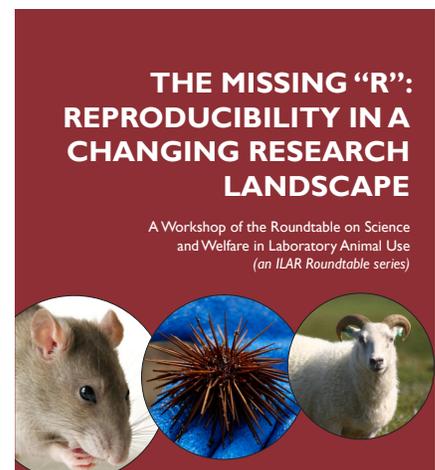


Reproducibility Issues in Research with Animals and Animal Models

Scientific progress is achieved by robust experiments that generate reliable and reproducible results to be used with confidence by the research community. Recent publications have drawn attention to an apparent and concerning prevalence in the number of peer-reviewed studies that cannot be reproduced, particularly those containing data from experiments using animals and animal models. At this workshop¹, researchers from around the world explored the many facets of animal-based research that could contribute to irreproducible results, including perspectives on improving experimental planning, design, and execution; the importance of reporting all methodological details; and efforts to establish harmonization principles of reporting on the care and use of animals in research studies. What follows is a factual summary of the presentations and discussions at the workshop.



June 4 - 5, 2014
2100 C Street NW, Washington, DC 20418
National Academy of Sciences Building, Room 125

THE REPRODUCIBILITY PROBLEM

In his introductory remarks Steven Niemi (American College of Laboratory Animal Medicine) discussed the practice of science focusing on the potential conflict between scientific achievements, career advances and reproducible experimental outcomes, a prevalent theme throughout the workshop. Malcolm Macleod (University of Edinburgh), in his opening talk, noted that irreproducible results cause a lack of faith in the research enterprise and cited Chalmers and Glasziou², who estimated that 85% of research investment/resources is ultimately “wasted.” Macleod noted that many studies fail to randomize or to use blinding, thus biasing results to false positives. In stroke research, he estimates that only 1/3 of published studies are either randomized or blinded and noted an inverse relationship between the reporting of randomization and the impact factor of the journal. Despite the bleak picture of wasteful research, Macleod ended by stating that understanding and acknowledging these problems is the first step to fixing them.

Henry Bourne (University of California - San Francisco) connected reproducibility issues to the competition and overwhelming incentives that scientists and researchers experience when applying for grants or submitting papers for publication. In contrast to previous decades, during which the NIH

¹ <http://nas-sites.org/ilar-roundtable/roundtable-activities/reproducibility>

² Chalmers I, Glasziou P. 2009. Avoidable waste in the production and reporting of research evidence. *The Lancet* 374(9683):86.

budget grew annually, today an expanding population of researchers is forced to compete for significantly fewer research funds, pages in journals, and faculty positions and promotions. Receiving grants and promotions depends on publishing both novel and positive results. Should an experiment produce negative results or results that have been previously shown, the researcher may feel pressure to selectively report data, or emphasize those results that most favorably support the hypothesis³. John P. A. Ioannidis (Stanford Prevention Research Center) expanded on the issues raised by Bourne and Martinson regarding reward systems and their relationship with reproducibility practices:

- Reward mechanisms focus on statistical significance and newsworthiness of results rather than study quality and reproducibility
- Promotion committees misplace emphasis on quantity rather than quality of publications
- With thousands of journals available, anything can get published
- Researchers are interested in publishing results that are characterized by funding agencies as innovative
- Researchers face few negative consequences for publishing flawed or incorrect results, or for making exaggerated claims⁴

PROBLEMS WITH STUDY DESIGN

According to Ioannidis, flawed study design significantly contributes to irreproducible results:

- Poor protocols and documentation
- Poor utility of information
- Statistical power and outcome misconceptions
- Lack of consideration of other (usually prior) evidence
- Subjective, non-standardized definitions subjected to “vibration of effects”⁵.

