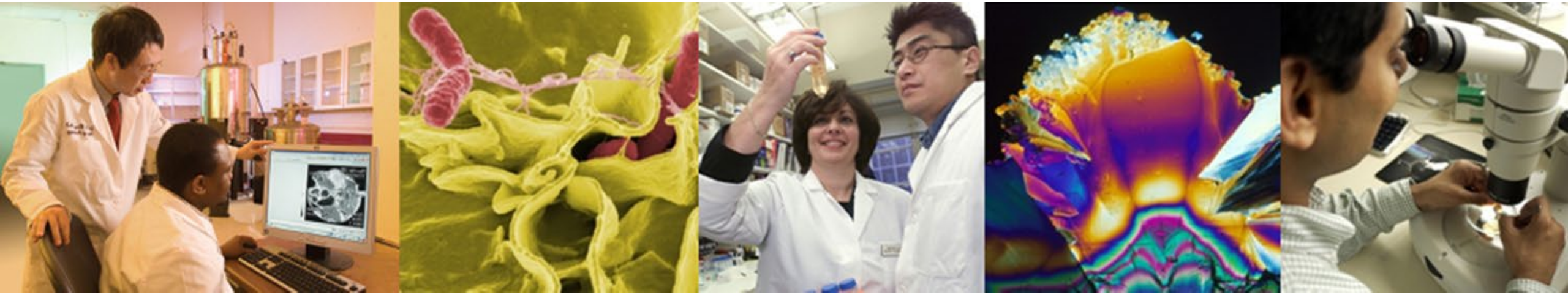
ACD Working Group on Enhancing Rigor, Transparency, and Translatability in Animal Research
Final Report

122st Meeting of the Advisory Committee to the Director (ACD)

June 11, 2021



Barbara Wold, PhD
Bren Professor of Molecular Biology
Allen V.C. Davis and Lenabelle Davis Leadership Chair
Director, Merkin Institute for Translational Research
California Institute of Technology

Lawrence A. Tabak, DDS, PhD
Principal Deputy Director, NIH
Department of Health and Human Services

Agenda

- **Review Charge**
- Background
- Recommendations

Charge

- **Identify gaps and opportunities** to improve the rigor, reproducibility, translational validity, and transparency of animal models studies
- Evaluate **how animal models of human disease are currently developed, validated, and accepted into routine use** and how this process could be improved
- Consider the process for **validating alternative models** to animal research
- Consider benefits and burdens of **registering animal studies** that aim to lead to first human trials
- Model **financial implications of potential changes in the average costs** of grants using animal models, the number of studies funded, or the need to develop consortia to achieve appropriate statistical power
- Consider how rigor in animal research is incorporated into **training**

ACD Enhancing Rigor, Transparency, and Translatability in Animal Research Working Group Members

EXTERNAL MEMBERS

Barbara Wold, PhD (Co-Chair)
Bren Professor of Molecular Biology
Merkin Institute for Translational Research
California Institute of Technology

Nancy Ator, PhD
Professor of Behavioral Biology
Johns Hopkins School of Medicine

Lais Berro, PhD
Postdoctoral Fellow
University of Mississippi Medical Center

Eliza Bliss-Moreau, PhD
Associate Professor; Core Scientist
University of California, Davis

Romer A. Gonzalez Villalobos, MD, PhD, FAHA
Senior Principal Scientist
Janssen Research and Development, LLC

F. Claire Hankenson, DVM, MS, DACLAM
Attending Veterinarian; Director, Campus Animal Resources; Professor
Michigan State University

Veronique Kiermer, PhD
Publisher & Executive Editor
PLOS

Keisa W. Mathis, PhD
Assistant Professor
University of North Texas Health Science Center

Sarah Nusser, PhD
Professor of Statistics
Iowa State University

Regina Nuzzo, PhD
Senior Advisor for Statistics
American Statistical Association

Eric Prager, PhD
Associate Director, External Affairs
Cohen Veterans Bioscience

F. Daniel Ramirez, MD, MSc
Cardiac Electrophysiology Fellow
University of Ottawa Heart Institute
CHU Bordeaux, IHU Liryc

Karen Svenson, PhD
Senior Scientific Program Manager and Research Scientist
Jackson Laboratory

ACD Enhancing Rigor, Transparency, and Translatability in Animal Research Working Group Members

USG MEMBERS

Lawrence A. Tabak, DDS, PhD (Co-Chair)
Principal Deputy Director, NIH

Brian Berridge, DVM, PhD, DACVP
Associate Director, National Toxicology Program;
Scientific Director Division National Toxicology Program
National Institute of Environmental Health Science, NIH

Paul Brown, PhD
Associate Director for Pharmacology and Toxicology
Office of New Drugs, Center for Drug Evaluation and Research, FDA

Janine Clayton, MD
Director
Office of Research on Women's Health, NIH

Joshua A. Gordon, MD, PhD
Director
National Institute of Mental Health, NIH

Michael Lauer, MD
Deputy Director for Extramural Research
Office of Extramural Research, NIH

Robyn Lee-Stubbs, MS, CPIA, PStat®
IACUC Chair/Statistician
United States Army Medical Research Institute of Chemical Defense

Glenn Merlino, PhD
Scientific Director for Basic Research
Center for Cancer Research, National Cancer Institute, NIH

Shai Silberberg, PhD
Director for Research Quality
National Institute of Neurological Disorders and Stroke, NIH

Carrie Wolinetz, PhD
Acting Chief of Staff; Associate Director
Office of Science Policy, NIH

Major Editors

Regina Nuzzo, PhD, is a freelance science writer, statistician, and professor at Gallaudet University. Her science journalism specialties center around data, probability, statistics, and the research process. Her work has appeared in Nature, Los Angeles Times, New York Times, Reader's Digest, New Scientist, and Scientific American, among others. Her communication and editing skills were invaluable to the working group in building the final synthesis and completing our report.

https://www.reginanuzzo.com/ and https://my.gallaudet.edu/regina-nuzzo

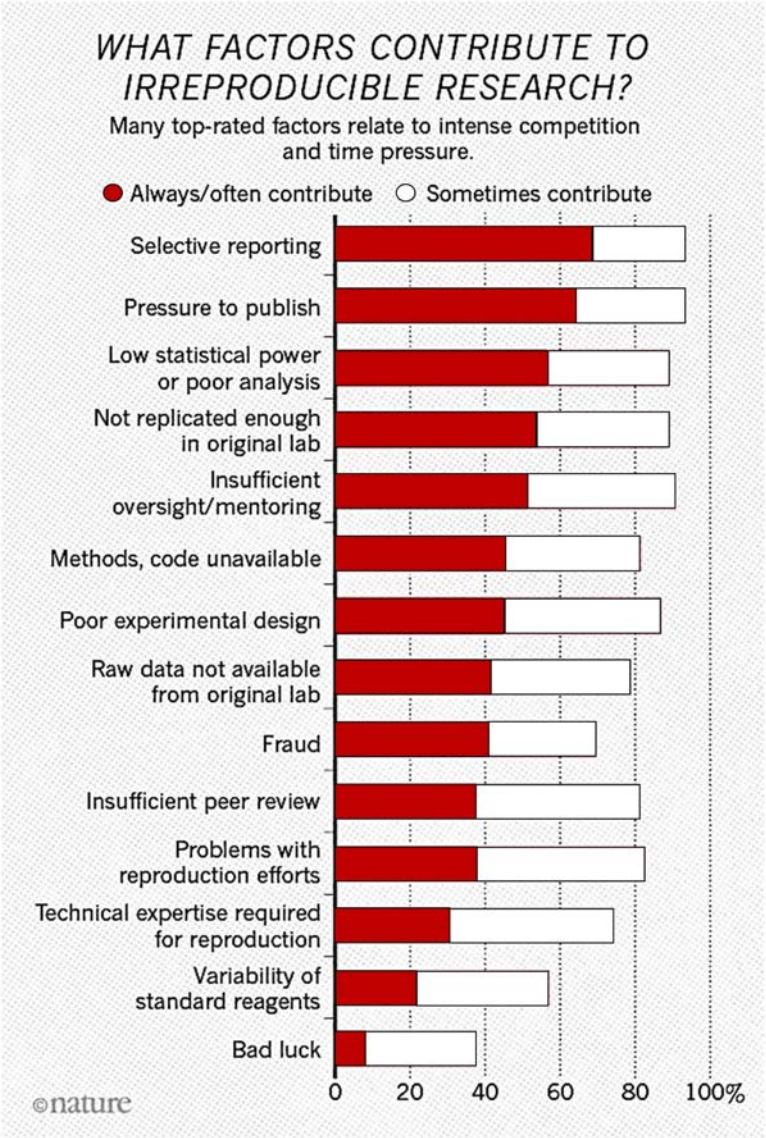
Jordan Gladman, PhD, was the Special Assistant to the NIH Principal Deputy Director and served as the Executive Secretary for our WG. He has joining the NINDS focused on Scientific Management and Operations.

