

Open Access to Clinical Trials Data

Harlan M. Krumholz, MD, SM; Eric D. Peterson, MD, MPH

Well-conducted randomized clinical trials (RCTs) are the gold standard for evaluating the safety and efficacy of medical therapeutics. Yet most often, a single group of individuals



Related article page 1024

who conducted the trial are the only ones who have access to the raw data, conduct the analysis, and publish the study results. This limited access does not typically allow others to replicate the trial findings. Given the time and expense required to conduct an RCT, it is often unlikely that others will independently repeat a similar experiment. Thus, the scientific community and the public often accept the results produced and published by the original research team without an opportunity for reanalysis. Increasingly, however, opinions and empirical data are challenging the assumption that the analysis of a clinical trial is straightforward and that analysis by any other group would obtain the same results.¹⁻³

In this issue of *JAMA*, Ebrahim et al⁴ report their findings based on a rigorous search of previously published reanalyses of RCTs. Their first surprising and disconcerting finding was just how infrequently data reanalysis has occurred in medical research. Searching the literature from 1966 to present, the authors found only 37 reports that met their criteria as an RCT reanalysis. Of these few reanalyses performed, the majority (84%) had overlapping authors from the original report. Thus, reanalyses are not only rare, but the majority that were reported were not fully independent of the original research group. Despite this overlap, Ebrahim et al report that about half of the reanalyses differed in statistical or analytic approaches, a third differed in the definitions or measurements of outcomes, and most important, a third led to interpretations and conclusions different than those in the original article. While the definition of what constituted different trial analyses, study end points, findings, and interpretations is subjective, the authors' general conclusions were consistent with an emerging literature that indicates RCT reanalysis can yield different results and conclusions from those originally published.

Even when the original investigators are presenting evidence in different venues it is not always consistent. For example, there is evidence from trials that data presented to the US Food and Drug Administration (FDA) may differ in important ways from those originally presented at scientific sessions or published in medical journals. Rising et al⁵ assessed clinical trial information provided to the FDA and reported a 9% discordance between the conclusions in the report to the FDA and in the published article. Not unexpectedly, all were in the direction favoring the drug.

Another example is discordance between what is reported in ClinicalTrials.gov and what is published in journal articles. Hartung et al² showed that in a random sample of phase 3 and 4 trials, in 15% the primary end point in the main article was different from the primary end point the trialists reported in ClinicalTrials.gov. Moreover, 22% reported the primary outcome value inconsistently, with some even having differences in the number of deaths. Other studies have found similar rates of discordance.⁶

When reanalyses by different groups obtain somewhat different results or reach alternative conclusions, the cause is not necessarily bias. Independent groups with individual patient-level data from trials will not always reach the same conclusions because every study involves discretionary decisions. This situation was exemplified in the evaluation of Medtronic's bone morphogenetic protein 2 (BMP-2) product. Two expert organizations with international reputations in the conduct of systematic reviews were given the same individual patient-level data from the BMP-2 trials, the same task, the same resources, and the same timeline.^{7,8} Nevertheless, in their final reports, their methods were not identical and their results and conclusions differed in important ways. Consistent with these findings, meta-analyses on similar topics and using similar data do not always show concordant results.⁹

The current review by Ebrahim et al, as well as other cited work, suggests several important next steps needed to ensure transparency and open access in RCTs. First, all RCTs, their prespecified study protocols and analytic plans, and their results should be registered and reported to the medical community, fulfilling the ethical promise made to those enrolled in the scientific experiment. Such a step would contribute to improvements in the standardization of trial registration and reporting of results, which remains varied despite governmental regulation and journal policies.¹⁰⁻¹² Full availability of trial registration data is essential to allow peer reviewers and journals to monitor trial protocols and analytic plans to ensure consistency and thereby reduce some of the variation that may occur in the reporting of results, particularly with respect to primary, secondary, and exploratory outcomes.

Second, raw data and metadata (all the information about the data) from the original trial should ideally be made available to those who seek the opportunity to replicate the findings. Such independent verification would markedly increase the scientific community's confidence in the study findings. Even when results differed importantly, it would allow for open dialogue that would promote a deeper understanding of the study and its interpretation. While the current study by Ebrahim et al demonstrates how infrequently

reanalyses have actually occurred, it is notable that several drug and device companies have recently taken steps toward this goal, and some National Institutes of Health institutes require that data from clinical trials they support be made publicly available.¹³

Several concerns have been raised about trial data sharing. There are concerns that reanalyses could compromise patient and investigator confidentiality. Yet the risks to study participants can be mitigated by database deidentification and other legal restrictions. Some may have concerns that data access will be used for commercial gain. Again, unauthorized use of the data could be controlled by research and data use agreements or alternatively, by the creation of protected data analytic workspaces in which trial reanalyses could be conducted by independent investigators in a safe setting. Some may have concerns that the conduct of research by those who did not generate the data will deprive the original investigators of the return on their time investment. However, the challenge to academic medicine is to fairly credit the progenitors of valuable data and not create incentives for investigators to sequester and shield data from widespread use. Others are concerned that disparate results obtained from reanalyses of the same data set by different groups could cause confusion and potential harm or be used by those to promote an agenda with analyses that do not represent good science.^{14,15} A contrary view is that the status quo, whereby data are held tightly and one interpretation prevails, may be just as or more harmful, perhaps codifying faulty messages and conclusions. Rather than causing confusion, making data available for independent

verification is essential to an open dialogue about trial results.