Glenn Begley (TetraLogic Pharmaceuticals) led a pragmatic discussion of examples of reproducibility issues in the published literature, including sample sizes (“n”) too small to derive statistically relevant conclusions, and exaggerated or downplayed results. He presented Amgen’s experience, whose research teams were unable to reproduce 47 of 53 seminal

publications in oncology drug discovery that claimed a new discovery⁶. Data from some of these papers could not even be reproduced by the original investigators in their own laboratory. The impact of these studies was substantial: multiple clinical trials were initiated and hundreds of secondary publications followed based on irreproducible data.

Begley contended that such issues do not challenge the validity or legitimacy of the scientific method, but result from individual sloppiness, scientific laziness, ignorance, exaggeration, or desperation as well as publication bias. He expanded on his six criteria for judging published scientific reports:

1. Were the studies blinded?
2. Were all results shown?
3. Were experiments repeated?
4. Were positive and negative controls shown?
5. Were reagents validated?
6. Were the statistical tests appropriate?

Victoria Stodden (University of Illinois-Urbana Champaign) argued that reproducibility can be divided into empirical, statistical and computational and pointed to the 2003 National Research Council report *Sharing Publication-Related Data and Materials*, whose first principle is that “authors should include in their publications the data, algorithms, or other information that is central or integral to the publication - that is, whatever is necessary to support the major claims of the paper and would enable one skilled in the art to verify or replicate the claims.” Stodden urged that all computational procedures be included in publications.

STUDY DESIGN, REPRODUCIBILITY AND THE PRINCIPLES OF THE 3 Rs

Michael Festing focused on the connection between reproducibility in research and the Three Rs (replacement, reduction, and refinement)⁷, which provide a structure for every investigator to consider while planning experimental strategy and methodology. According to Festing, a key aspect of study planning and design that affects both reproducibility and the application of the Three Rs is determination of the correct sample size. Although well-designed experiments should give repeatable results that do not depend on sample size (but are subject to specified levels of sampling variation), applying the principle of Reduction would allow the use of the minimum number of animals consistent with the scientific

3 Bourne echoed points from the 2014 PNAS article *Rescuing Biomedical Research from its Systemic Flaws* (Alberts et al.): “The longheld but erroneous assumption of never-ending rapid growth in biomedical science has created an unsustainable hypercompetitive system.....making it difficult for seasoned investigators to produce their best work.”

4 Ioannidis JPA, et al. 2014. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 383:166-175.

5 Ibid.

6 Begley CG, Ellis LM. 2012. Drug development: Raise standards for preclinical cancer research. *Nature* 483, 531-533.

7 Russell WMS, Burch RL. 1959. *The Principles of Humane Experimental Technique*. London: Methuen and Co.

objective. Further, he noted that most false positive results are due to faulty experimental design and incorrect analysis, and that researchers should be deliberately inserting sex and strain differences into their experimental sample pools. He further observed that the use of “factorial design—which enables investigators to study the effect of two or more factors on the response variable—can make experiments both more comprehensive and more efficient”, thus producing added knowledge and a higher degree of precision with the same numbers of observation.

Festing’s list of methodological issues that lead to false positive results includes:

- Selective publication of positive results
- Incorrect randomization
- Failure to blind wherever possible
- Pseudo-replication and incorrect identification of the animal used
- Failure in quality control of experimental materials
- Inadequate external validity
- Inadequate description of methods (e.g., strain nomenclature)
- Incorrect statistical analysis (e.g., no analysis; multiple testing without adjustment; wrong statistical model; incorrect treatment of outliers; cherry-picking the data).

SYSTEMIC ISSUES AND REPRODUCIBILITY

Brian Martinson⁸ (HealthPartners Institute for Education and Research) stated that, historically, the scientific community has focused on “bad scientists” who intentionally alter their findings and less on institutional and systemic factors that threaten the integrity of science. He noted that the reproducibility problem appears to be too large to be caused only by single individuals and that it is irresponsible of the scientific community to not look for larger, overarching causes, such as hypercompetitive systems. Martinson further argued that (work) environments may directly influence the quality and integrity of people’s work by fostering or undermining the integrity of their behavior⁹. Noting that competition for ideas has been substituted by competition for resources and career survival, Martinson cited five factors conducive to cheating, identified in James Lange’s book *Cheating*

Lessons: Learning from Academic Dishonesty:

1. A strong emphasis on performance
2. Very high stakes
3. Extrinsic motivation
4. A low expectation of success
5. A peer culture that accepts or endorses corner-cutting or cheating.

