APPENDIX 1: R3PEATS Replication Landscape Analyses

Contents

APPENDIX 1: R3PEATS Replication Landscape Analyses 1
Executive Summary 2
Results of NIH-wide Survey on Ideas for a Replication Program 3
Analysis of Prior Replications 4
Landscape Analysis of Educational Initiatives 6
Scientometric and Bibliometric Indices 8
Literature Review (Conducted 2/13/2024, Updated 11/19/2025) 10
Reference List 14

Executive Summary

The Working Group conducted a landscape analysis using: 1) NIH-sourced ideas for a replication program, 2) lessons learned from prior replication efforts, 3) collected education and training resources, 4) collected metrics related to scientometrics and bibliometrics, and 5) a literature review.

1) Results from an NIH survey found an interest in validating methods, incorporating existing big databases into research studies, emphasizing reporting of null results, and generalizing results. Responses indicated that a program should focus on subjects such as: clinical/human subjects, basic research, methods and research culture, and policy.
2) Replicability requires rigorously performed research and transparent reporting. Exact direct replications are impossible, so replication research is needed to define how strict one needs to adhere to the previous protocol and how much variability should be built in to determine robustness. Replications are time-consuming and expensive, and it is unclear which experimental factors play important roles in replicability.
3) Many resources and standards on replication are available but are dispersed and difficult to navigate. Efforts are needed to encourage replicability through ongoing training, pro-social data sharing, transparent reporting, and incentivization approaches. The effectiveness of existing policies needs to be assessed and expanded. Enhanced public engagement would build community trust in the role of replicability in improved health outcomes.
4) A handful of existing indices measure critical features of rigor, transparency, and scientific impact at different scales (e.g. journals vs researchers). None measure or predict replicability well.
5) Failure of clinical research to translate effectively into the clinic has been attributed to the lack of rigor and replicability in preclinical research, the complexity of human diseases, and the logistics of therapeutic development. Recommendations to enhance replicability include detailed reporting, standardized protocols, data sharing, and the use of technology-based approaches such as software for study design, randomization and data collection, and analysis workflows. The scientific community's historically limited compliance with these recommendations and the need for incentives or requirements to align researcher behavior remain significant challenges.

Results of NIH-wide Survey on Ideas for a Replication Program

Becky Krupenevich, NIA

Response:

- Wide response. 15 ICOs represented across 30 responses. Some were general, others suggested specific foci. Some proposed specific studies, others standalone initiatives.

Major Ideas Represented:

- Methods for enhancing rigor and reproducibility standards
- Bioinformatics tools and informational models may lack reproducibility
- Big data, secondary analysis, and meta-analysis
    - Usage of existing data sources for predictive models
    - Real world data to replicate and validate critical findings from clinical trials (e.g., for drugs, preventive health screening)
    - Synthetic data/cohorts for rapid, scalable replications
- Emphasis on null studies
- Responses thought a program should address: clinical/human subjects, basic research, methods and research culture, policy

Specific Studies Proposed for Replication:

- Cocoa Supplement and Multivitamin Outcomes Study finding that daily multivitamin/mineral supplementation for two years significantly improved episodic memory in older adults
- Validation of the THX humanized mouse model that used estrogen to support immune cell differentiation
- Validation of genotype contributions to drug response and toxicity (e.g., replication of findings such as the DPYD HapB3 haplotype paper).

Specific Topics Proposed for Replication:

- SSRI effects on maternal and fetal health
- Confirm the causal role of methylation changes in autism spectrum disorder (ASD)
- HrPQCT for as biomarker for fracture risk in individuals with rare bone diseases
- Microplastics assessments

Suggested Broad Areas for Replication:

- Digital health interventions – replication across versions of tech (e.g., sensors)
- Human-based biomarkers
- Disease-mechanistic studies that implicate a specific cellular pathway or molecular target
- fMRI - strong need for methodological rigor
- Psychedelic research
- Behavior change interventions e.g., physical activity, diet

Open Questions:

- How might some of the responses fit into replication center themes?
- Is generalizability in scope? (e.g., replication across populations or contexts)

Analysis of Prior Replications

Devon Crawford & Mariah Hoye, NINDS

Approach:

- Analysis of a selection of 20 papers from many fields and methods [1-20]. Papers included meta-analyses of interventions, large replication project reports, and commentary on lessons learned. Refer to the bibliography for a list of references.