Agenda

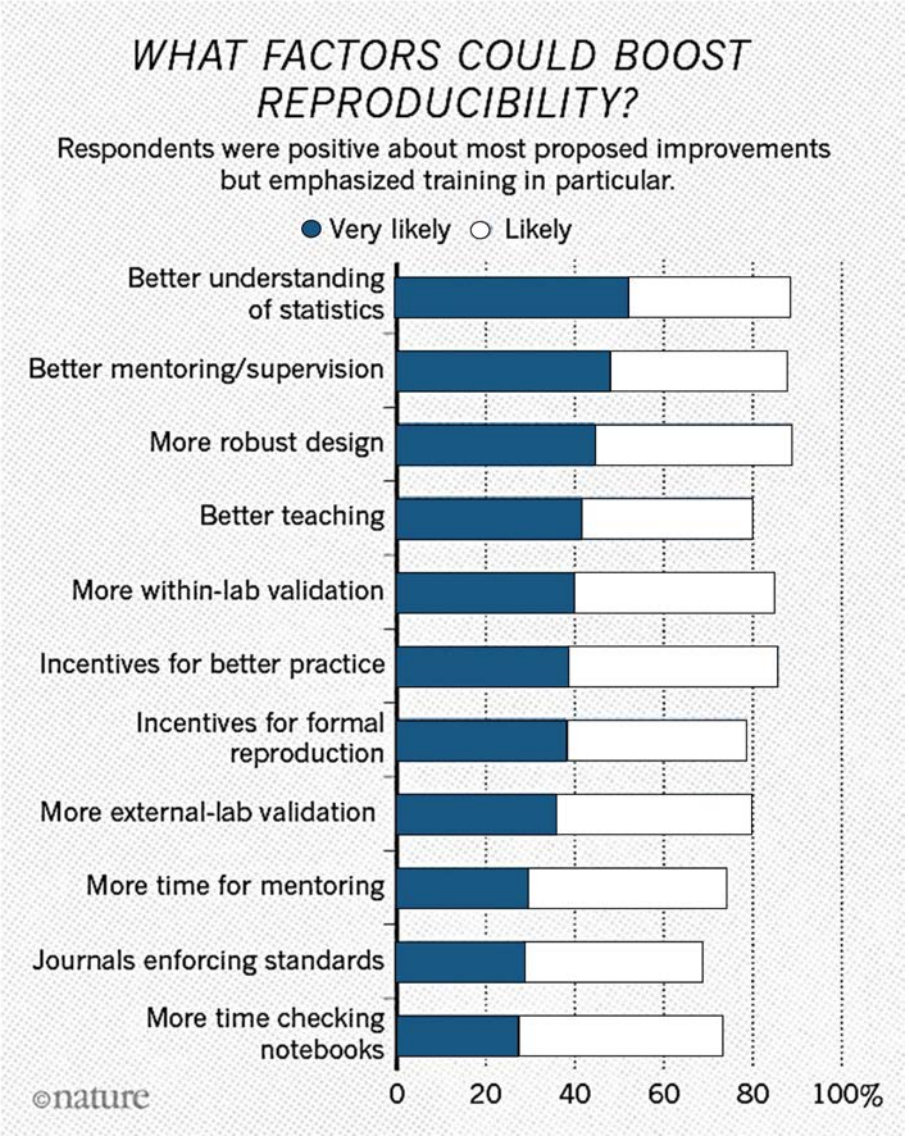
- Review Charge
- **Background**
- Recommendations

Impetus: Rigor and Reproducibility Shortfall

Community view that we can do better; request training, incentives



- Nature's survey of 1,576 researchers
- Questionnaire on reproducibility in research



Baker M. 1,500 scientists lift the lid on reproducibility. Nature. 2016 May 26;533(7604):452-4. doi: 10.1038/533452a. PMID: 27225100.

AWG Framework

- **Rigor** - Application of the scientific method to ensure unbiased and well-controlled experimental design, methodology, analysis, interpretation, and reporting of results
- **Reproducibility**
  - **Methods reproducibility** - Providing enough procedural detail and data to repeat successfully
  - **Results reproducibility** - Getting the same results from a new study with procedures as close to the original as possible
  - **Inferential reproducibility** - Drawing similar conclusions or making knowledge claims of similar strength from study replications and re-analyses

Scope: Why Animal Research Rigor? Which Animals?

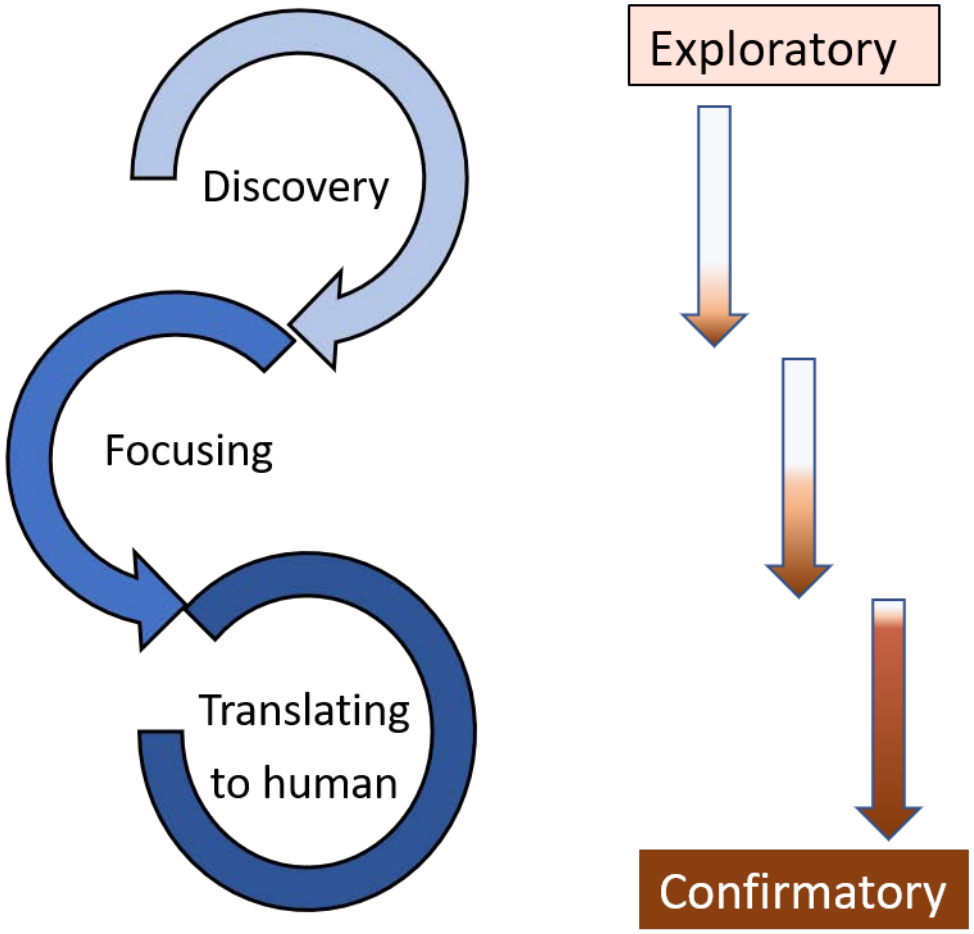
Animals: Additional challenges for rigor and reproducibility