REPRODUCIBILITY ISSUES AND PUBLIC TRUST IN SCIENCE

Robert Bazell, a former journalist now at Yale University, observed that the public was not yet aware of or concerned with reproducibility issues. A bigger problem is the publication of conflicting scientific findings. For example, despite hundreds of papers published over a decade, no one really knows whether dietary carbohydrates are “good” or “bad” for health. The overexpansion of science, which has led to 25,000 journals publishing 3,000 papers daily, has only contributed to the multitude of conflicting research outcomes.

Jan Piotrowski (The Economist), author of the 2013 article “Trouble in the Lab”, noted the absence of reporting on scientific reproducibility until fairly recently. He predicted that additional papers would likely be written on this topic as new research shows deficiencies in scientific rigor and as more scientists focus on this problem. Piotrowski acknowledged that science reporters contribute to the public’s misconceptions about science because they report on the “snazziest of the snazziest” published papers, often those with very large biological effects, which, as pointed out by other speakers, can often be plagued by reproducibility problems.

Roger Reeves (Johns Hopkins University) agreed that scientific literature is misunderstood by the public, the popular press, and -to some degree- by scientists themselves. “It seems we have come to a point where anything that is published in a scientific journal is supposed to be absolute, infallible, and incontrovertible truth.” Instead, Reeves added, journals should be the “stock market” of ideas for scientific discovery that will succeed or fail based on objective parameters, contrary to the assumption that any deviation from absolute truth must be the product of fraud, deceit or incompetence.

Reeves said there is little reward and much cost for doing science properly—i.e., formulate a hypothesis, do everything possible to disprove it, and publish the results so that others can try to disprove or improve it. In contrast, there is substantial reward and enormous pressure to publish.

⁸ Brian Martinson is a member of the Committee revising the 1992 NRC report *Responsible Science: Ensuring the Integrity of the Research Process*. The views presented are his own and do not reflect deliberations of that Committee.

⁹ Institute of Medicine and National Research Council Committee on Assessing Integrity in Research Environments. 2002. *Integrity in Scientific Research: Creating an Environment that Promotes Responsible Conduct*. The National Academies Press - Washington, DC.

REPRODUCIBILITY IN RESEARCH WITH ANIMAL MODELS

Despite the advent of many new in vitro research options, animal models remain central to conducting research.

Mice

Reeves noted that, since 2012, genome editing technologies, such as CRISPR/Cas9, have enabled faster, less expensive and more reliable methods of transgenic manipulation of the mouse genome. However, optimal use of the mouse as a genetic model system is often impacted by the lack of justification for the use of inbred strains vs. outbred stocks when modeling human conditions, as some traits are much more variable between individuals of inbred than of outbred strains. Reeves also discussed the problem of variable protocols across laboratories that limit comparisons between studies; as he said, “not all tests with the same name are the same test”. To rectify this situation, he recommended the creation of minimum standard phenotyping tests for behavior, pharmacology and metabolism, such as those used by the International Mouse Phenotyping Consortium¹⁰ to build the first comprehensive, functional catalogue of phenotypes for every gene in the mouse genome.

Zebrafish

Monte Westerfield (University of Oregon) explained the unique advantages of zebrafish as a model system to investigate certain human biology and disease phenotypes, which have led to a dramatic increase of the number of NIH grants that fund research using mutated zebrafish. However, the use of morpholinos (i.e., antisense oligonucleotides used to create knockdown gene functions) affects reproducibility of zebrafish-based data:

- In high doses morpholinos can be toxic and induce stress that confounds experimental results
- The effect of a genetic manipulation induced at the one-cell stage can become diluted over time
- Gene targeting relies upon binding of the morpholinos to RNA, so the presence of genomic polymorphisms can alter the binding; thus polymorphisms in a stock or between stocks of different laboratories can affect the results of the experiments.