Christakis and Zimmerman suggested some criteria for the conduct of reanalyses.¹⁴ In their view, the reanalyses should be prespecified, formalized as a detailed study protocol, and registered in the same manner as the original trial. Additionally, it should be expected, even required, that the results be shared publicly through peer-reviewed publications, scientific meetings, and registration sites. It is also appropriate etiquette to share the results with the original investigators prior to public dissemination. There is no question that for this approach to flourish there will be a need to codify standards and resolve many of the challenges that include distribution of data, authorization and oversight of those who have access, and evaluation of the program. Nevertheless, these challenges can be addressed, and any program should improve iteratively over time.

Replication is a vital part of the scientific method. Fields outside of medicine have already embraced sharing experimental data, as have the basic biological sciences within medicine. The culture of clinical research in medicine will need to evolve for open science to succeed. The recognition that one trial can potentially lead to different findings and conclusions depending on many discretionary decisions that are made about the data and reanalyses almost mandates that those choices are transparent and described in detail—and that others have the chance to replicate them. Rather than the rare exception, open science and replication should become the standard for all trials and especially those that have high potential to influence practice.

ARTICLE INFORMATION

Author Affiliations: Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut (Krumholz); Duke University Medical Center, Durham, North Carolina (Peterson); Associate Editor, *JAMA* (Peterson).

Corresponding Author: Eric D. Peterson, MD, MPH, Duke University Medical Center, PO Box 17969, Durham, NC 27715 (eric.peterson@duke.edu).

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Krumholz reports board membership with UnitedHealthcare and grants or pending grants with the Centers for Medicare & Medicaid Services, Medtronic, and Johnson & Johnson. The Duke Clinical Research Institute, of which Dr Peterson is the Executive Director, conducts clinical trials and other research with a number of governmental and industry partners. A full list of all relationships is available at <http://www.dcri.org/about-us/conflict-of-interest>.

REFERENCES

- Ross JS, Krumholz HM. Ushering in a new era of open science through data sharing: the wall must come down. *JAMA*. 2013;309(13):1355-1356.
- Hartung DM, Zarin DA, Guise JM, McDonagh M, Paynter R, Helfand M. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. *Ann Intern Med*. 2014;160(7):477-483.
- Krumholz HM. Open science and data sharing in clinical research: basing informed decisions on the totality of the evidence. *Circ Cardiovasc Qual Outcomes*. 2012;5(2):141-142.
- Ebrahim S, Sohani ZN, Montoya L, et al. Reanalyses of randomized clinical trial data. *JAMA*. doi:10.1001/jama.2014.9646.
- Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. *PLoS Med*. 2008;5(11):e217.
- Becker JE, Krumholz HM, Ben-Josef G, Ross JS. Reporting of results in ClinicalTrials.gov and high-impact journals. *JAMA*. 2014;311(10):1063-1065.
- Simmonds MC, Brown JV, Heirs MK, et al. Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: a meta-analysis of individual-participant data. *Ann Intern Med*. 2013;158(12):877-889.
- Fu R, Selph S, McDonagh M, et al. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. *Ann Intern Med*. 2013;158(12):890-902.
- Thompson R, Bandera E, Burley V, et al. Reproducibility of systematic literature reviews on food, nutrition, physical activity and endometrial cancer. *Public Health Nutr*. 2008;11(10):1006-1014.
- Update on trials registration: clinical trial registration: looking back and moving ahead. www.icmje.org/news-and-editorials/clinical_trial_reg_jun2007.html. June 2007. Accessed August 19, 2014.
- Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007-2010. *JAMA*. 2012;307(17):1838-1847.
- Ross JS, Tse T, Zarin DA, Xu H, Zhou L, Krumholz HM. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. *BMJ*. 2012;344:d7292.
- Ross JS, Krumholz HM. Ushering in a new era of open science through data sharing: the wall must come down. *JAMA*. 2013;309(13):1355-1356.
- Spertus JA. The double-edged sword of open access to research data. *Circ Cardiovasc Qual Outcomes*. 2012;5(2):143-144.
- Christakis DA, Zimmerman FJ. Rethinking reanalysis. *JAMA*. 2013;310(23):2499-2500.