- Investigators must balance smallest possible number of animals for ethical reasons versus large enough number to achieve statistical power
- Animal husbandry design, data, reporting

Animal scope: Vertebrate and cephalopod species

Scope: Intersection with Translatability

Translation: Applying results from preclinical research, usually via late-stage preclinical animal studies, to justify, design and inform trials in humans



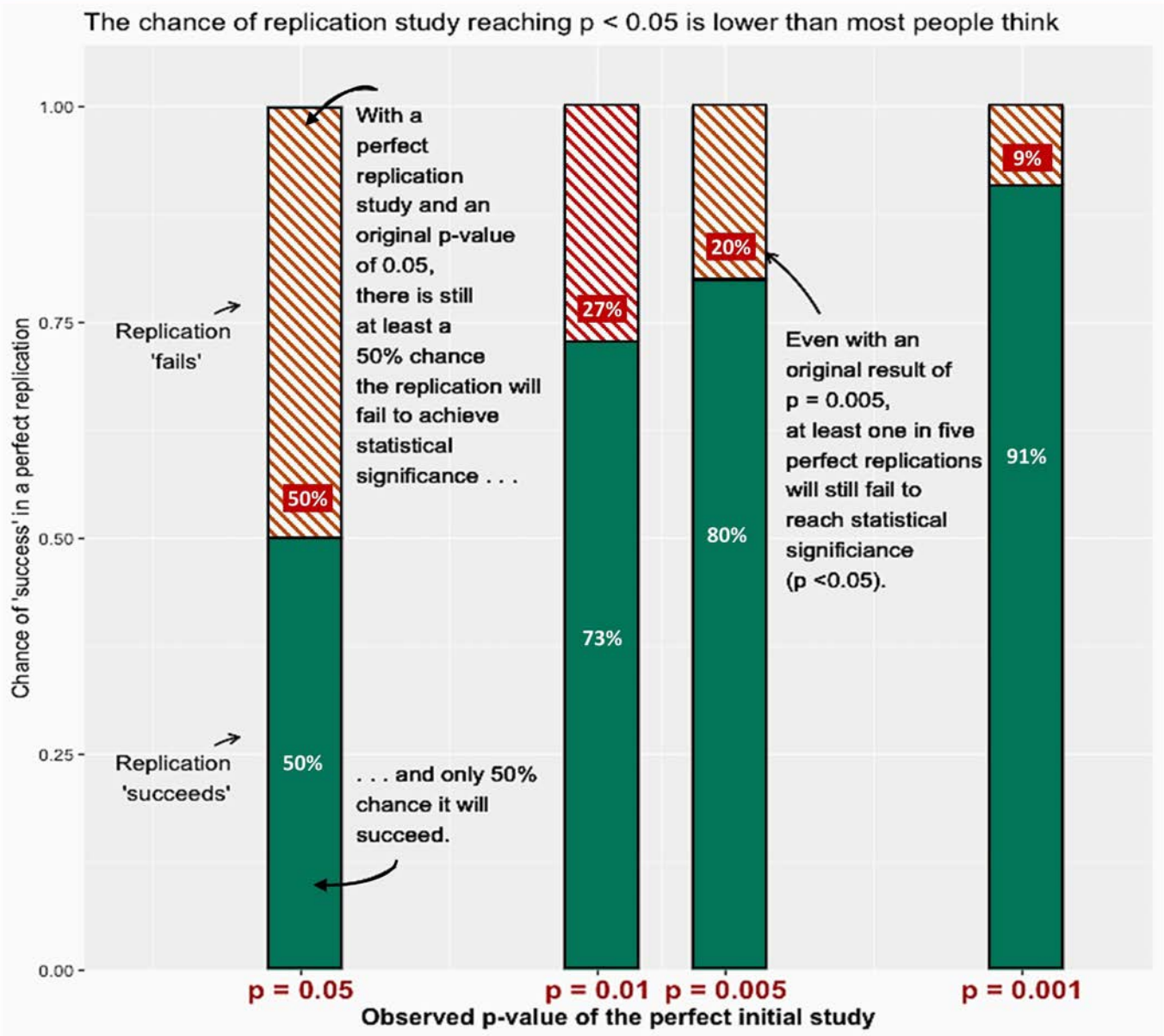
Translatability Challenges

- Validity of a given animal to model specific human biology or disease
- High dollar and opportunity costs when translation fails
- Proximal to translation, rigor tools and solutions may differ

Outcomes of High Rigor
Managing Expectations of Statistical Analysis

Statistical significance is not enough to judge reproducibility.

Given a statistically significant initial study, the chance of a replication "succeeding" (another statistical significance; $p < 0.05$) is surprisingly low.

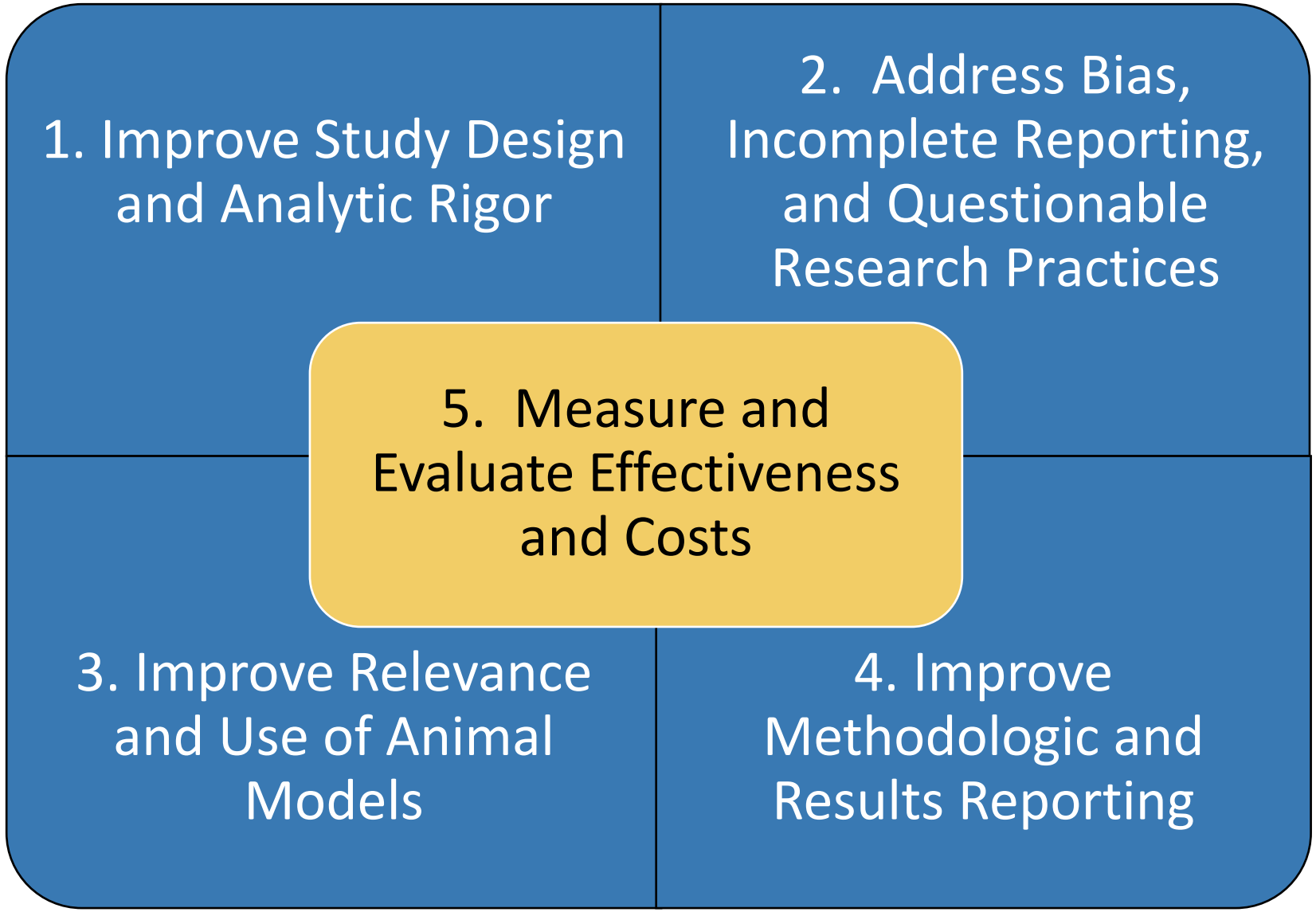


For details, see: Goodman, Steven N. *Statistics in medicine* (1992).