Westerfield observed that the reproducibility of zebrafish-derived experimental data is improved by

¹⁰ www.mousephenotype.org

the use of advanced genomic editing tools, like CRISPR/Cas9, and the advent of free, open-access repositories (e.g., Zebrafish Mutation Project¹¹), to create precise models of human patients.

Nonhuman primates

Reproducibility of studies using nonhuman primates is influenced by differences in methods and protocols across laboratories (echoing Roger Reeves’ prior presentation), small sample sizes, and genetic differences among study subjects said Jeff Rogers (Baylor College of Medicine). The primary reason for small sample sizes in -otherwise carefully designed primate studies- is cost. Primate species differ significantly from each other: e.g., prostate specific antigen (PSA) cannot be studied in New World monkeys because they do not have the PSA gene. While such diversity across populations can be valuable, it creates reproducibility problems if the origin of the animals is not explicitly identified, as the genome sequencing of 144 rhesus macaques from research facilities revealed more genetic diversity than shown by data from the human 1000 Genomes project. Rogers recommended that researchers a) evaluate the background and genetics of their study animals and b) choose populations and individuals appropriate to their research questions.

HARMONIZATION OF PRINCIPLES OF ANIMAL CARE, USE AND REPORTING

Coenraad Hendriksen discussed a 2009 ILAR workshop, which he chaired, on the challenges of conducting animal research globally. According to Hendriksen, one of the workshop’s messages was that harmonization, not standardization, of laboratory animal care and use is needed. “Each country needs to establish an animal welfare oversight system that reflects its own culture, tradition, religion, laws and regulations”, he said.

Gilly Griffin (Canadian Council on Animal Care) noted that the 2011 Montreal Declaration on Synthesis of Evidence to Advance the 3Rs Principles in Science¹² called for a change in the culture of planning, executing, reporting, reviewing, and translating of animal-based research because of three key concerns regarding animal-based studies:

- Animal studies are poorly designed and reported (Kilkenny et al. 2009)
- Animal studies are not translating into clinical treatments (Macleod et al. 2005)

¹¹ <https://www.sanger.ac.uk/resources/zebrafish/zmp>

¹² http://www.wc8.ccac.ca/files/WC8_Declaration_of_Montreal_FINAL.pdf

- Animal lives are being wasted. (Hooijmans et al. 2010; Leenars et al. 2012)

Also in 2011, Jeff Everitt (GlaxoSmithKline) chaired the NRC committee that authored the *Guidelines for Scientific Publications Involving Animal Studies*¹³, intended primarily for editors of scientific journals. Everitt argued that, because reporting of methodological information relating to laboratory animal care and use differs based on the field of study, the journal, and the type of study, journal editors could customize their needs but should publish their journal's expectations; issue guidance to editors, authors and reviewers; and articulate clear policies on animal use and ethical review.

In 2013, Griffin said, the International Council for Laboratory Animal Science (ICLAS¹⁴) established a working group to conduct an analysis of available international reference documents (e.g., the 2011 aforementioned NRC Guidelines; the 2010 ARRIVE Guidelines¹⁵; the 2010 Gold Standard Publication Checklist¹⁶; the MIBBI Project¹⁷; various professional societies' guidelines and other checklists, e.g., Landis et al. 2012) and develop a set of harmonization principles of reporting on the care and use of animals in experimental procedures (see Box 1.)