Major Takeaways:

- Good replicability can be achieved.
  - Needs transparent and comprehensive reporting.
  - Small design choices affect replicability (e.g., dose/timing, equipment parameters, analysis approaches).
  - Success linked to using standardized methods/materials, strong bias prevention (e.g., blinding/quality control/randomization/data management/power), and multi-site designs.
- There is no such thing as an exact direct replication.
  - Even with standardization there will be differences from the original study (e.g. reagent/animal stocks, local IACUC rules, experimenter skill, unknown confounding factors, etc.).
  - Strict adherence to protocol may limit ability to test robustness of result.
  - It is not always clear which changes in protocol contribute to differences in outcome.
- Not all issues with replicability can be solved.
  - Even in the same study, some tests/outcomes and models may be replicable, and some may not, due to chance, inherent variability, biological differences, and/or context.
- Replications are expensive and time-consuming.
  - Projects often need extra time to work out protocol tweaks and to reach sufficient sample sizes.
  - Data management can be a challenge, and long-term data storage should be considered.
- There is a gap in knowledge about how cultural practices affect replicability.
  - There is evidence that open science interventions improve transparency, but less is known about how policies, incentives, resources, etc. directly impact replicability of research findings.

Open Questions:

1. Where should the dividing line(s) be between the definitions of direct replication, robustness check, conceptual replication, and generalization?
   - Do measures of test-to-test, sample-to-sample, and lab-to-lab variability factor into definitions?
   - How should investigators approach missing protocol information?
   - If strict standardization isn't possible, when does this mean that the experiment should not be replicated?

2. How are we planning to define a successful replication across all pieces of R3PEATS?
    - Similar effect size in the same direction? Statistical threshold? Measures of variability? Will this be defined by the Coordinating Center? Individual studies/replication centers?
    - How does this definition interact with the "replicability index" (i.e., is this index a human behavior signifier or a characteristic of a specific experiment/result) being developed by the Metascience Testbed?
3. What kinds of replications does this initiative want to prioritize?
    - Experiments with translational implications? Widely contested findings? Method validation?

Landscape Analysis of Educational Initiatives

Tiffany Cook, NEI

Approach:

Research: Review of NIH-, web-, and publication-based education, training, and reporting resources for rigor, reproducibility and replication studies

Public: Review of resources on public trust of biomedical research and clinical outcomes

Challenges:

Research rigor and reproducibility compliance:

- Gaps in training, mentorship, expertise, oversight across the research community (trainees, PIs, institutions, journals, editors, reviewers, media)
- Misaligned incentives (e.g., publish-or-perish) tied to funding, reputation, and promotion
- Limited standardization of reagents, protocols, study design, analysis, and reporting
- Cost and access barriers to experimental/analytical tools, raw data, and research materials
- Inconsistent record-keeping, data management, interpretation, and reporting practices
- Questionable research practices, citation inaccuracies, and selective or incomplete reporting

Public trust:

- Imbalanced information sources and coverage
- Limited communication and accessible, plain-language summaries of biomedical research findings

Major Takeaways and Recent Advances in Replication Research:

Research: Numerous and varied resources and opportunities exist inside and outside NIH, but they are dispersed and their effectiveness is difficult to evaluate

Examples:

- NIH contributions:
  - Funding initiatives (e.g., CENTER & METER at NINDS)
  - Competitions and prizes (e.g. S-index, Replication)
  - Training material and links to external resource
  - Grant requirements (e.g., Rigor & Reproducibility, DMSP, Authentication, HSS, RCR)
- Journals: Strengthened submission and reporting requirements
- Guidelines: Clearer, centralized resources across multiple biomedical domains
- Training: Multiple online options in statistics and methods (free/paid, videos/slides, programming language-based)

Public: Growing use of social media to explore health issues and contested scientific findings

- Heightened awareness within the scientific community of the need to bridge communication gaps with public

Opportunities

- Centralize and spotlight NIH-supported resources and policies
- Assess and track compliance of policies
- Strengthen R3 mentoring and education (e.g. standardized training for trainees, established investigators, peer reviewers, and the public)
- Align incentives to promote research rigor and impact
- Expand data sharing and detailed method reporting (including public data sharing portals)
- Promote stronger journal oversight, submission standards, pre-registration, and reporting of negative/null studies
- Increase meaningful public engagement

Open Questions

- How can we improve and coordinate available training resources?
- How can researchers, universities, journals develop, standardize, and adopt new/existing solutions for sustained culture change?
- How can we foster public trust in biomedical research?

