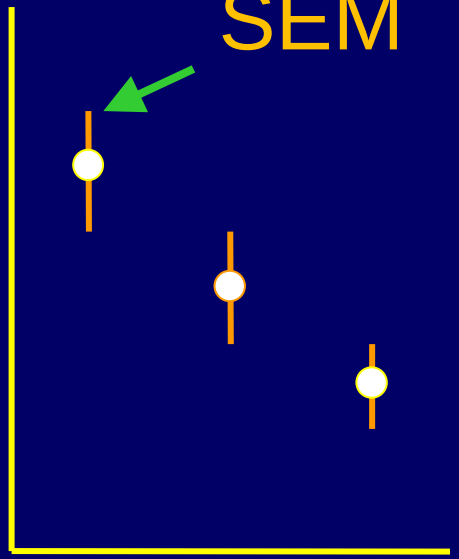
The Lab Rat: A Natural history (Google Ratlife Berdoy)

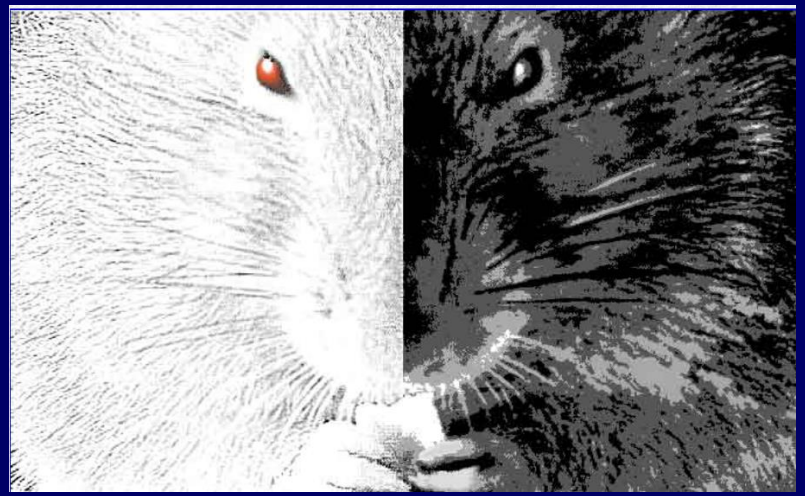
Variation (Noise)

Confidence

$$SEM = \frac{SD}{\sqrt{N}}$$

Sample size
(number of animals)





INCREASE WELFARE

- Reduce pain
- Perform procedures competently
- Reduce Stress (incl. husbandry)
- Understand normal behaviour
- Standardisation fallacy

GOOD SCIENCE: (Reproducible, Efficient and Ethical)

Design of experiments We are not doing a great job !

Welfare of animals Ethical aspects integral to good science
&
Education of researchers / animal users.

Some solutions for the road ahead

Some solutions:

ARRIVE Guidelines: *Animal Research: Reporting of In Vivo Experiments.*

A checklist of information to include in publications describing animal research.

Education, Training and Standards in Animal Use.