BOX 1

The ICLAS Harmonization Principles

1. Ethical statement
2. Details of scientific objectives (rationale for the experimental approach)
3. Details relating to study funding
4. Details of study design
5. Details of animal subjects (including numbers and attrition)
6. Details of experimental protocols (perhaps include justification of withholding pain medication)
7. Details of housing and husbandry

Additional principles requested by journal editors

8. Details of permits, consent for client-owned animals
9. Description of how raw data will be shared
10. Completed checklist at time of submission

13 <http://www.nap.edu/catalog/13241/guidance-for-the-description-of-animal-research-in-scientific-publications>

14 <http://iclas.org/>; An international scientific organization dedicated to advancing human and animal health by promoting the ethical care and use of laboratory animals in research worldwide

15 <https://www.nc3rs.org.uk/arrive-guidelines>

16 http://www.3rs-reduction.co.uk/assets/applets/Gold_standard_guidelines.pdf

17 <http://www.dcc.ac.uk/resources/metadata-standards/mibbi-minimum-information-biological-and-biomedical-investigations>

Learning from Clinical Research

Ethicist Jonathan Kimmelman (McGill University) pointed out that animal experiments establish a cause-and-effect relationship between a drug and a disease response that will generalize to human patients. He argued that accurate assessment of this relationship is critical: as sentient beings, animals experience suffering, therefore their sacrifice should be for a greater benefit to mankind; preclinical research establishes the rationale and justification to expose patients to unproven and potentially harmful drugs; animal research directly informs healthcare practices, as physicians often rely on information from preclinical studies to treat idiosyncratic conditions. Kimmelman described mechanisms used in clinical studies to assess validity threats:

Planning: In Canada, clinical studies are required to undergo systematic review prior to initiation.

Design: Clinical studies have adopted a system of practices to address validity threats and minimize the effects of bias. As discussed by other presenters, preclinical studies have only sporadically adopted these practices.

Reporting: In clinical research, two mechanisms ensure accurate and complete reporting: (1) clinical trial registration and (2) reporting guidelines, both of which are either absent or just getting underway in preclinical research. Preclinical research further suffers from difficulties in accessing information, especially from commercial studies, as well as incomplete reporting.

Uptake of findings: In clinical medicine, a whole suite of measures is used to promote uptake of findings, including standardization of research practices, which enables side-by-side comparison and aggregation of findings in systematic reviews.

EFFORTS TO DATE ARE FALLING SHORT

Several participants noted that efforts to curb or reverse reproducibility problems to date have not met expectations and concerns have intensified rather than lessened. Furthermore, the number of variables within animal experiments that may be contributing to the issue has grown as the causes of these problems are being more extensively studied.

Several participants said that, while the ARRIVE guidelines are very useful there is still a gap between their endorsement, adoption and execution. Kimmelman noted that some actors have employed innovative strategies to ensure replication, e.g., the journal *Cortex* accepts submissions (called *Registered*

Reports¹⁸) on the basis of their experimental protocol before experiments are conducted. Other stakeholders are not very active, including regulators, such as the United States Food and Drug Administration, which does not prioritize rigor and validity of preclinical studies. He further argued that although Institutional Review Boards and Institutional Animal Care and Use Committees (IACUCs) are charged with maintaining a favorable risk-benefit ratio based on the quality of evidence for initiating studies, they have not used their authority to adequately address these issues.

Jerry Collins (Yale University) similarly observed that risk-benefit analysis of animal-based experiments is a primary IACUC responsibility as is the broader oversight of all elements within an Animal Care and Use Program that may influence the health and well-being of research animals. He further argued that IACUC functions contribute to enhanced reproducibility by ensuring humane handling, care and treatment of research animals and that the committee should evaluate scientific aspects of research protocols affecting the welfare and use of laboratory animals, including hypothesis testing, sample size determination, and adequacy of controls.

IMPROVING THE RELIABILITY OF PUBLISHED RESULTS

Elizabeth Marincola (PLOS) noted that recently PLOS introduced a new policy requiring that all data leading to findings in a paper under consideration for publication be provided at the time of submission, preferably in a repository, but if not, as a supplement. Her colleague, Damian Pattinson (PLOS), presented a 2011 *Science* survey¹⁹ about difficulty in accessing data: 50% of it resides in laboratories, 38.5% on university servers, and only about 10% is in community repositories or otherwise accessible. Pattinson said that a number of journals, including PLOS, EMBO, BMC, F1000 and eLife, are working on an open data “badge”.