Scientometric and Bibliometric Indices

Alicia Cavanaugh, OD

- Existing indices measure critical features of rigor and replicability efforts such as *internal validity* (e.g. statistical power and bias), *rigor and transparency* in reporting research elements that enable future replication, and *impact* of publications and researcher productivity (see Table 1). An upcoming metric will measure how effective a researcher is at *sharing* data.
- *Reliability*. The Z-curve is a meta-analysis metric that measures the statistical power of published studies, the likely number of missing studies, and the amount of bias in published studies [21]. It is intended to measure the reliability of large sets of studies with heterogenous design with a common element (e.g., same journal, program, year).
- *Rigor and Transparency*. There are 3 indices focused on ensuring the reporting of elements needed for replication. (1) SciScore is an index that identifies how many elements of rigor and transparency are included in a paper [22]. The average SciScore is the (2) Rigor and Transparency Index for a given group of papers [23, 24]. These elements are specific to biological sciences. Similarly, while not a scored index, (3) AlzPED's Rigor Report Card is a framework that identifies whether a paper has 24 recommended and 9 critical elements of rigor and reproducibility [25].
- *Impact*. The *h*-index measures the productivity and citation impact of publications and individuals [26]. Each database produces a different *h-index*.
- *Sharing*. The forthcoming S-index will measure how effective a researcher is at sharing data in a way that will be useful for future study [27].

Table 1: Scientometric and Bibliometric Indices Description

Name	Purpose	Measurement	Notes
Z-Curve	Meta-analysis of many papers (100+) to determine: 1)The **statistical power** of published studies 2)The likely number of missing studies 3)The amount of **bias**	Models the distribution of z-scores derived from p-values of published studies.	Intended for large sets of studies with heterogenous hypothesis and design, with a common element (e.g., same journal, same program)
SciScore	Identify whether a paper has elements associated with increased **Rigor and Transparency**	Using a 10-point scale, points are given for the presence of Rigor (5 pts) and Transparency in Biological Resources Used (5 pts) features.	Identifies whether these elements are present, not indicative of quality.

Rigor and Transparency Index (RTI)	Identify whether a group of papers has elements associated with increased **Rigor and Transparency**.	The mean SciScore for a group of papers.	Elements are specific to **biological sciences.** Used to compare groups, or a single group over time.
AlzPED's Rigor Report Card	Identify whether a paper has 24 recommended experimental design elements, and 9 elements that are critical for scientific **rigor and reproducibility**	Recommended elements: *see list here* Core elements: *Power/Sample Size Calculation, Inclusion/Exclusion Criteria, Blinded for Treatment, Study Balanced for Sex, Randomized into Groups, Blinded for Outcome Measures, Genetic Background, Conflict of Interest, Sex as a Biological Variable*	Not an index/score, but a framework to enhance rigor, reproducibility, and translatability
h-index	Measures the **productivity and citation impact** of publications, and individuals	The maximum value of *h* such that the given author/journal has published at least *h* papers that have each been cited at least *h* times	Each database produces a different h for the same scholar due to coverage. First developed for individuals, then applied to journals. A number of indices have been developed to complement the h-index [28].
S-Index	Forthcoming index on how effective a researcher is at **sharing data in a way that has utility for future study**	TBD	Developing through NEI's Data Sharing Index Challenge. Inspired by h-index.