Agenda

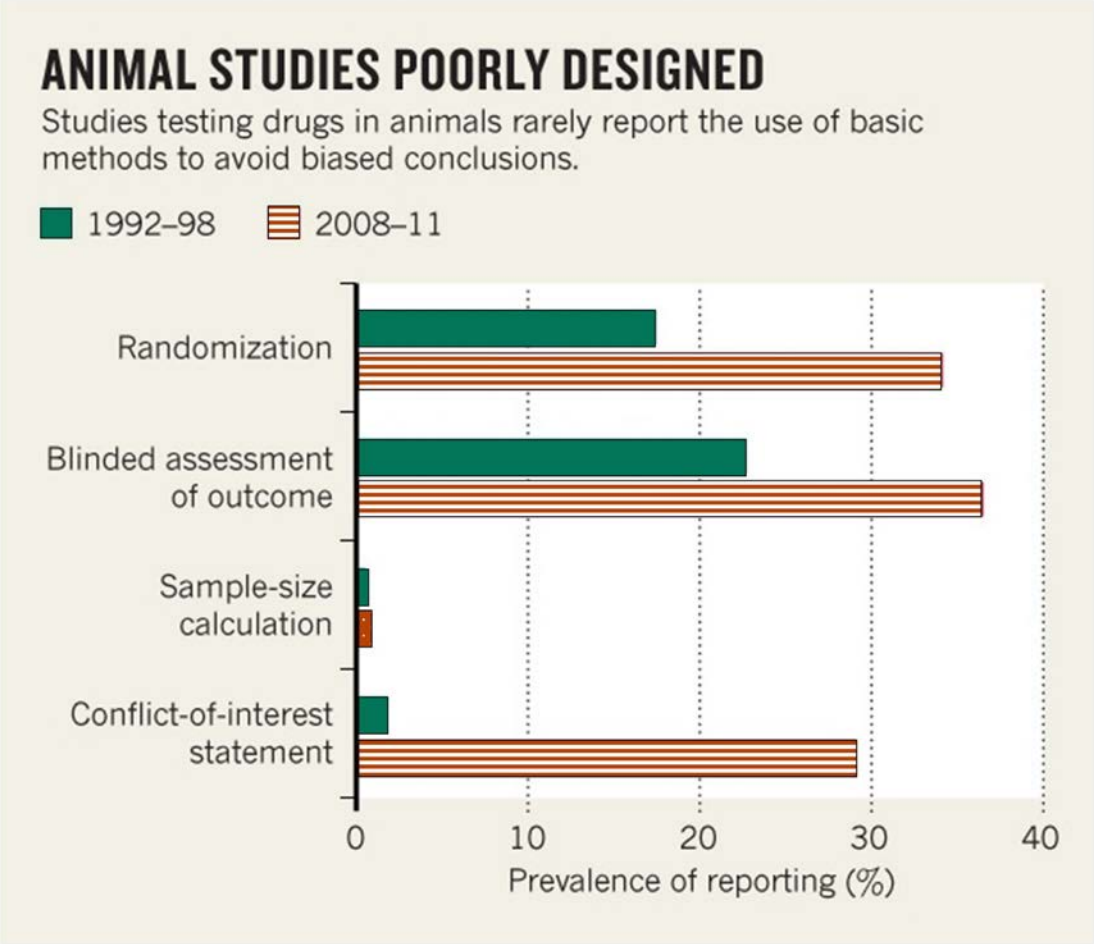
- Review Charge
- Background
- **Recommendations**

Recommendations: Five Themes



Theme 1: Improve Study Design and Analytic Rigor

Motivating Problem



Crassey D. Poorly designed animal experiments in the spotlight. Nature News. 2015 Oct 13. doi:10.1038/nature.2015.18559.

Approach

- Improve study design and data analysis to:
  - Reduce researcher bias
  - Control random variability and confounders
  - Improve reproducibility
  - Enhance *ethical* animal use
- Enable NIH to assess design quality for review and funding incentives
- Provide education, tools, and manpower to help researchers improve their designs and analyses

Theme 1: Improve Study Design and Analytic Rigor

- **Recommendation 1.1: NIH should improve and expand statistical training for animal researchers**
  - **Sub-recommendation 1.1A:** NIH should partner with other organizations to develop modern and innovative statistics curricula relevant to animal researchers.
  - **Sub-recommendation 1.1B:** NIH should develop statistical resources specifically for animal researchers.
  - **Sub-recommendation 1.1C:** NIH should require statistical training for trainees doing animal research and strongly encourage it for team members involved in study design and data analysis.

Power calculation for causal test via CRISPR genome editing (R markdown file)

ApoE x Tomm40 - Power analyses and future plans

David J Anderson September 23, 2019

Introduction: What is this analysis?

Before we conduct statistical tests for effects of genotype, age, and sex on CRISPR mouse CFA/CFM, we need to make sure these tests are well-powered. Really, this should be done at the outset of experiments to determine how many mice we need to breed, phenotype and harvest, but as far as I can tell, only rule-of-thumb estimates were used. This is understandable as I myself have had a fair bit of difficulty tracking down the requisite information to run this analysis (namely, expected effect sizes for each variable), but the following summarizes my attempts to figure out what sample sizes we'll need to feel confident moving forward with our analysis. I begin with my calculations of effect sizes for genotype, age, and sex on CFA and CFM tasks based on Neuner's papers. I then calculate the sample size necessary for each group for our analyses to be well-powered. Finally, I summarize these data and compare the sample sizes required to the number of data points we already have and the number of mice we have coming through the pipeline currently. This should serve as a roadmap for what mice to breed and phenotype going forward.

Calculation of effect sizes: I calculated expected effect sizes for genotype, age, and sex from the F-ratios and N's reported in Neuner's Neuron paper using the calculators available here. Across AD-BXD mice, there was a significant effect of Apoe allele ($F(1,354) = 4.7$; $p = 0.03$), age ($F(1,354) = 12.3$; $p = 0.001$), and sex ($F(1,354) = 17.9$; $p < 0.001$) on CFA.

IV	Variable	*F* statistic	Effect Size (*d*)
Genotype	CFA	4.7	0.226/0.85*
Age	CFA	12.3	0.283
Sex	CFA	17.9	0.348
Genotype	CFM	20.9	0.476/1.79*
Age	CFM	86.2	0.75
Sex	CFM	4.9	0.182

Calculating group size: From the table, we can see that the effect sizes of genotype and age are smaller for CFA while the effect size for sex is smaller for CFM. As CFM/CFA tests are paired (i.e., any mouse that gets scored for one will be scored for the other), and as we want to be able to test for differences in both variables, I feel that it is fair to use the *smaller* effect sizes for each variable when calculating desired group sizes. Let me know if this doesn't make sense; I'm not sure if I'm explaining my reasoning well. Running power analysis for genotype and age is easy enough if we just assume simple ANOVA models; there are ways to do this for more complex models, but I'm not well versed in them: this might be something we should talk about going forward.

Genotype power analysis: To start with, I calculated group sizes per genotype with $d = 0.226$. I'm assuming the number of groups (k) to be 12; we have *many* genotypes currently, and I'm assuming we'll ultimately want to look for differences between all of them.

Balanced one-way analysis of variance power calculation **pwr.anova.test**(k=12, f=0.226, power=0.8)

k =12
n =28.24219
f =0.226, sig.level =0.05 ,power=0.8

This calls for ~28 mice to be phenotyped in each genotype
Bred, enter pipeline (all blind until final data for all QCed)

Toward best practices as part of PhD training

Theme 1: Improve Study Design and Analytic Rigor

- **Recommendation 1.2: NIH should facilitate collaboration between statisticians and animal researchers.**
  - **Sub-recommendation 1.2A:** NIH should expand research collaborations between statisticians and animal researchers.
  - **Sub-recommendation 1.2B:** NIH should fund training for statisticians on domain-specific subject matter and on challenges faced by animal researchers.
  - **Sub-recommendation 1.2C:** NIH should increase animal researchers' access to statistical consulting through funding opportunities.
  - **Sub-recommendation 1.2D:** NIH should incentivize research in statistical methods for animal study design and analysis.

