

Limited Reproducibility of Research Findings: Implications for the Welfare of Research Participants and Considerations for Institutional Review Boards

In a number of biomedical science fields, a lack of reproducibility of research results has caused alarm about wasted resources, both human and financial.¹ The inability to replicate findings has significant implications not only for the reliability of science but also for research subjects. A problem that has not drawn sufficient attention is that flawed irreproducible science can also have a negative impact on the welfare of prospective research participants by interfering with risk minimization and risk-benefit comparison, especially when participants are from vulnerable populations.

“Reproducibility” is an umbrella term covering repeatability (the question of whether repeating an original study yields the same findings), replicability (whether different data sets and their accumulation in meta-analyses by independent investigators yield the same findings),² and what is called “validation” (a general term for consistency with laboratory and clinical tests, guidelines, or predictive measurement instruments). Some estimates show that only 22% to 23% of published results in the biomedical sciences can be validated.³ Others suggest that fewer than half can be.⁴ There are numerous possible reasons for the inability to reproduce research methods and findings: unrecognized study variables, poor study design, poor documentation of findings, outcome reporting bias that falsely inflates the benefits of a new study, inadequate statistical analyses of study data, investigators’ errors or research misconduct, and omission of Food and Drug Administration findings of data falsification or fabrication in the peer-reviewed literature.⁵

Lack of reproducibility by independent investigators may signal that research misconduct took place and that, as a result of fraudulent data, current research

participants and subsequent patients could be harmed if research and medical practices are based on such data. For example, charges of data fabrication and research misconduct have been raised against a researcher who published a family of studies about perioperative use of beta blockers for noncardiac surgery in patients with ischemic heart disease. The findings of these studies had been incorporated into practice guidelines even though systematic reviews six years earlier had identified effect sizes too large to be true.⁶ Later analysis revealed that use of these medications appeared to significantly increase perioperative mortality.⁷ Attempts to replicate this family of studies should have occurred much earlier, when systematic reviews revealed danger signals.

Core protections for research participants—a reasonable risk-benefit ratio, the existence of equipoise, and voluntary participation based on informed consent—assume that the research results of prior research upon which a new study is justified are valid and reproducible. With no system in place to detect that evidentiary claims justifying new studies with research participants cannot be reproduced, there is the danger of a cumulative inaccurate risk-benefit profile that could result in research-related harms to study participants.

Efforts to undertake reproducibility studies are proceeding on many fronts. For example, distressed by a lack of therapies reporting successful interventions for spinal cord injury (SCI), the National Institute of Neurological Disorders and Stroke recently funded studies to replicate several published studies.⁸ Concurrently, a group of researchers established rigorously defined standard data for experiments using animal models to guide SCI research with humans.⁹ Poor reproducibility caused by misidentification or contamination of cell lines is being addressed with newly required validation procedures.¹⁰ Stunned by a series of high-profile cases

Redman BK and Caplan AL. Limited reproducibility of research findings: Implications for the welfare of research participants and considerations for institutional review boards. *IRB: Ethics & Human Research* 2016;38(4):8-10.

of research misconduct, leading researchers in social psychology organized initiatives to replicate published studies in their field. Of the 100 prominent papers analyzed, only 39% could be replicated unambiguously.¹¹ And to provide a basis upon which researchers can undertake replication studies, several journals are beginning to demand that for the peer-review process, authors submit their full data sets along with their manuscripts.

Limited reproducibility is a risk those charged with the protection of research participants need to be aware of and seek to address. Institutional review boards (IRBs) ought to require that research protocols contain explicit probability statements about likely risks and benefits, based on a comprehensive review of prior studies and meta-analyses addressing reproducibility. Such estimates are essential for IRB judgment about minimizing risk, for determining an appropriate risk-benefit ratio for presentation as a part of the informed consent process, and in seeking to facilitate the informed choices of potential research subjects.

Cumulative meta-analyses can help document the degree of reproducibility in prior trials. When researchers submit their protocol for research with humans for IRB review, they should be required to include the results of a search for such studies, as is the case for researchers conducting studies with animals.¹² Such summaries should include the risks and benefits that prior studies identified, including what is known about the likelihood, magnitude, and duration of research risks. This information will be useful in addressing whether the proposed trial will correct methodological or other issues from past trials and will contribute to improving reproducibility. Information from cumulative meta-analysis may also be useful if, during the course of the trial, information from other studies alters the initial risk-benefit ratio. For instance, a summary of more than 1,500 cumulative meta-analyses of clinical intervention studies showed that systematic assessment by researchers of what was already known and replicated could have resulted in less exposure of trial participants to less effective treatments and in some instances could have undermined the rationale for new trials.¹³

In addition, concerns about serious flaws in the reliability of reagents such as antibodies could lead an IRB to require that the proposal outline steps to be taken to ensure their specificity. Rigorous experimental design and transparency in reporting specifics of data collection and analysis not only undergird reproduc-

ible science but also protect current and future research participants from unnecessary risk.