Literature Review (Conducted 2/13/2024, Updated 11/19/2025)

Brionna Hair and Alicia Cavanaugh, OD

Key Takeaways:

- The failure of promising clinical research to translate successfully in the clinic is a result of not only the lack of reproducibility and rigor in preclinical research, but also because of the complexity of human disease and the inability for animals to accurately model complex human diseases [29, 30]
    - Improving research requires improving both rigor/reproducibility and translatability [29]
- Many recommendations to improve rigor and reproducibility focus on more detailed reporting of research, including standardized protocols for experimentation and reporting, preregistration of research, and guidelines by journals/professional societies to improve reporting of animal research [29, 31-33]
- Data sharing is also a recognized approach to improve rigor and reproducibility, with many entities suggesting or requiring data be deposited in repositories [29]
- However, the potential for these recommendations to improve rigor and reproducibility is limited by lack of compliance by the scientific community, including limited compliance with reporting guidelines endorsed by journals and other entities [33]; and researchers not depositing full primary data into repositories [31].
    - Incentives or requirements may be required to align researcher behavior with actions needed to improve rigor and reproducibility [34]
- Recommendations for funding agencies to improve rigor and reproducibility include education and training initiatives, conducting replication studies of published work, collecting information on indicators of rigor in funded research, requiring discussion of reproducibility issues in grant applications, adding to the review process criteria related to reproducibility, and funding technology that can improve reproducibility [34, 35]
- Several technology-based approaches to improving the rigor and translatability of research [36], include:
    - To improve poorly designed studies, use of software programs that provide unbiased guidance on study design and analysis
    - To improve the characterization and randomization of animals, use of digital technologies that can automate the collection of baseline data and assist with block and stratified randomization
    - To improve data sharing, use of Internet of Things sensors to facilitate the automatic capture of study data and experimental conditions
    - To reduce experimenter bias, use of biomonitoring technologies that collect digital biomarkers
    - To improve the measurement of functional outcomes in animal studies (and therefore improve translatability to clinical studies), use of biomonitoring, and
    - To improve translatability of research, use of artificial intelligence to create digital markers that can be standardized between preclinical and clinical studies

Replication Literature Notes:

- There is no universal definition of reproducibility or replication (see Table 2)

Table 2. Categories and Usage of definitions on Reproducibility and Replication [34, 37-39]

	No Distinction	Distinction 1	Distinction 2
Reproducibility	Reproducibility is used as an umbrella term.	Original data & methods to reach same findings	New data & methods to reach same findings.
Replicability	Reproducibility is used as an umbrella term.	New data & methods to reach same findings.	Original data & methods to reach same findings
Government Organization Usage	2014 NIH Plan EPA	NSF Subcommittee on Replicability in Science NASEM	2025 NIH Gold Standard Science

Given the variability of definitions, the following notes use the terms that were used by the authors.

- Rigor and reproducibility vs translatability
    - The failure of promising preclinical research to translate in the clinic is multifactorial and includes the complexity of chronic human disease, the inability of animals to accurately model these complex diseases, and the irreproducibility of basic research [30]
    - Work can be rigorous and reproducible and still not be translatable from animal models to human; improving research requires improving both rigor/reproducibility and translatability [29]
- Failure to reproduce a computational finding can occur because of a lack of transparency, particularly with regards to the code and data used [34]
- Failure to reproduce scientific findings can occur because of:
    - False positives (by chance or because of bias in the study design) [33]
    - Methodological errors [34]
    - Publication bias [33]
    - An unknown variable that influences the phenomenon being studied [33]
- Recommendations for improving rigor and reproducibility:
    - Detailed reporting of research and data sharing
        - Detailed reporting of research methods is necessary for the replication of research, but also to assess the rigor of research. Recommendations include:
            - Standardized protocols for experimentation and reporting (especially at design stage of research) [29, 33]
            - Preregistration (templates that request information about experimental design, samples, analyses) of research before data collection begins and that can be evaluated [29, 33]
        - Data sharing through the use of repositories [29, 34]
        - Actions by journals to promote transparency, rigor, and reproducibility