Marincola and Pattinson said that PLOS has always encouraged the publication of negative (or neutral) results and is exploring the use of mechanisms like clinicaltrials.gov, which encourage pre-registration of experimental protocols. As one procedure to correct literature post-publication, PLOS is working on a mechanism to link original papers to related subsequent work in order to enable researchers to understand the full trajectory of a study. This may be supplemented by post-publication peer review

Authors and reviewers of all 14 journals

18 <http://www.elsevier.com/reviewers-update/story/innovation-publishing/registered-reports-a-step-change-in-scientific-publishingobserved>

19 <http://www.sciencemag.org/content/331/6018/692.short>

published by the American Physiological Society are instructed to verify that a study can be replicated, said Gaylen Edwards. Results of an internal survey of the journals’ editors supported publication of negative data from sound and properly reported experimental methods but expressed concern about the potential diminishing effect on the journals’ impact factor from such articles as they may be cited less often. Edwards said that the biggest concern was the number of animals used in experiments and argued that the Reduction principle and the Three Rs could be replaced by “the Three Os”: Optimize animal numbers so that Outcomes are reliable to help Overcome issues of reproducibility.

ADDRESSING THE PROBLEM

Kathryn Bayne (AAALAC International) pointed to the recent National Science Board recommendation to “develop standard operating procedures and a single set of guidelines that can be cited in IACUC protocols”²⁰, but many workshop speakers and participants expressed concerns about additional regulatory burden by funding agencies to rectify lack of reproducibility and to preserve the public’s trust and investment in science.

Paul Braunschweiger (CITI Program) cited a 2009 meta-analysis showing increased frequency in the reporting of questionable research practices, including incomplete data reporting and the use of inappropriate statistical methods. Reflecting back on the presentations of Begley and Festing, he also thought that reproducibility issues are not equivalent to misconduct or fraud, but he concluded that lack of reproducibility and sloppiness in planning, execution and reporting of animal research is unprofessional and that promoting data integrity and reproducibility of science should be a shared responsibility across the community.

Glenn Begley’s belief is that investigators and their institutions are ultimately responsible for and should be accountable for the quality of their publications. Funding agencies can play a significant role by raising the standards for grants and publications, e.g., by requiring the presentation of preclinical proposals in advance; by recognizing the value of publishing confirmatory data; and by rewarding findings that refute high profile studies.

Jonathan Kimmelman suggested a formal process for developing reporting and practice guidelines for preclinical research. He further clarified that preclinical testing should emulate the stages of clinical research: exploratory (i.e., early phase studies with flexible designs, small sample sizes, and surrogate

20 www.nsf.gov/pubs/2014/nsb1418/nsb1418.pdf

endpoints to measure response) and confirmatory (i.e., late stage studies with adequately powered samples using clinical endpoints and a pre-specified design). He urged that regulatory, funding, institutional and publishing structures be established to encourage researchers to change their approach to planning and uptake of findings in preclinical research.

Widespread use of new computational tools should be encouraged, said Victoria Stodden, but she also cautioned that verification of computational procedures would depend on a fixed version of software, as recommended in the 2012 Institute of Medicine report *Evolution of Translational Omics: Lessons Learned and the Path Forward*²¹.

John Ioannidis discussed the role of proactive planning in ensuring reliable data from preclinical research studies. In his 2005 study he showed that the odds of a research outcome being true diminish in the presence of bias, small effect size or small studies; in “hot” fields with significant competition among research teams; when results are highly anticipated; when datasets are not targeted; and when statistical analyses are more flexible. In the world of animal studies, randomization and blinded assessment of outcomes are probably two of the most important preventive actions against irreproducible results. He stressed that anticipating the magnitude of the effect-to-bias ratio is necessary in order to decide whether the proposed research is even justified. Setting minimal design prerequisites by journals and funding agencies would further help reduce the effect-to-bias threshold to acceptable levels²².

Ioannidis described additional opportunities to improve study design:

- Public availability and/or registration of protocols or complete documentation of the exploratory process
- A priori examination of the utility of information
- Consideration of both prior and ongoing evidence
- Standardization of measurements, definitions, and analysis²³.

