

COMMENTARY

Open Access



Clinical research data sharing: what an open science world means for researchers involved in evidence synthesis

Joseph S. Ross^{1,2,3}

Abstract

The International Committee of Medical Journal Editors (ICMJE) recently announced a bold step forward to require data generated by interventional clinical trials that are published in its member journals to be responsibly shared with external investigators. The movement toward a clinical research culture that supports data sharing has important implications for the design, conduct, and reporting of systematic reviews and meta-analyses. While data sharing is likely to enhance the science of evidence synthesis, facilitating the identification and inclusion of all relevant research, it will also pose key challenges, such as requiring broader search strategies and more thorough scrutiny of identified research. Furthermore, the adoption of data sharing initiatives by the clinical research community should challenge the community of researchers involved in evidence synthesis to follow suit, including the widespread adoption of systematic review registration, results reporting, and data sharing, to promote transparency and enhance the integrity of the research process.

Keywords: Data sharing, Systematic reviews, Clinical trials

Background

In January 2016, the International Committee of Medical Journal Editors (ICMJE) announced a bold proposal to require data generated by interventional clinical trials that are published in its member journals to be responsibly shared with external investigators, both to fulfill an ethical obligation to the research participants and to enhance the integrity of the clinical research process by enabling independent confirmation of results [1]. In response, more than 300 comments were submitted by researchers and organizations, at least some of which were unresponsive [2]. Thus, while any guidance from the ICMJE on this issue is yet to be finalized, we can reflect upon what the overall movement toward a research culture that supports data sharing might mean for the community of researchers involved in evidence

synthesis and the design, conduct, and reporting of systematic reviews and meta-analyses.

Main text

Support for data sharing as a means for improving the reliability of medical evidence and reinforcing evidence-based practice has grown steadily over the past 5 to 10 years. Over this time, several major research funders, including the US National Institutes of Health, the US Patient-Centered Outcomes Research Institute, the UK Medical Research Council, and the Bill and Melinda Gates Foundation, have adopted policies supporting or mandating clinical research data sharing [3, 4]. Private industry, by far the biggest funder of clinical research, has similarly adopted supportive policies [5], perhaps even stronger than those adopted by government and non-profit funders [6]. Academic clinical trialists, while expressing some concerns, have also demonstrated growing support for data sharing [7–9].

Correspondence: joseph.ross@yale.edu

¹Section of General Internal Medicine and the Robert Wood Johnson Foundation Clinical Scholars Program, Yale School of Medicine, New Haven, CT, USA

²Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA

Full list of author information is available at the end of the article



© 2016 The Author(s). **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

In January 2015, the Institute of Medicine (IOM) of the US National Academies further strengthened these efforts with its report, “Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risks,” recommending that stakeholders foster a culture in which data sharing is the expected norm and commit to responsible strategies aimed at maximizing benefits, minimizing risks, and overcoming challenges of sharing clinical trial data [10]. The 2016 ICMJE proposal clearly aligns with the IOM recommendations, requiring as a condition of consideration for publication of a clinical trial report in a member journal that the authors share with others the de-identified individual patient-level data (IPD) underlying the results presented in the article within 6 months of publication and include a plan for data sharing as a component of clinical trial registration.

A clinical research infrastructure that requires data sharing, and is increasingly transparent, holds great promise for the future of evidence synthesis [11]. Among the many challenges to rigorous systematic review and meta-analysis is non-publication of research, prohibiting the identification and inclusion of all relevant research. In addition to selective publication, other challenges include selective outcome reporting and lack of published data for potentially relevant patient subgroups. As prior studies have demonstrated, somewhere on the order of one half to one third of clinical research studies are never published [12–18], and far fewer report all collected endpoints of clinical importance, especially safety endpoints [19–21]. Data sharing may help to address publication and outcome reporting biases to improve the accuracy of systematic reviews and ensure that each is based on all relevant research studies. Researchers involved in evidence synthesis will now be able to request access to the IPD from a clinical trial to investigate briefly reported safety endpoints or examine efficacy endpoints among important subgroups, allowing fuller examination treatment effect heterogeneity for patient populations of potential interest to clinical communities, professional organizations, and guideline committees. However, it is important to note that the ICMJE proposal only applies to clinical trials published in member journals; for unpublished clinical research to be identified and included in systematic reviews, researchers will continue to rely upon research funders’ data sharing initiatives as well as trial registration and results reporting requirements.