        - Many journals have guidelines to promote transparency in the reporting of animal research (recently many journals require data be submitted to a repository) [32]
        - However, compliance with following guidelines needs to be assessed; additionally, guidelines may need to be tested to ensure researchers understand what they are being asked to report on [33] and better incentives or requirements by journals are recommended [34]
    - Training
        - Institutions should include training on the use of statistical analysis and inference [34]
    - Recommendations for funding agencies
        - Education and training initiatives to ensure researchers can conduct rigorous work and describe methods clearly [34]
        - Conduct tests of the reproducibility of published research; conduct replication studies [34]
        - Require discussion on grant applications on issues relevant to reproducibility and replicability and how uncertainties will be evaluated; review criteria should include review of reproducibility and replicability guidelines [34]
        - Collect information on indicators of rigor to assess funded research projects [35]
        - Determine whether current policies are adequate to achieve transparency [35]
    - Technological solutions
        - Funding agencies should invest in research and development of open-source tools and infrastructure that support computational reproducibility [34]
        - To address the challenge of poorly designed studies, use of software programs that can provide unbiased guidance to facilitate study design and analysis.
        - To address the challenge of poor characterization and randomization of animals into treatment groups, there are digital technologies that can facilitate automated collection of baseline data and enhance data quality. Software tools are available to assist with block and stratified randomization [36]
        - To address challenges involving data sharing, the use of “Internet of Things” sensors and cloud can enhance data collection and sharing, including the automatic capture of study data and experimental conditions [36]
        - To address challenge of experimenter bias, biomonitoring technologies that collect digital biomarkers (limiting the bias associated with subjective measurement) [36]
        - Biomonitoring can also facilitate the measurement of functional outcomes in animal studies (rather than biomarkers or histology) [36]
        - Use of AI to create “translationally relevant digital biomarkers that can be standardized between preclinical and clinical studies” [36]

- Implementation as a challenge
    - An important barrier to improving rigor and reproducibility is lack of compliance with recommendations
        - In the field of pharmacology, widely endorsed recommendations to improve translatability of animal studies (including use of rigorous research plans and training in use of statistics and performance or requisite analyses) were not being adopted by research community [31]
        - Lack of compliance also an issue with data repositories (researchers are not depositing the full primary data online) [31]
- Nonhuman primate research (special consideration may need to be made for primate research)
    - There are similarities between nonhuman primates and humans that make them good models for the study of human disease [29]
    - Ethical constraints for animal research, including replacing animals in research and reducing sample size, are difficult to balance with rigor and reproducibility needs (including increased sample size to improve rigor and reproducibility) [29]
    - Effective use of pilot grant programs and data collaboration/sharing may be particularly important for nonhuman primate research [29]
- NIH activities
    - 2012: NINDS convened meeting to develop recommendations to improve preclinical research [32]
    - 2014: NIH held workshop with Nature and Science to discuss rigor and reproducibility in research articles [32]
    - 2016: NIH released Rigor and Reproducibility policy and guidelines for [32]
    - 2019: NIH released instructions on rigor and reproducibility to training grant applicants [43]
    - 2019: NIH launched Stroke Preclinical Assessment Network which includes replication activities to demonstrate feasibility of network model and interventions [44].
    - 2022: NINDS launched the Initiative to Improve Education in the Principles of Rigorous Research to improve training in rigorous research [45].
    - 2022: NIH launched the HEAL Initiative which included NINDS' Preclinical Screening Platform for Pain which tests the efficacy of pain therapies [46].
    - 2023: NINDS launched the Rigor Champions Prize which promote rigor and transparency practices and cultivate a culture that promotes robust neuroscience [47].
    - 2024: NIH Common Fund launched the Replication to Enhance Research Impact Initiative to provide support for the replication of NIH-funded research and validate novel tools, technologies, and methods [48].
    - 2025: Autism Data Science Initiative funds studies that explore the causes of autism, and how treatments and interventions are used. ADSI will fund independent replication and validation of ADSI findings [49].