Theme 1: Improve Study Design and Analytic Rigor

Recommendation 1.3: NIH should add a single page to the research strategy for animal research that is dedicated to the description of critical elements of study design, including inclusion/exclusion criteria, sample size estimation, data analysis plan, blinding, and randomization, to reduce the risk of bias and chance observations.

This page will be additional to the current research strategy page limit and apply to vertebrate and cephalopod studies.

Elements of Rigor Proposed Research Plan Page

1. **Inclusion/exclusion criteria:** Describe the criteria that will be used for inclusion or exclusion of samples or animals during the experiments and for data used in analysis.

2. **Sample-size estimation:** Provide planned sample sizes for each group and how they were derived.

3. **Data analysis plan:** Describe plans for data analysis, including statistical methods as appropriate, designed to answer the proposed scientific questions.

4. **Blinding:** Describe measures planned to blind the investigators during group allocation, the conduct of the experiment and the analysis, where applicable. If none taken and blinding is not appropriate to the study design, provide justification.

5. **Randomization:** Describe methods planned for random allocation to comparison groups and strategies for random sample processing and collection of data where applicable. Provide a rationale if a randomization scheme is not used.

Theme 1: Improve Study Design and Analytic Rigor

- **Recommendation 1.4:** NIH should evaluate where in the pre-study research process experts could assess the quality of study design and data plans, then implement pilot studies at the most plausible stage(s) and incentivize adoption.

Theme 1: Improve Study Design and Analytic Rigor

- Candidate stages in the pre-study research process
  - ***Researcher study design stage:*** Applications could be strongly encouraged to use the NC3R's Experimental Design Assistant tool and include its flowchart in the vertebrate animal section of the application.
  - ***Grant peer review stage:*** All study panels with animal studies could include at least one trained reviewer who can evaluate the application's statistical elements, including study design and analysis plans.
  - ***Grant post-peer review stage:*** A statistical review panel composed of applied statisticians can be formed to evaluate proposals with animal studies that have received the highest scores in the previous peer review stage.

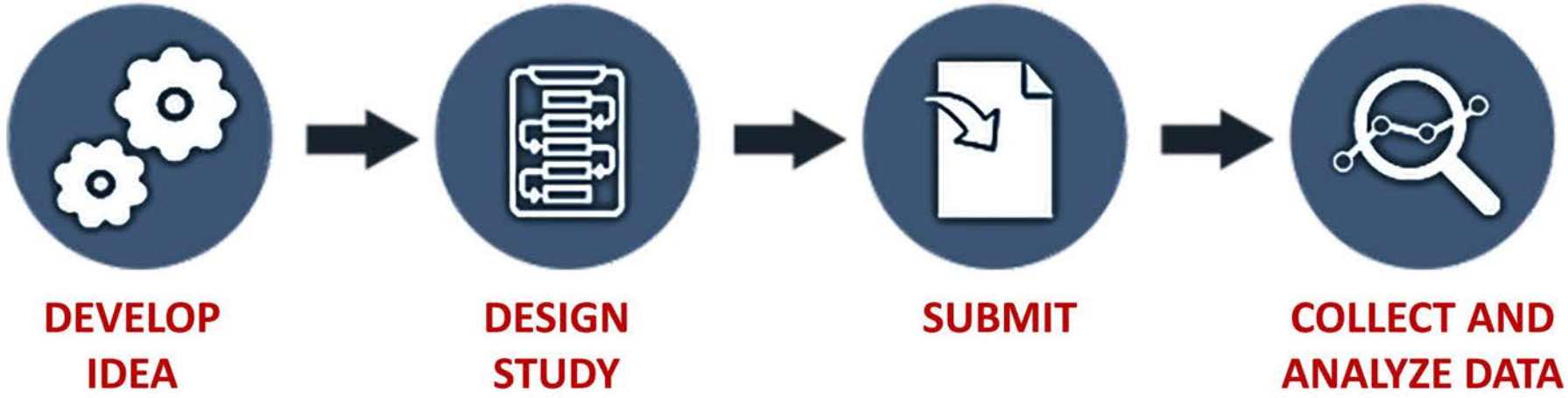
Theme 2: Address Incomplete Reporting and Questionable Research Practices

Motivating Problems

- Investigator and publishing biases: reliability of published research depends on minimizing bias
- Reproducibility depends on full reporting of research processes and outcomes

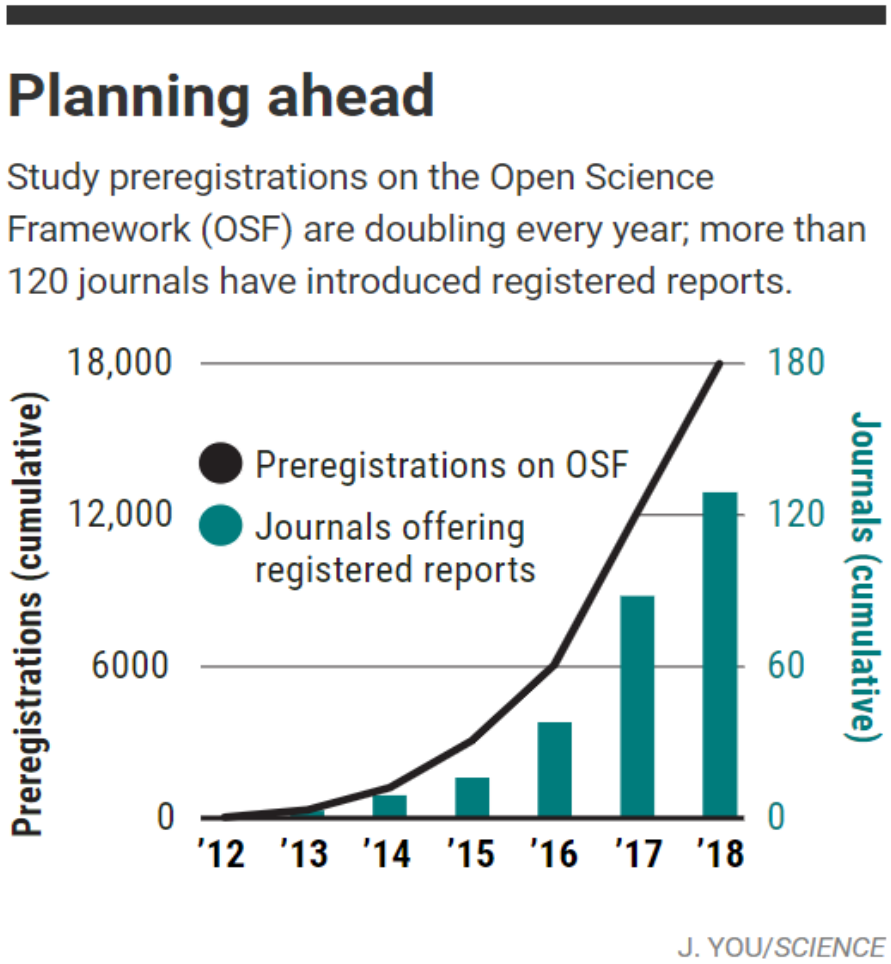
Approach

- Improve research and publication practices through complete, transparent and unbiased reporting of methods and results



Theme 2: Address Incomplete Reporting and Questionable Research Practices

- **Recommendation 2.1:** NIH should launch a campaign to raise awareness and understanding of prospectively documenting study design and analysis plans.
  - Prospective registration
  - Registered Reports



Kupferschmidt, K. More and more scientists are preregistering their studies. Should you? Science. 2018 Sept 21; doi:10.1126/science.aav4786

What Is Prospective Registration?

Prospective Registration: Creating a permanent record of a study design, analytic plan, and primary outcome before the data are collected. Prospective registration allows retrospective assurance against selective reporting and outcome switching. The registered research plan can be embargoed for a limited time to protect intellectual property, and when the registration is published, it allows to identify and mitigate publication bias.