At the same time, a clinical research infrastructure that requires data sharing poses challenges for evidence synthesis. First, the universe of applicable databases that need to be searched to identify all relevant research continues to expand, potentially taxing researchers or requiring the involvement of information specialists as they scrutinize a multitude of sources that now includes data sharing platforms, such as DRYAD, ClinicalStudyDataRequest.com,

and the Yale University Open Data Access Project [22–24]. This widening universe also requires that investigators carefully cross-check among identified trials, to ensure that the same trial is not erroneously included multiple times because of discordant information reporting in multiple databases. Second, for published trials, the ICMJE proposal clearly states that all de-identified IPD underlying the results presented in the article are required to be shared. However, the availability of IPD might be considered a double-edged sword, as researchers may now face the burden of validating published trial results using IPD prior to including the summary findings in a systematic review or meta-analysis. Third, the proposal is silent on IPD from the clinical trial that does not underlie results presented in the article. For instance, this requirement could be interpreted to preclude sharing of data on the efficacy and safety outcomes, whether common or rare, not reported in the main trial publication, as well as results for patient subgroups that were not reported. Researchers must still do the legwork to obtain information on these unpublished outcomes and patient subgroups. Fourth, the ICMJE proposal may lead to a situation in which authors who publish multiple articles from a trial prepare multiple IPD files from the same clinical trial, creating version control issues and potential confusion among researchers who attempt to reconcile trial findings across the multiple IPD files. Related to this point, the availability of results from multiple sources, including articles, trial registries, and IPD, necessitates that researchers involved in evidence synthesis resolve discrepancies that are likely to exist among all reported results from included individual trials [25, 26].

A larger challenge also exists for the evidence synthesis community, including organizations such as the Cochrane Collaboration that have long advocated for stronger data sharing policies [27]. These researchers should now take on the same challenges being imposed on clinical trial researchers, enhancing efforts in the field to promote transparency and share clinical research data. As a condition of publication, in 2005, the ICMJE began requiring clinical researchers to prospectively register clinical trials in a public trial registry that allowed reporting of key trial information and pre-specification of primary and secondary outcomes, as well as safety endpoints [28]. While adherence to this policy has not been perfect [29–31], it led to remarkable increases in clinical trial registration [32] and made clinical research far more transparent. A similar effort to require systematic review registration was begun in 2012 in order to reduce unplanned duplication of reviews, provide transparency, and minimize reporting bias [33]. Efforts to ensure adherence to these requirements are needed, as the ICMJE does not currently require prospective systematic review registration as a condition of publication and some issues in primary outcome reporting have been identified [34]. However, with

more than 10,000 registered systematic reviews in the largest international public registry [35], the field is clearly moving toward a more transparent scientific process.

Evidence synthesis researchers and organizations should similarly adopt policies supporting results reporting and research data sharing. The ICMJE has supported trial results reporting since 2007, when the US Food and Drug Administration Amendments Act initiated the requirement for medical product intervention trials, clarifying that brief, structured results posted in the same clinical trials register in which the initial registration resides would not be considered a prior publication [36]. Systematic review registries should be modified to accommodate results reporting, not only linkages to published articles, so that the results of unpublished systematic reviews can be better disseminated throughout the research community. Time frames for results reporting should also be specified, ideally within 12 months of systematic review completion or to coincide with publication, whichever comes first. A structured format for the reporting of results should be developed by researchers and stakeholder organizations in the field.

Efforts should also be undertaken to facilitate the sharing of libraries of articles that were aggregated for systematic reviews and meta-analyses, along with the results, summary data, and IPD that underlie the work. Resources will need to be allocated to create repositories for this information, such as the Open Science Framework [37], which could also potentially be managed by systematic review registries. To overcome copyright issues, aggregated articles might be listed on registries with linkages to the publications at journal publishers' websites or PubMed at the US National Library of Medicine. Such a system is quite similar to how institutional repositories currently work and, if linked with publishers' websites, might encourage publishers to enhance publications' meta-data and facilitate retrievability and usability. To overcome IPD data ownership issues, explicit details from where and how data access was obtained could be provided, including linkages to data sharing repositories. Time frames for sharing aggregated articles, results, summary data, and IPD sharing should be specified and align with those recommended by the IOM and proposed by the ICMJE, and formats for distribution will need to be standardized. Finally, as in the clinical research community, some means of providing appropriate credit to those researchers sharing data needs to be developed and recognized in the academic community. While each of these suggestions to share data requires additional time and effort on the behalf of researchers, these are the same burdens that have been imposed on clinical trial researchers. Moreover, one could argue that the potential to reduce unplanned duplicative searching, aggregation, and analysis in the evidence synthesis community is likely far greater than the

potential to reduce unplanned duplicative clinical trials in the clinical research community, although the burden of the latter is greater because it involves patients.