Reference List

1. Ashoor, I., et al., *Standardization and Cross Validation of Alloreactive IFNγ ELISPOT Assays Within the Clinical Trials in Organ Transplantation Consortium*. American Journal of Transplantation, 2013. **13**(7): p. 1871-1879.
2. Dudda, L., et al., *Open science interventions to improve reproducibility and replicability of research: a scoping review*. Royal Society Open Science, 2025. **12**(4): p. 242057.
3. Errington, T.M., et al., *Challenges for assessing replicability in preclinical cancer biology*. eLife, 2021. **10**: p. e67995.
4. Gould, E., et al., *Same data, different analysts: variation in effect sizes due to analytical decisions in ecology and evolutionary biology*. BMC Biology, 2025. **23**(1): p. 35.
5. Laboratory, I.B., et al., *Reproducibility of in vivo electrophysiological measurements in mice*. eLife, 2025. **13**: p. RP100840.
6. Keslar, K.S., et al., *Multicenter evaluation of a standardized protocol for noninvasive gene expression profiling*. American Journal of Transplantation, 2013. **13**(7): p. 1891-1897.
7. Lemaitre, J., et al., *A retrospective analysis of 400 publications reveals patterns of irreproducibility across an entire life sciences research field*. bioRxiv, 2025: p. 2025.07.07.663460.
8. Lu, O.D., et al., *A multi-institutional investigation of psilocybin's effects on mouse behavior*. bioRxiv, 2025.
9. Lyden, P.D., et al., *A multi-laboratory preclinical trial in rodents to assess treatment candidates for acute ischemic stroke*. Science Translational Medicine, 2023. **15**(714): p. eadg8656.
10. Niepel, M., et al., *A Multi-center Study on the Reproducibility of Drug-Response Assays in Mammalian Cell Lines*. Cell Systems, 2019. **9**(1): p. 35-48.e5.
11. Prinz, F., T. Schlange, and K. Asadullah, *Believe it or not: how much can we rely on published data on potential drug targets?* Nature Reviews Drug Discovery, 2011. **10**(9): p. 712-712.
12. Radabaugh, H.L., et al., *Translational Outcomes Project in Neurotrauma (TOP-NT) Pre-Clinical Consortium Study: A Synopsis*. Journal of Neurotrauma, 2025. **42**(9-10): p. 898-912.
13. Reed, E.F., et al., *Comprehensive Assessment and Standardization of Solid Phase Multiplex-Bead Arrays for the Detection of Antibodies to HLA*. American Journal of Transplantation, 2013. **13**(7): p. 1859-1870.

14. Rychert, J., et al., *Multicenter comparison of laboratory performance in cytomegalovirus and Epstein–Barr virus viral load testing using international standards*. Clinical Transplantation, 2014. **28**(12): p. 1416-1423.

15. Scott, S., et al., *Design, power, and interpretation of studies in the standard murine model of ALS*. Amyotroph Lateral Scler, 2008. **9**(1): p. 4-15.

16. Steward, O., et al., *Replication and reproducibility in spinal cord injury research*. Experimental Neurology, 2012. **233**(2): p. 597-605.

17. The Brazilian Reproducibility, I., et al., *Estimating the replicability of Brazilian biomedical science*. bioRxiv, 2025: p. 2025.04.02.645026.

18. Voelkl, B., et al., *Reproducibility of preclinical animal research improves with heterogeneity of study samples*. PLOS Biology, 2018. **16**(2): p. e2003693.

19. von Kortzfleisch, V.T., et al., *Do multiple experimenters improve the reproducibility of animal studies?* PLOS Biology, 2022. **20**(5): p. e3001564.

20. Westlake, H., et al., *Reproducibility of Scientific Claims in Drosophila Immunity: A Retrospective Analysis of 400 Publications*. bioRxiv, 2025: p. 2025.07.07.663442.

21. Schimmack, U. *Z-Curve.3.0 Tutorial*. 2025; Available from: https://replicationindex.com/2025/07/08/z-curve-tutorial-introduction/.

22. Bandrowski, A. and M. Roelandse. *SciScore, a tool that can measure rigor criteria presence or absence in a biomedical study*. in *RExPO22*. 2022. ScienceOpen.

23. Menke, J., et al., *Establishing institutional scores with the rigor and transparency index: large-scale analysis of scientific reporting quality*. Journal of Medical Internet Research, 2022. **24**(6): p. e37324.

24. Menke, J., et al., *The rigor and transparency index quality metric for assessing biological and medical science methods*. Iscience, 2020. **23**(11).

25. (NIA), N.I.o.A. *AlzPED Analytics*. 2025; Available from: https://alzped.nia.nih.gov/alzped-analytics.

26. Hirsch, J.E., *An index to quantify an individual's scientific research output*. Proceedings of the National academy of Sciences, 2005. **102**(46): p. 16569-16572.

27. (NIAID), N.I.o.A.a.I.D. *NIH Data Sharing Index (S-index) Challenge*. 2025; Available from: https://www.challenge.gov/?challenge=nih-data-sharing-index-s-index-challenge.