Time-stamped, read-only version of research plan

1. Hypotheses
2. Sampling plan
3. Variables
4. Design plan
5. Analysis

https://www.cos.io/initiatives/prereg

What Are Registered Reports?

Registered Reports: Journal article type in which the detailed study protocol is submitted for peer review before the data are collected. Upon successful review, the journal guarantees publication of the article regardless of the study findings. Like prospective registration, Registered Reports mitigates selective reporting and outcome switching. In addition, the protocol peer review can help improve experimental design and mitigate experimental and analysis bias.

Traditional Publication Model

Registered Report Publication Model

https://www.cos.io/initiatives/registered-reports

Theme 2: Address Incomplete Reporting and Questionable Research Practices

- Education and awareness campaign:
  - Clear definitions of and distinction between of prospective public (or embargoed) registration and Registered Reports
  - Clear articulation of the benefits of prospective registration and Registered Reports
  - Mitigation measures (e.g., embargo periods)
  - Awareness of prospective registration as a means to bring benefit to hypothesis-testing studies, particularly focusing on in vivo studies that are intended to directly inform clinical trials

Theme 2: Address Incomplete Reporting and Questionable Research Practices

- **Recommendation 2.2:** NIH should develop and implement a pilot program to generate data on and evaluate the effects of solutions that involve the prospective documentation of study design and analysis plans in preclinical animal studies.
  - **Sub-recommendation 2.2A:** NIH should develop and incentivize projects that generate data on the Impact of prospective registration and Registered Reports.
  - **Sub-recommendation 2.2B:** NIH should set up a dedicated program to evaluate the data generated from the projects recommended in 2.2a and guide future adoption of prospective registration practices in preclinical animal studies.

Detailed Pilot Planning & Evaluation

Theme 3: Improve Selection, Design, and Translational Relevance of Animal Models

Animal studies contribute to significant findings and breakthroughs in both basic and translational research.

Motivating Problems

When intended as models of human disease, poor design or improper model selection compromises and misdirects translation.

Approach

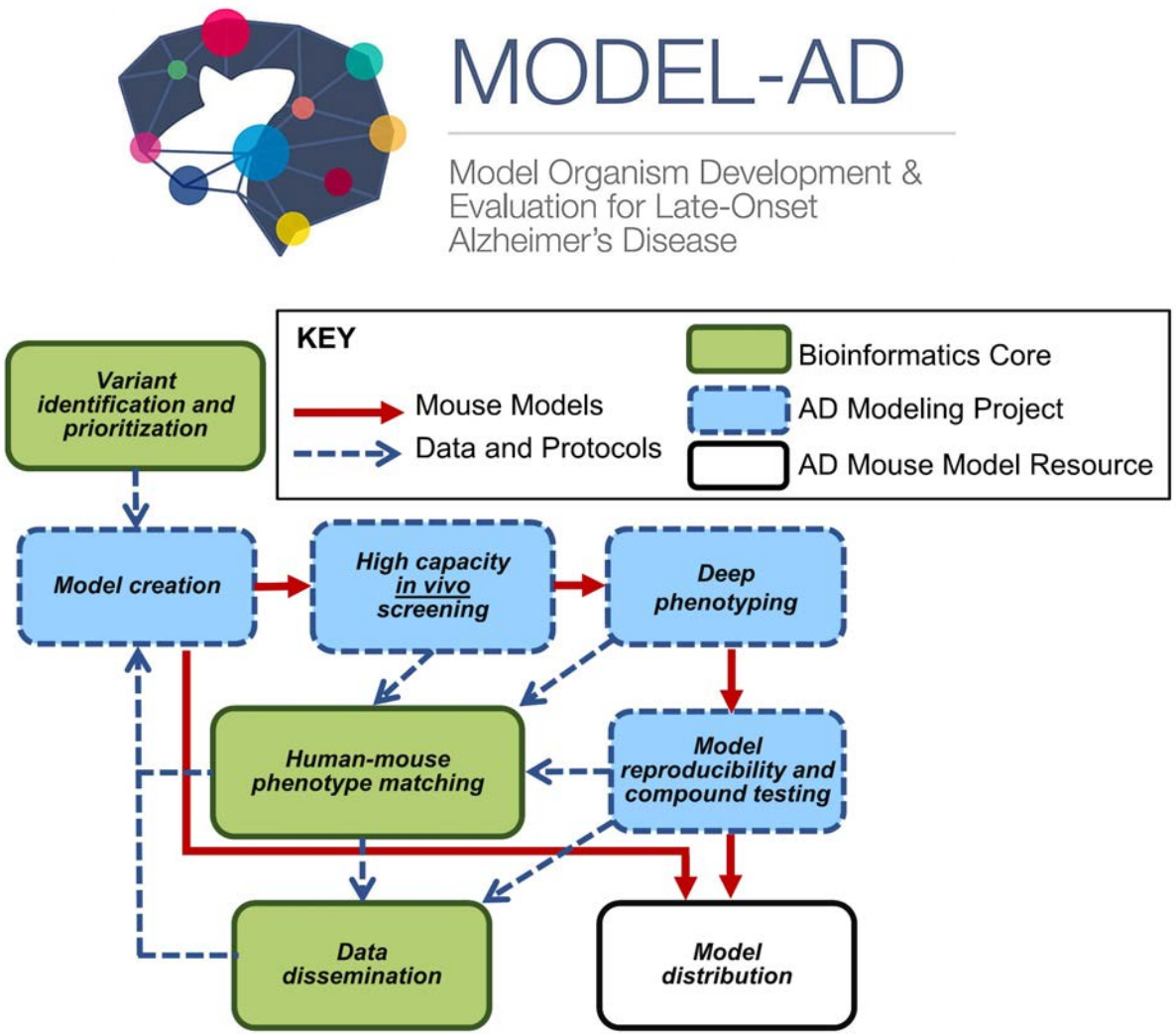
Theme 3 recommendations address how NIH can:

- Help fields to design, evaluate, and adopt better and, where applicable, more translationally relevant animal models
- Help investigators to select the most appropriate model *for a given question*
- Advance the development and appropriate uses of non-animal models

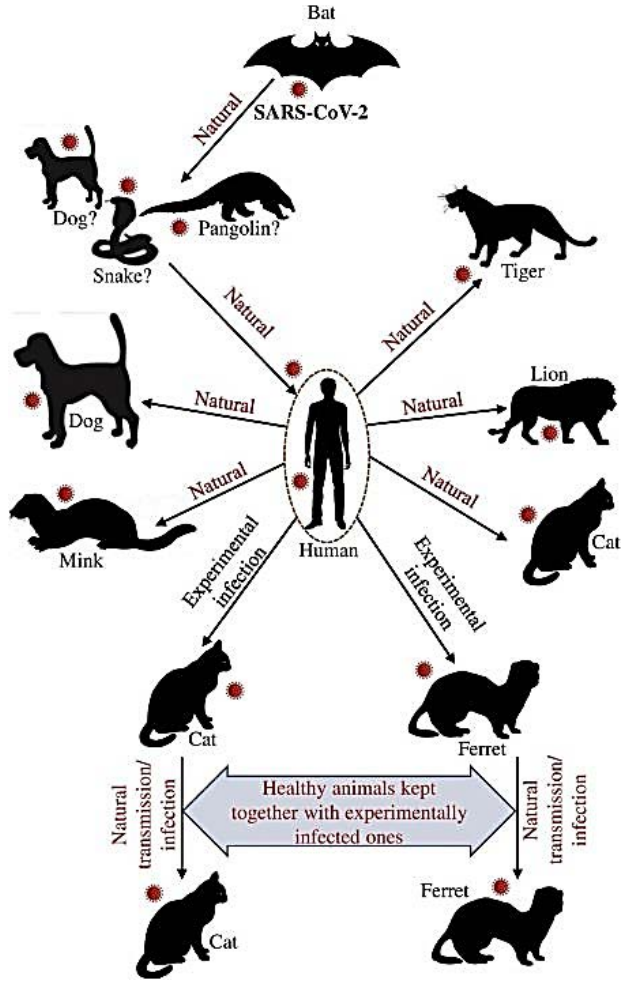
Theme 3: Improve Selection, Design, and Translational Relevance of Animal Models

- **Recommendation 3.1:** NIH should establish a framework for rationalizing the scientific and, when appropriate, translational (human) relevance of an animal model and its selection. This framework should be used to as part of the justification for animal uses in grant applications and included in ethical review processes and in journal reports.
- **Recommendation 3.2:** NIH should establish or identify venues to exchange information related to animal model design and characterization, study design, and general best practices.
- **Recommendation 3.3:** NIH should work to improve the design of animal models through the funding of focused research programs that enhance understanding of comparative human-animal biology.