Conclusion

Data sharing is about more than minimizing duplicative data collection efforts. Sharing reduces research costs and lowers human participant burden. Sharing maximizes the value of collected data by enabling follow-up studies of secondary research questions by a multitude of investigators using existing data. Sharing encourages multiple examinations and interpretations of data, both protecting against faulty analyses and contributing to replication, refinement, or refutation of prior work. Sharing positions research data as a public good [38]. The entire biomedical research community, from the basic sciences to the clinical sciences to the evidence synthesis sciences, has much to be gained from data sharing efforts. The ICMJE proposal is the next important step in the overall movement toward a research culture that supports data sharing. The community of researchers involved in evidence synthesis should embrace the challenges being imposed on clinical trial researchers in order to promote transparency and share research data.

Abbreviations

ICMJE: International Committee of Medical Journal Editors; IOM: Institute of Medicine; IPD: Individual patient-level data

Acknowledgements

None

Funding

This article was not supported by any specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The author assumes full responsibility for the accuracy and completeness of the ideas presented.

Availability of data and materials

Not applicable

Author's contributions

JSR was solely responsible for the article content and design and drafting of the manuscript.

Author's information

Dr. Ross is an Associate Professor of Medicine (General Internal Medicine) and Public Health (Health Policy and Management) at Yale University. Mail correspondence should be directed to the following: Joseph S. Ross, MD, MHS, Section of General Internal Medicine, Yale University School of Medicine, P.O. Box 208093, New Haven, CT 06520-8093, USA; email correspondence: joseph.ross@yale.edu.

Competing interests

Dr. Ross declares (currently or formerly) receiving support through Yale University from Medtronic, Inc., and Johnson and Johnson to develop methods of clinical trial data sharing, from the Blue Cross Blue Shield Association (BCBSA) to better understand medical technology evidence generation, from the Centers of Medicare and Medicaid Services (CMS) to develop and maintain performance measures that are used for public reporting, and from the Food and Drug Administration (FDA) to develop methods for post-market surveillance of medical devices.

Consent for publication

Not applicable

Ethics approval and consent to participate

Not applicable

Author details

¹Section of General Internal Medicine and the Robert Wood Johnson Foundation Clinical Scholars Program, Yale School of Medicine, New Haven, CT, USA. ²Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA. ³Center for Outcomes Research and Evaluation, Yale–New Haven Hospital, New Haven, CT, USA.