28. Bihari, A., S. Tripathi, and A. Deepak, *A review on h-index and its alternative indices*. Journal of Information Science, 2023. **49**(3): p. 624-665.

29. Bliss-Moreau, E., et al., *Improving rigor and reproducibility in nonhuman primate research*. Am J Primatol, 2021. **83**(12): p. e23331.

30. Zeiss, C.J. and L.K. Johnson, *Bridging the Gap between Reproducibility and Translation: Data Resources and Approaches*. Ilar j, 2017. **58**(1): p. 1-3.

31. Begley, C.G. and J.P.A. Ioannidis, *Reproducibility in Science*. Circulation Research, 2015. **116**(1): p. 116-126.

32. Lu, H.S. and A. Daugherty, *Key Factors for Improving Rigor and Reproducibility: Guidelines, Peer Reviews, and Journal Technical Reviews*. Front Cardiovasc Med, 2022. **9**: p. 856102.

33. Macleod, M. and S. Mohan, *Reproducibility and Rigor in Animal-Based Research*. Ilar j, 2019. **60**(1): p. 17-23.

34. National Academies of Sciences, E., and Medicine, et al., *Understanding Reproducibility and Replicability*, in *Reproducibility and Replicability in Science*. 2019, National Academies Press (US)

Copyright 2019 by the National Academy of Sciences. All rights reserved.: Washington (DC).

35. GAO), U.S.G.A.O.U.S., *Research Reliability: Federal Actions Needed to Promote Stronger Research Practices*. 2022.

36. Defensor, E.B., M.A. Lim, and L.R. Schaevitz, *Biomonitoring and Digital Data Technology as an Opportunity for Enhancing Animal Study Translation*. Ilar j, 2021. **62**(1-2): p. 223-231.

37. Barba, L.A., *Terminologies for reproducible research*. arXiv preprint arXiv:1802.03311, 2018.

38. Plesser, H.E., *Reproducibility vs. Replicability: A Brief History of a Confused Terminology*. Frontiers in Neuroinformatics, 2018. **Volume 11 - 2017**.

39. Antunes, B. and D.R.C. Hill, *Reproducibility, Replicability and Repeatability: A survey of reproducible research with a focus on high performance computing*. Computer Science Review, 2024. **53**: p. 100655.

40. Goodman, S.N., D. Fanelli, and J.P. Ioannidis, *What does research reproducibility mean?* Science translational medicine, 2016. **8**(341): p. 341ps12-341ps12.

41. Schmidt, S., *Shall we really do it again? The powerful concept of replication is neglected in the social sciences*. Review of general psychology, 2009. **13**(2): p. 90-100.

42. Force, A.D.R.T., *How Can Scientists Enhance Rigor in Conducting Basic Research and Reporting Research Results?* 2015, American Society for Cell Biology (ASCB).

43. LaPolla, F.W.Z., et al., *Rigor and reproducibility instruction in academic medical libraries*. J Med Libr Assoc, 2022. **110**(3): p. 281-293.

44. (SPAN), S.P.A.N. *Stroke Preclinical Assessment Network*. 2024; Available from: https://spannetwork.org/.

45. (NINDS), N.I.o.N.D.a.S. *Initiative to Improve Education in the Principles of Rigorous Research*. 2025; Available from: https://www.ninds.nih.gov/current-research/trans-agency-activities/rigor-transparency/initiative-improve-education-principles-rigorous-research.

46. (NINDS), N.I.o.N.D.a.S. *Preclinical Screening Platform for Pain (PSPP)*. 2025; Available from: https://www.ninds.nih.gov/current-research/trans-agency-activities/ninds-role-heal-initiative/preclinical-screening-platform-pain-pspp.

47. Stroke, N.I.o.N.D.a. *NINDS Rigor Champions Prize*. 2025; Available from: https://www.ninds.nih.gov/current-research/trans-agency-activities/rigor-transparency/rigor-champions-and-opportunities/ninds-rigor-champions-prize.

48. Fund, N.C. *Replication to Enhance Research Impact Initiative*. 2025; Available from: https://commonfund.nih.gov/replication-initiative.

49. Division of Program Coordination, P.a.S.I.D. *Autism Data Science Initiative*. 2025; Available from: https://dpcpsi.nih.gov/autism-data-science-initiative.