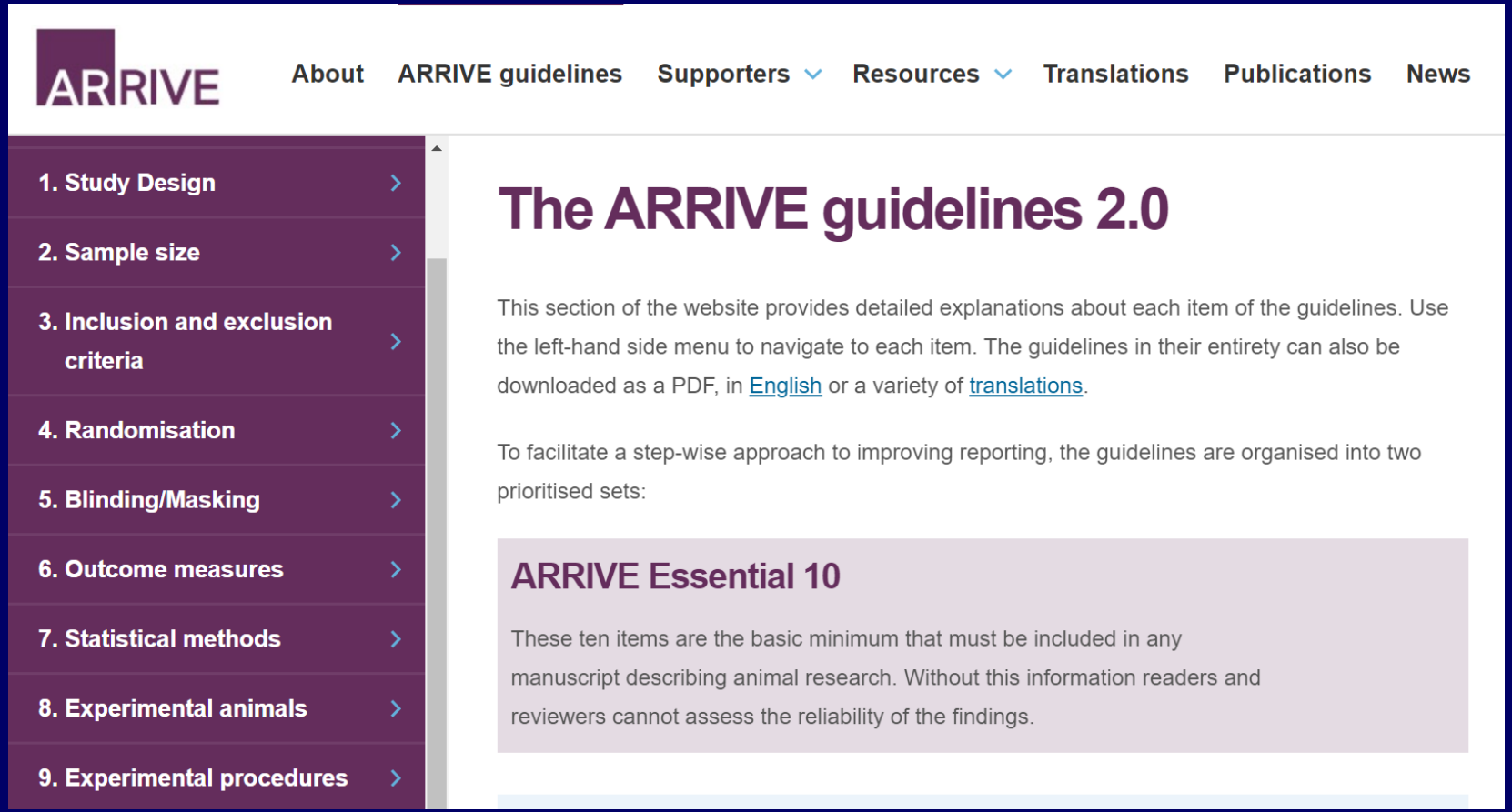
Examples of mechanisms in Europe and the UK:

- **Training before given being able to conduct animal experiments**
- Training is Accredited (UK, FELASA)
- Supervision after training
- Mechanism of competence insurance within institution
- Training of Animal Technicians
- Animal Care standards

Some solutions:

ARRIVE Guidelines: Animal Research: Reporting of In Vivo Experiments.

A checklist of information to include in publications describing animal research.



Some solutions:

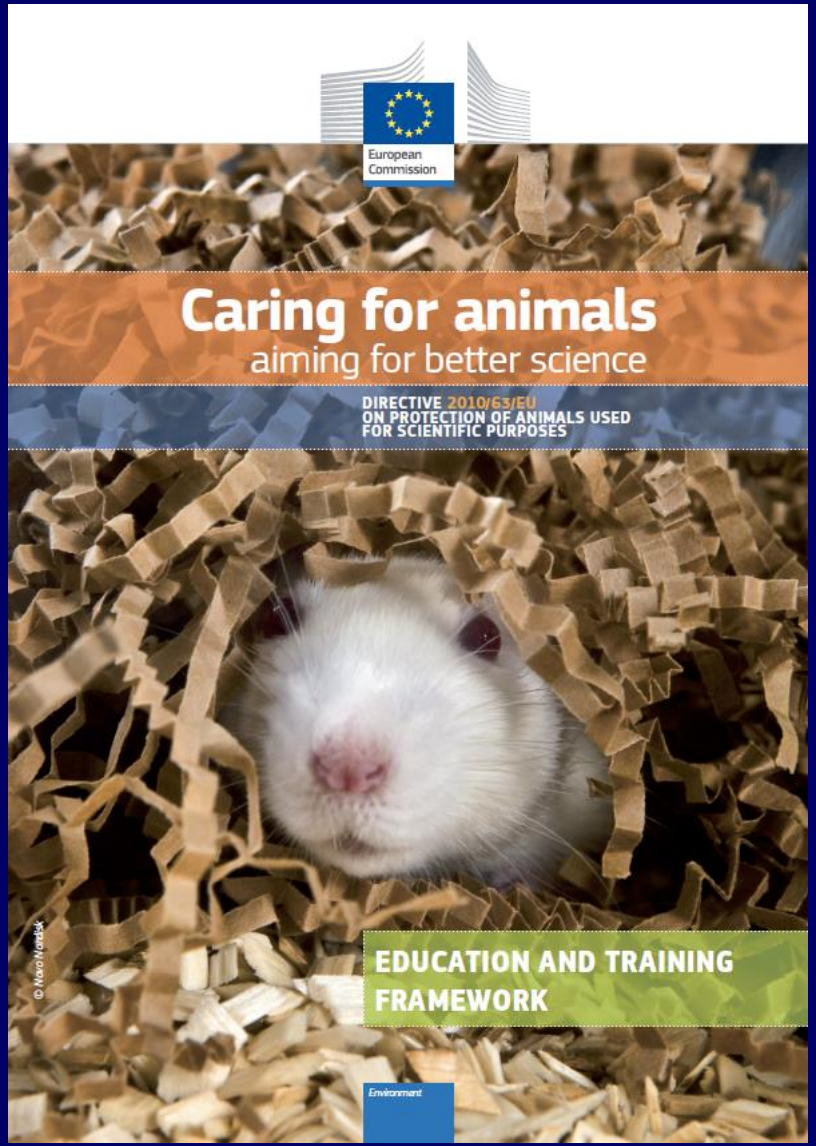
ARRIVE Guidelines: *Animal Research: Reporting of In Vivo Experiments.*

A checklist of information to include in publications describing animal research.

Education, Training and Standards in Animal Use.

Examples of mechanisms in Europe and the UK:

- **Training before given being able to conduct animal experiments**
- Training is Accredited (UK, FELASA)
- Supervision after training
- Mechanism of competence insurance within institution
- Training of Animal Technicians
- Animal Care standards



http://ec.europa.eu/environment/chemicals/lab_animals/pdf/guidance/inspections/en.pdf

EU Framework: Modular Structure

Module Title	EU No		Species Specific
National legislation	1	*(also called module L in the UK)*	
Ethics, animal welfare and the Three Rs (level 1)	2	*(also called module E1 in the UK)*	
Basic and appropriate biology (theory)	3.1		*
Basic and appropriate biology (practical)	3.2		*
Animal care, health and management (theory)	4		*
Recognition of pain, suffering and distress	5		*
Humane methods of killing (theory)	6.1	*(also called module K in the UK)*	*
Humane methods of killing (skills)	6.2	*(also called module K in the UK)*	*
Minimally invasive procedures without anaesthesia (theory)	7		*
Minimally invasive procedures without anaesthesia (skills)	8		*
Anaesthesia for minor procedures	20		
Advanced anaesthesia e.g. for surgical or prolonged procedures	21		
Principles of surgery	22		
Ethics, animal welfare and the Three Rs (level 2)	9	*(also called module E2 in the UK)*	
Design of procedures and projects (level 1) (experimental design)	10		
Design of procedures and projects (level 2) (design and management)	11		

EU Framework: New Modular Structure

Module Title	EU No		Species Specific
National legislation	1	*(also called module L in the UK)*	
Ethics, animal welfare and the Three Rs (level 1)	2	*(also called module E1 in the UK)*	
Basic and appropriate biology (theory)	3.1		*
Basic and appropriate biology (practical)	3.2		*
Animal care, health and management (theory)	4		*
Recognition of pain, suffering and distress	5		*
Humane methods of killing (theory)	6.1	*(also called module K in the UK)*	*
Humane methods of killing (skills)	6.2	*(also called module K in the UK)*	*
Minimally invasive procedures without anaesthesia (theory)	7		*
Minimally invasive procedures without anaesthesia (skills)	8		*
Anaesthesia for minor procedures	20		
Advanced anaesthesia e.g. for surgical or prolonged procedures	21		
Principles of surgery	22		
Ethics, animal welfare and the Three Rs (level 2)	9	*(also called module E2 in the UK)*	
Design of procedures and projects (level 1) (experimental design)	10		
Design of procedures and projects (level 2) (design and management)	11		