Theme 3: Improve Selection, Design, and Translational Relevance of Animal Models



Host range of SARS-CoV-2 and animals susceptible to SARS-CoV-2



Theme 3: Improve Selection, Design, and Translational Relevance of Animal Models

Small Animals

ACCELERATING COVID-19 THERAPEUTIC INTERVENTIONS AND VACCINES (ACTIV)

FNIH

Species	Modification	Model Name/ Nomenclature	Vaccines	Antivirals	Neutralizing Antibodies	Other Therapies	Infectivity	Transmission	Disease Enhancement	Disease Manifestation & Pathology	Extent of Disease
Ferret	Outbred Stock	Ferret	✓		✓		Y	Y	TBD	Viral titers in nasal washes; fever	Mild
Guinea Pig	Wild Type	Guinea Pig					N	N	N	Lung lesions	None to minimal
Hamster	Inbred Strain	Syrian Golden	✓	✓	✓		TBD	Y	TBD	Lung lesions; interstitial pneumonia; recovery	Mild to moderate
Hamster	Transgenic	Tg(K18-hACE2)					TBD	TBD	TBD	TBD	TBD
Mouse	ACE2 Tranduced	Adenovirus transduced hAcE2	✓	✓	✓		Y	TBD	TBD	Lung lesions; interstitial pneumonia; weight loss; recovery	Mild
Mouse	Inbred Strain	BALB/c (adapted virus)	✓		✓		Y	TBD	TBD	Lung lesions; interstitial pneumonia; recovery	Mild
Mouse	Knock-In	C57BL/6-ACE2$^{em1(ACE2)Yowa}$	✓		✓		Y	TBD	TBD	Lung lesions; interstitial pneumonia; recovery	Mild
Mouse	Knock-In	B6.129S2(Cg)-ACE2$^{tm1(ACE2)Dwnt/J}$	✓				Y	TBD	TBD	TBD	Mild

Theme 3: Improve Selection, Design, and Translational Relevance of Animal Models

- **Recommendation 3.4:** NIH should provide adequate research support for larger and long-lived non-rodent species when justified.
  - **Sub-recommendation 3.4A:** NIH should create policy to accommodate longer time frames and higher budgets for larger and long-lived non-rodent species.
  - **Sub-recommendation 3.4B:** NIH should continue to develop national resources to produce large and long-lived animals.
- **Recommendation 3.5:** NIH should educate the public on the value of animal research, including the important roles of long-lived, non-rodent mammals for translation to improved human health and disease.
- **Recommendation 3.6:** NIH should charter a high-level working group on "non-animal modeling systems in biomedical research" to complement the activities and recommendations of this ACD working group.

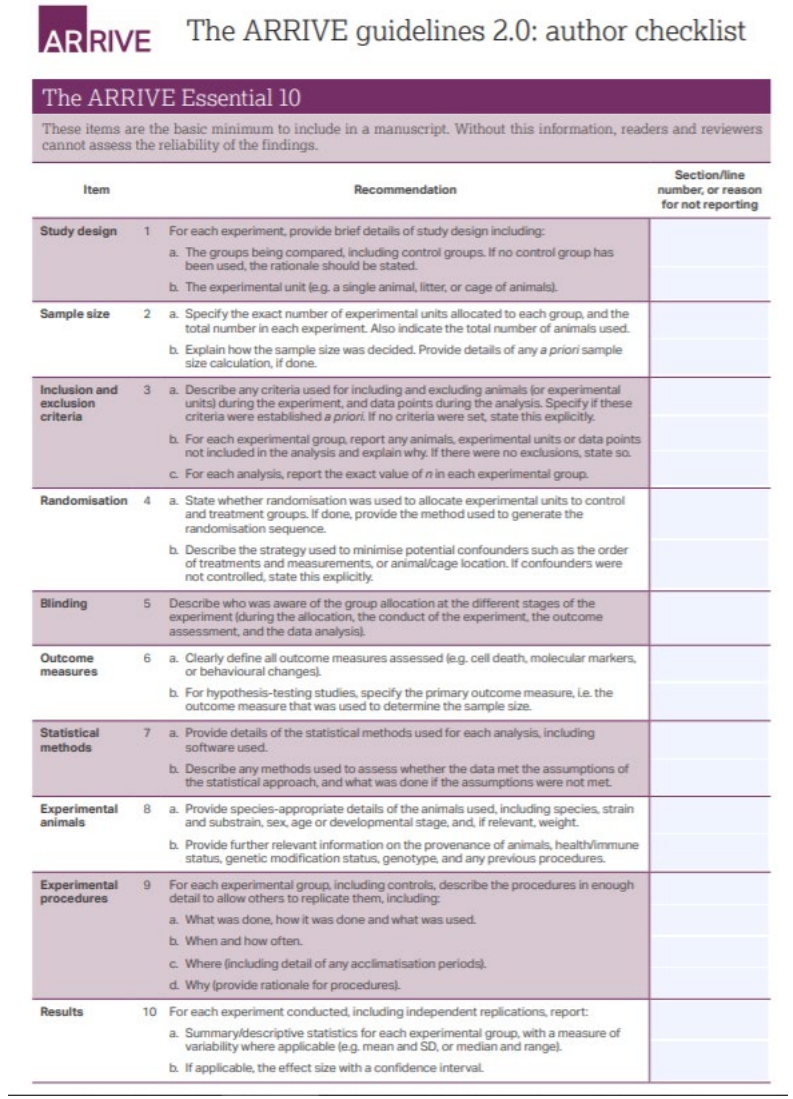
Theme 4: Improve Methodological Documentation and Results Reporting

Motivating Problems

- Transparent reporting of research methods and findings is essential, yet there are frequent failures and shortfalls.
- Failure to record and report degrades reproducibility
- Completeness and granularity of reporting for animal husbandry are a quality issue and a research topic

Theme 4: Improve Methodological Documentation and Results Reporting

- **Recommendation 4.1:** NIH should expect that key supporting data reported on animal research submitted in support of grant applications will include measures of quality and/or uncertainty for reported estimates and an interpretation of effect sizes within the context of the field.
- **Recommendation 4.2:** NIH should expect all vertebrate and cephalopod animal research to include the ARRIVE 2.0 Essential 10 at the publication stage.



ARRIVE The ARRIVE guidelines 2.0: author checklist

The ARRIVE Essential 10

These items are the basic minimum to include in a manuscript. Without this information, readers and reviewers cannot assess the reliability of the findings.

Item		Recommendation	Section/line number, or reason for not reporting
Study design	1	For each experiment, provide brief details of study design including: a. The groups being compared, including control groups. If no control group has been used, the rationale should be stated. b. The experimental unit (e.g. a single animal, litter, or cage of animals).	
Sample size	2	a. Specify the exact number of experimental units allocated to each group, and the total number in each experiment. Also indicate the total number of animals used. b. Explain how the sample size was decided. Provide details of any *a priori* sample size calculation, if done.	
Inclusion and exclusion criteria	3	a. Describe any criteria used for including and excluding animals (or experimental units) during the experiment, and data points during the analysis. Specify if these criteria were established *a priori*. If no criteria were set, state this explicitly. b. For each experimental group, report any animals, experimental units or data points not included in the analysis and explain why. If there were no exclusions, state so. c. For each analysis, report the exact value of *n* in each experimental group.	
Randomisation	4	a. State whether randomisation was used to allocate experimental units to control and treatment groups. If done, provide the method used to generate the randomisation sequence. b. Describe the strategy used to minimise potential confounders such as the order of treatments and measurements, or animal/cage location. If confounders were not controlled, state this explicitly.	
Blinding	5	Describe who was aware of the group allocation at the different stages of the experiment (during the allocation, the conduct of the experiment, the outcome assessment, and the data analysis).	
Outcome measures	6	a. Clearly define all outcome measures assessed (e.g. cell death, molecular markers, or behavioural changes). b. For hypothesis-testing studies, specify the primary outcome measure, i.e. the outcome measure that was used to determine the sample size.	
Statistical methods	7	a. Provide details of the statistical methods used for each analysis, including software used. b. Describe any methods used to assess whether the data met the assumptions of the statistical approach, and what was done if the assumptions were not met.	
Experimental animals	8	a. Provide species-appropriate details of the animals used, including species, strain and substrain, sex, age or developmental stage, and, if relevant, weight. b. Provide further relevant information on the provenance of animals, health/immune status, genetic modification status, genotype, and any previous procedures.	
Experimental procedures	9	For each experimental group, including controls, describe the procedures in enough detail to allow others to replicate them, including: a. What was done, how it was done and what was used. b. When and how often. c. Where (including detail of any acclimatisation periods). d. Why (provide rationale for procedures).	
Results	10	For each experiment conducted, including independent replications, report: a. Summary/descriptive statistics for each experimental group, with a measure of variability where applicable (e.g. mean and SD, or median and range). b. If applicable, the effect size with a confidence interval.	