Received: 1 June 2016 Accepted: 9 September 2016

Published online: 20 September 2016

References

- Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, Fletcher J, Frizelle FA, Groves T, Haileamlak A, et al. Sharing clinical trial data: a proposal from the International Committee of Medical Journal Editors. *Ann Intern Med.* 2016;164:505–6.
- Comments on the ICMJE's proposals for sharing clinical trial data. [http://icmje.org/news-and-editorials/sharing_clinical_trial_data_comments_feed.html]. Accessed 14 July 2016.
- Ross JS, Krumholz HM. Ushering in a new era of open science through data sharing: the wall must come down. *JAMA.* 2013;309:1355–6.
- Hudson KL, Collins FS. Sharing and reporting the results of clinical trials. *JAMA.* 2015;313:355–6.
- Principles for responsible clinical trial data sharing. Our commitment to patients and researchers. [<http://phrma.org/sites/default/files/pdf/PhRMAPrinciplesForResponsibleClinicalTrialDataSharing.pdf>]. Accessed 14 July 2016.
- Krumholz HM, Gross CP, Blount KL, Ritchie JD, Hodshon B, Lehman R, Ross JS. Sea change in open science and data sharing: leadership by industry. *Circ Cardiovasc Qual Outcomes.* 2014;7:499–504.
- Rathi V, Dzara K, Gross CP, Hrynaskiewicz I, Joffe S, Krumholz HM, Strait KM, Ross JS. Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ.* 2012;345:e7570.
- Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, Read E, Manoff M, Frame M. Data sharing by scientists: practices and perceptions. *PLoS One.* 2011;6:e21101.
- Tenopir C, Dalton ED, Allard S, Frame M, Pjesivic I, Birch B, Pollock D, Dorsett K. Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS One.* 2015;10:e0134826.
- Institute of Medicine of the National Academies of Science. *Sharing clinical trial data: maximizing benefits, minimizing risks.* Washington, DC: National Academies Press; 2015.
- Ross JS, Gross CP, Krumholz HM. Promoting transparency in pharmaceutical industry-sponsored research. *Am J Public Health.* 2012;102:72–80.
- Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E, Decullier E, Easterbrook PJ, Von Elm E, Gamble C, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One.* 2008;3:e3081.
- Dwan K, Altman DG, Clarke M, Gamble C, Higgins JP, Sterne JA, Williamson PR, Kirkham JJ. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. *PLoS Med.* 2014;11:e1001666.
- Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet.* 1991;337:867–72.
- Ross JS, Mulvey GK, Hines EM, Nissen SE, Krumholz HM. Trial publication after registration in ClinicalTrials.gov: a cross-sectional analysis. *PLoS Med.* 2009;6:e1000144.
- 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.
- Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med.* 2008;358:252–60.
- Chen R, Desai NR, Ross JS, Zhang W, Chau KH, Wayda B, Murugiah K, Lu DY, Mittal A, Krumholz HM. Publication and reporting of clinical trial results: cross sectional analysis across academic medical centers. *BMJ.* 2016;352:i637.
- Doshi P, Jefferson T, Del Mar C. The imperative to share clinical study reports: recommendations from the Tamiflu experience. *PLoS Med.* 2012;9:e1001201.
- Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervolgyi V, Kohlepp P, Kamphuis M, Grouven U. Completeness of reporting of patient-relevant clinical trial outcomes: comparison of unpublished clinical study reports with publicly available data. *PLoS Med.* 2013;10:e1001526.
- Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med.* 2009;169:1756–61.
- DRYAD. [<http://datadryad.org/>]. Accessed 14 July 2016.
- Yale University Open Data Access Project. [<http://yoda.yale.edu/>]. Accessed 14 July 2016.
- ClinicalStudyDataRequest.com. [<https://www.clinicalstudydatarequest.com/>]. Accessed 14 July 2016.
- Becker JE, Krumholz HM, Ben-Josef G, Ross JS. Reporting of results in ClinicalTrials.gov and high-impact journals. *JAMA.* 2014;311:1063–5.
- 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;160:477–83.
- Press release: Cochrane Collaboration statement on access to clinical trial data. [<http://www.cochrane.org/features/clinical-trials-statement-press-release>]. Accessed 14 July 2016.
- DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, Kotzin S, Laine C, Marusic A, Overbeke AJ, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA.* 2004;292:1363–4.
- Boccia S, Rothman KJ, Panic N, Flacco ME, Rosso A, Pastorino R, Manzoli L, La Vecchia C, Villari P, Boffetta P, et al. Registration practices for observational studies on ClinicalTrials.gov indicated low adherence. *J Clin Epidemiol.* 2016;70:176–82.
- Dal-Re R, Ross JS, Marusic A. Compliance with prospective trial registration guidance remained low in high-impact journals and has implications for primary end point reporting. *J Clin Epidemiol.* 2016;75:100–7.
- Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database—update and key issues. *N Engl J Med.* 2011;364:852–60.
- Zarin DA, Tse T, Ide NC. Trial registration at ClinicalTrials.gov between May and October 2005. *N Engl J Med.* 2005;353:2779–87.
- Booth A, Clarke M, Dooley G, Ghera D, Moher D, Petticrew M, Stewart L. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev.* 2012;1:2.
- Tricco AC, Cogo E, Page MJ, Polisen J, Booth A, Dwan K, MacDonald H, Clifford TJ, Stewart LA, Straus SE, Moher D. A third of systematic reviews changed or did not specify the primary outcome: a PROSPERO register study. *J Clin Epidemiol* 2016; doi: 30010.31016/j.jclinepi.32016.30003.30025. [Epub ahead of print].
- Welcome to PROSPERO, international prospective register of systematic reviews. [<http://www.crd.york.ac.uk/PROSPERO/>]. Accessed 14 July 2016.
- Laine C, De Angelis C, Delamothe T, Drazen JM, Frizelle FA, Haug C, Hebert PC, Horton R, Kotzin S, Marusic A, et al. Clinical trial registration: looking back and moving ahead. *Ann Intern Med.* 2007;147:275–7.
- Open Science Framework. [<https://osf.io/>]. Accessed 14 July 2016.
- Rodwin MA, Abramson JD. Clinical trial data as a public good. *JAMA.* 2012;308:871–2.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