Research participants reasonably expect that their contribution to scientific research will yield progress in knowledge to solve a problem. Since prior irreproducibility can undermine this goal, it is incumbent on researchers and IRBs to flag, as best they can, concerns about reproducibility. Through their ability to obtain input from various experts, IRBs must do their best to seek evidence of reproducibility or lack of it in the line of research for individual protocols they are being asked to approve. Indeed, IRBs are more frequently being asked to judge the validity of findings from preclinical animal studies so as to make more reliable judgments about the potential risks and benefits of

With no system in place to detect that evidentiary claims justifying new studies with research participants cannot be reproduced, there is the danger of a cumulative inaccurate risk-benefit profile that could result in research-related harms to study participants.

translational trials with humans.¹⁴ These efforts and those to assess and improve reproducibility follow from acknowledgment that scientific validity is a basic ethical requirement of research.¹⁵ Of special concern are the high percentage of positive research findings reported in the literature and the under-reporting of harms, which make the IRB's responsibility to minimize risks and ensure reasonable risks in relation to anticipated benefits all the more difficult.

Researchers tend to overestimate potential benefits and underestimate risks, and individuals frequently don't fully understand the consequences of participating in research, making IRB protection essential. Yet one study found that research ethics committees (RECs) in Europe did not have a clear and systematic approach toward assessment of proportionality of research risks.¹⁶ Incorporating reproducibility into REC or IRB performance can help to systematize a more rigorous estimate of proportionality.

Other research has documented that IRBs are confused about whether and how to assess the quality of the science of proposed studies and that IRB members vary in the extent to which they feel that they can ask researchers to alter a study protocol to improve the

scientific quality of the research.¹⁷ While clarification from regulators would be helpful,¹⁸ given the evidence of the need for data from replicated studies, institutions should encourage their IRBs to conduct rigorous scientific analyses of protocols on their own or in conjunction with separate scientific review bodies.

The reproducibility issue regarding the family of beta blocker studies and other research whose results have not been replicated demonstrate an urgent need for a rapid-response safety system to apprise IRBs of concerns about reproducibility that may be associated with unexpected serious harm to research participants.¹⁹ Research funders like the National Institutes of Health are calling for stakeholders to take the steps necessary to reset the self-corrective process of scientific inquiry.²⁰ In the meantime, as the entities that are required to protect the welfare of research participants, IRBs need to recognize the problem of reproducibility and take what steps they can to ensure that the studies individuals are recruited to participate in are designed and carried out on the basis of valid prior scientific findings.

■ **Barbara K. Redman, PhD, MBE**, is an associate of the Division of Medical Ethics at New York University Langone Medical Center, and **Arthur L. Caplan, PhD**, is the director of the Division of Medical Ethics at New York University Langone Medical Center.

References

1. Ioannidis J. How to make more published research true. *PLoS Medicine* 2014;11(10):1001747.
2. Dolgin E. Drug discoverers chart path to tackling data irreproducibility. *Nature Reviews Drug Discovery* 2014;13:875-876.
3. Ioannidis J. How not to be wrong. *New Scientist* 2014;22:32-33.
4. Ioannidis J. Improving validation practices in “omics” research. *Science* 2011;334:1230-1232.
5. Seife C. Research misconduct identified by the US Food and Drug Administration: Out of sight, out of the peer-reviewed literature. *JAMA Internal Medicine* 2015;175:567-577.
6. Chopra V, Eagle KA. Perioperative mischief: The price of academic misconduct. *American Journal of Medicine* 2012;125:953-955.
7. Bouri S, Shun-Shin M, Cole GD, et al. Meta-analyses of secure randomized controlled trials of beta-blockers to prevent perioperative death in non-cardiac surgery. *Heart* 2014;100:456-464.
8. Steward O, Popovich P, Diedrich W, et al. Replication and reproducibility in spinal cord injury research. *Experimental Neurology* 2012;233:597-605.
9. Lemmon V, Ferguson AR, Popovich PG, et al. Minimum information about spinal cord injury experiment. *Journal of Neurotrauma* 2014;31:1-8.
10. Lorsch JR, Collins FS, Lippincott-Schwartz J. Fixing problems with cell biology. *Science* 2014;346:1452-1453.
11. Bohannon J. Many psychology papers fail replication test. *Science* 2015;349:910-911.
12. Animal Welfare Act 9 CFR Pt 1. Section 2.31diii.
13. Clark M, Brice A, Chalmers I. Accumulating research: A systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. *PLoS One* 2014;9(7):e102670.
14. Kimmelman J, London AJ. Predicting harms and benefits in translational trials: Ethics and uncertainty. *PLoS Medicine* 2011;8(3):e1001011.
15. Borgerson K. Redundant, secretive, and isolated: When are clinical trials scientifically valid? *Kennedy Institute of Ethics Journal* 2014;24(4):385-411.
16. Simonsen, S. *Acceptable Risk in Biomedical Research: European Perspectives*. New York: Springer, 2012.
17. Klitzman, R. How good does the science have to be in proposals submitted to IRBs? *Clinical Trials* 2013;10(5):761-766.
18. See ref. 17, Klitzman 2013.
19. Neuman M, Bosk C, Fleisher L. Learning from mistakes in clinical practice guidelines: The case of perioperative beta-blockade. *BMJ Quality and Safety* 2014;23:957-964.
20. Collins FS, Tabak LA. NIH plans to enhance reproducibility. *Nature* 2014;505:612-613.