Theme 4: Improve Methodological Documentation and Results Reporting

- **Recommendation 4.3:** NIH should encourage and support work to better understand, monitor, record, and report important extrinsic factors related to animal care that may affect research results.
  - **Sub-recommendation 4.3A:** NIH should provide education about the importance of extrinsic factors to the research community, provide a method to report such factors and incentivize pilot studies to further identify which extrinsic factors are impactful to reproducibility.
  - **Sub-recommendation 4.3B:** NIH should establish a task force to implement the cataloging of extrinsic factors as data from pilot studies are gathered.
  - **Sub-recommendation 4.3C:** NIH should dedicate funds for controlled randomized trials to test the effect of potentially high-value extrinsic factors identified from pilot studies and task force recommendations.

Theme 4: Improve Methodological Documentation and Results Reporting

- **Recommendation 4.4:** NIH should provide support for documenting larger and longer-lived animals' longitudinal experimental, medical, and husbandry histories.
  - **Sub-recommendation 4.4A:** NIH should formalize funding mechanisms to longitudinally record and manage animal-level experimental, medical, and husbandry history metadata for larger and longer-lived animals.
  - **Sub-recommendation 4.4B:** NIH should identify minimal animal-level experimental, medical, and husbandry history metadata that would be longitudinally recorded.
  - **Sub-recommendation 4.4C:** NIH should encourage the sharing of animal-level experimental, medical, and husbandry history.

Theme 5: Evaluate Effectiveness and Costs of Improving Rigor, Reproducibility, and Translatability

Motivating Questions

What will success look like, and what will it cost NIH and the wider research community?

Difficult-to-predict costs from changes to improve rigor, reproducibility, transparency, and translatability:

- Financial costs
- Opportunity costs
- Time costs
- Savings from reduced waste

Improved research quality and outcomes are expected, but they are not all readily quantifiable:

- Signatures of rigor in grant applications and publications
- Increased methodological and results reproducibility
- Increased competencies and workforce in statistics and analysis
- Increased success and efficiency of translation

How will NIH measure outcomes and use the results to guide midcourse corrections?

How will outcomes and evolving best practices be communicated effectively to researchers and the public?

Theme 5: Evaluate Effectiveness and Costs of Improving Rigor, Reproducibility, and Translatability

- **Recommendation 5.1:** NIH should develop an evaluation program to assess the progress in implementing the report recommendations, their effects on NIH and the research community, and challenges that arise in implementing recommendations.

Awareness

- Elevate awareness of rigor issues and opportunities in the animal research context.

Design and Early Analysis

- Propose hypotheses, gather data types, and test them. Use grant administrative data, centers of excellence, and results of animal and rigor research.

Implementation and Further Data Acquisition

- Implement incentives and requirements to achieve change.

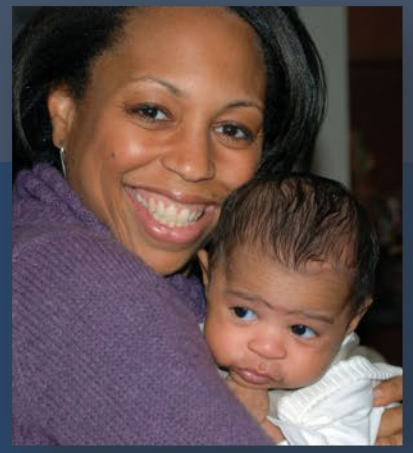
Data Evaluation and Course Adjustment

- Conduct data-driven analysis to modify, phase out, or reject as needed.

Measure costs and effectiveness of efforts to improve rigor, reproducibility, translatability

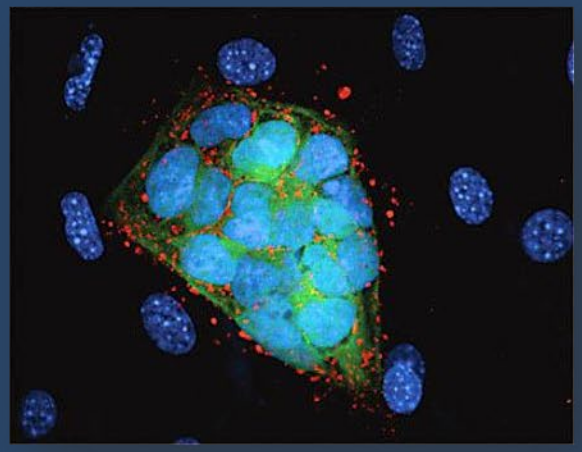
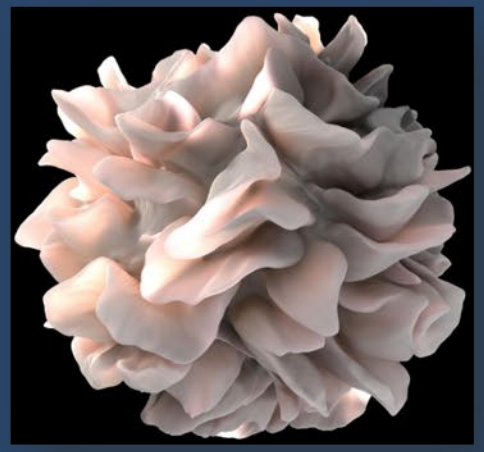
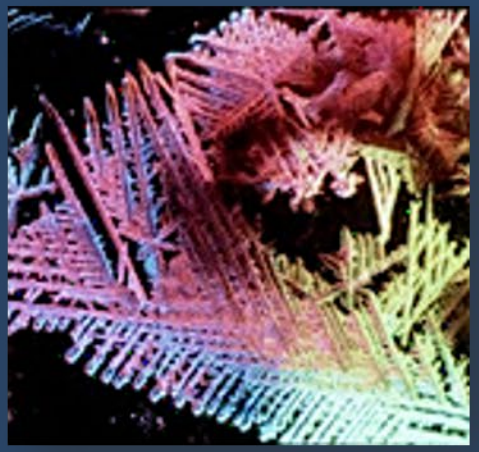
Theme 5: Evaluate Effectiveness and Costs of Improving Rigor, Reproducibility, and Translatability

- **Recommendation 5.2:** NIH should externally support and internally conduct analyses on elements of rigor and transparency in grant applications and publications to examine their financial costs, opportunity costs, and impact on portfolio balance
  - **Sub-recommendation 5.2A:** NIH should identify and collect computationally extractable information from grant proposals and reports on potentially important variables, including publication metrics, methodological rigor, funding, investigator career stage, involvement of statisticians, experimental design descriptions, and numbers and species of animals, and conduct extensive analyses on these data.
  - **Sub-recommendation 5.2B:** NIH should allow applicants to include text in the budget justification section on how projected animal budgets are linked to efforts to enhance transparency, rigor, and reproducibility.
- **Recommendation 5.3:** NIH should identify scientists who demonstrate the highest levels of transparency and rigor to help define enterprise best practices.




NIH…

Turning Discovery Into Health

Lawrence.Tabak@nih.gov