Ghislaine Poirier described the development of a strategy at GlaxoSmithKline to improve the use of animals and animal models internally as well as in external collaborations by increasing the sharing of animal-derived data. Like many speakers before her, she emphasized that cultural change is happening across the research community as both individuals and institutions (funding, academic, private) recognize the diverse benefits gained by data sharing: reputational, scientific and on behalf of the animals. She, however, cautioned that data sharing carries certain risks (e.g., misinterpretation of data, loss of competitiveness, loss of data integrity) that should not be disregarded.

Kent Lloyd (University of California - Davis) concluded the workshop by proposing a convocation of all stakeholders committed to overcoming the challenges of reproducibility. While meetings focused on one or more aspects of the causes of irreproducible research have taken place, according to Lloyd there is a need to bring all groups together in an open forum to create a community-wide, bottom up consensus and agreement on next steps. Bringing scientists, institutional representatives, research veterinarians, journal editors, funders, and members of the public together will allow discussions in areas of broad agreement and divergence, and help strengthen the commitment of the scientific community to uphold and enact principles of reproducibility in research involving animals.

21 <http://www.nap.edu/catalog/13297/evolution-of-translational-omics-lessons-learned-and-the-path-forward>

22 Ioannidis JPA, et al. 2014. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 383:166-175.

23 *Ibid.*

WORKSHOP INFORMATION

Planning Committee for Reproducibility Issues in Research with Animals and Animal Models: **Kent Lloyd** (*Co-Chair*), University of California - Davis; **Steven Niemi** (*Co-Chair*), American College of Laboratory Animal Medicine; **Bonnie V. Beaver**, Texas A&M University; **Brian R. Berridge**, GlaxoSmithKline; **Pamela Chamberlain**, Food and Drug Administration; **Carol Clarke**, United States Department of Agriculture; **Margaret S. Landi**, GlaxoSmithKline; **Malcolm Macleod**, University of Edinburgh, United Kingdom; **Brian C. Martinson**, HealthPartners Institute for Education and Research; **Susan Brust Silk**, National Institutes of Health.

Staff: **Lida Anestidou**, Director, ILAR Roundtable; **Angela Kolesnikova**, Administrative Assistant; **Jenna Ogilvie**, Senior Program Assistant; **Bethelhem Mekasha**, Financial Associate.

DISCLAIMER: This Workshop in Brief has been prepared by Nancy Huddleston, Jenna Ogilvie and Lida Anestidou as a factual summary of what occurred at the meeting. The committee's role was limited to planning the meeting. The statements made are those of the authors or individual meeting participants and do not necessarily represent the views of all meeting participants, the planning committee, the ILAR Roundtable, or the National Academies.

The summary was reviewed in draft form by Gilly Griffin, Brian Martinson, Timo Nevalainen and Emily Sena to ensure that it meets institutional standards for quality and objectivity. The review was coordinated by Janet Garber, Private Consultant, and the comments and draft manuscript remain confidential to protect the integrity of the process.

Statement of Task

An ad hoc committee will organize and conduct a public workshop to discuss fundamental aspects of experimental design of research using animals and animal models, aimed at improving reproducibility. The workshop will include invited speakers to provide background and context on how to ensure that animal studies are of sufficient quality and relevance, and described in adequate detail for the findings to provide an evidence base for any decision to proceed into clinical trials or other outcomes of public importance. The ad hoc committee will develop a workshop agenda, select and invite speakers and discussants, and moderate the workshop discussions. An individually-authored summary of the presentations and discussions at the workshop will be prepared by a designated rapporteur in accordance with institutional guidelines.

About the Roundtable on Science and Welfare in Laboratory Animal Use

The ILAR Roundtable was created to promote the responsible use of animals in science, provide a balanced and civil forum to stimulate dialogue and collaboration, and help build trust and transparency among stakeholders. Roundtable members comprise entities with strong interests in the use of laboratory animals in research, testing and education, including government agencies, leading pharmaceutical and consumer product companies, contract research organizations, animal advocacy groups, professional societies, and prominent academic institutions.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.

www.national-academies.org