EU Framework: New Modular Structure

Module Title	EU No		Species Specific
National legislation	1	*(also called module L in the UK)*	
Ethics, animal welfare and the Three Rs (level 1)	2	*(also called module E1 in the UK)*	
Basic and appropriate biology (theory)	3.1		*
Basic and appropriate biology (practical)	3.2		*
Animal care, health and management (theory)	4		*
Recognition of pain, suffering and distress	5		*
Humane methods of killing (theory)	6.1	*(also called module K in the UK)*	*
Humane methods of killing (skills)	6.2	*(also called module K in the UK)*	*
Minimally invasive procedures without anaesthesia (theory)	7		*
Minimally invasive procedures without anaesthesia (skills)	8		*
Anaesthesia for minor procedures	20		
Advanced anaesthesia e.g. for surgical or prolonged procedures	21		
Principles of surgery	22		
Ethics, animal welfare and the Three Rs (level 2)	9	*(also called module E2 in the UK)*	
Design of procedures and projects (level 1) (experimental design)	10		
Design of procedures and projects (level 2) (design and management)	11		

EU Framework: New Modular Structure

Module Title	EU No		Species Specific
National legislation	1	*(also called module L in the UK)*	
Ethics, animal welfare and the Three Rs (level 1)	2	*(also called module E1 in the UK)*	
Basic and appropriate biology (theory)	3.1		*
Basic and appropriate biology (practical)	3.2		*
Animal care, health and management (theory)	4		*
Recognition of pain, suffering and distress	5		*
Humane methods of killing (theory)	6.1	*(also called module K in the UK)*	*
Humane methods of killing (skills)	6.2	*(also called module K in the UK)*	*
Minimally invasive procedures without anaesthesia (theory)	7		*
Minimally invasive procedures without anaesthesia (skills)	8		*
Anaesthesia for minor procedures	20		
Advanced anaesthesia e.g. for surgical or prolonged procedures	21		
Principles of surgery	22		
Ethics, animal welfare and the Three Rs (level 2)	9	*(also called module E2 in the UK)*	
Design of procedures and projects (level 1) (experimental design)	10		
Design of procedures and projects (level 2) (design and management)	11		

Learning outcomes

Trainees should be able to:

10.1. Describe the concepts of fidelity and discrimination (e.g. as discussed by Russell and Burch and others).

10.2. Explain the concept of variability, its causes and methods of reducing it (uses and limitations of isogenic strains, outbred stocks, genetically modified strains, sourcing, stress and the value of habituation, clinical or sub-clinical infections, and basic biology).

10.3. Describe possible causes of bias and ways of alleviating it (e.g. formal randomisation, blind trials and possible actions when randomisation and blinding are not possible).

10.4. Identify the experimental unit and recognise issues of non-independence (pseudo-replication).

10.5. Describe the variables affecting significance, including the meaning of statistical power and "p-values".

10.6. Identify formal ways of determining of sample size (power analysis or the resource equation method).

10.7. List the different types of formal experimental designs (e.g. completely randomised, randomised block, repeated measures [within subject], Latin square and factorial experimental designs).

10.8. Explain how to access expert help in the design of an experiment and the interpretation of experimental results

Some solutions:

ARRIVE Guidelines: *Animal Research: Reporting of In Vivo Experiments.*

A checklist of information to include in publications describing animal research.

Education, Training and Standards in Animal Use.

Examples of mechanisms in Europe and the UK:

- **Training before given being able to conduct animal experiments**
- Training is Accredited (UK, FELASA)
- Supervision after training
- Mechanism of competence insurance within institution
- Training of Animal Technicians
- Animal Care standards

GOOD SCIENCE: **(Reproducible, Efficient and Ethical)**

Design of experiments **We are not doing a great job !**

Welfare of animals **Ethical aspects integral to good science**
&
Education of researchers / animal users.

Some solutions for the road ahead

Manuel Berdoy

manuel.berdoy@bms.ox.ac.uk

UNIVERSITY OF OXFORD

Thank you.