



The MRCT Center of Harvard and Brigham and Women's Hospital
14 Story Street, 4th Floor
Cambridge, MA 02138 USA

Submitted electronically to: scientificdata@nature.com

Scientific Data
The Macmillan Building
4 Crinan Street
London N1 9XW
United Kingdom

September 7, 2015
RE: Editorial (Open)
Article number: 150034 (2015)
Let's be pragmatic about clinical data (07 July 2015)

Dear Editor:

The Multi-Regional Clinical Trials (MRCT) CENTER of Harvard University and Brigham and Women's Hospital has been working toward pragmatic solutions for Clinical Trial Data Sharing since holding our first multi-stakeholder conference in May 2013.¹

Scientific Data is an open access journal for descriptions of scientifically valuable datasets and through which publications are associated with their corresponding publicly available datasets. To date, non-public datasets have not had a publication outlet for linking these valuable assets. While well-publicized efforts such as YODA and CSDR may yield many inquiries from external researchers, publicly funded investigators or smaller industry players who wish to prepare their datasets for sharing may be reluctant to invest those resources without knowing that they will be utilized. Smaller data generators have virtually no outlet to publicize the fact that their datasets exist and receive due academic credit for such data sharing. *Scientific Data's* innovative approach may fill this void.

Scientific Data's open access journal would permanently link each publication to its dataset including the details of the dataset. This represents an effort that has the potential to enable vastly expanded data discoverability of these previously inaccessible datasets and publicize newly available data. The description of the dataset and data descriptors would be subject to formal peer review, thus eliminating datasets that are incomplete, poorly curated. Importantly, this process serves to create a permanent public record through publication and citation.

¹ *Issues and Case Studies in Clinical Trial Data Sharing: Lessons and Solutions*, Conference, Harvard Law School, May 17 2013, Cambridge, MA

We applaud the journal's effort in encouraging authors to work with existing data repositories. In concert with and parallel to the efforts of *Scientific Data*, the MRCT Center is working with a multi-stakeholder group that includes the Institute of Medicine, the Wellcome Trust, the Laura and John Arnold Foundation and industry to create a portal and a central user interface to search all clinical trials (including participant-level data sets where available), a federated portal that would allow data generators to host their own data or deposit data onto this or other platforms. In addition, the platform would provide the capability to perform data analysis with or without downloading, and to process data requests via independent review. We hope to consolidate efforts to create a trans-national non-profit entity that would have responsibility for this platform, to include policies and practices to maximize data utilization with attention to interoperability, participant privacy, fair business practices, and flexibility. We welcome the opportunity to work with and alongside like-minded collaborators to align priorities, integrate feedback and address interests of stakeholders invested in supporting and facilitating data sharing (see <http://mrccenter.org/news/launch-data-sharing-working-groups>).

Further, we suggest *Scientific Data* consider:

- 1) Description, delineation and publication of the determinants and qualities of a trial dataset that would be applied in peer review of each dataset.
- 2) The logistics surrounding the responsibilities of the peer reviewer(s) must be clearly outlined. For instance, it is unclear whether the peer reviewer will be responsible for verifying publication claims against the dataset provided and concordance with the Joint Declaration of Data Citation Principles. If this is the responsibility of the reviewer, each review may require the inclusion of an experienced statistician. The cost of a potentially resource-intensive plan must be considered in terms of time, necessary expertise in peer review, permanence, quality, ease of discoverability and persistence of links, among other issues.
- 3) The journal encourages editors to determine compliance with data sharing policies for every submission and to check that the research is consistent with the original proposal and with the DUA. Given the limited resources of small journals, one might consider if it would be sufficient for each author to attest that the submission is compliant with the outlined conditions.
- 4) That ultimately this proposal may provide publication outlet for negative trials and potentially ease the currently well-known publication bias of positive results.
- 5) Although leadership at *Scientific Data* do not claim to directly address issues surrounding incentives for data sharing with this proposal, we suggest that the essence of this plan, permanently linking publications to datasets, may indeed align incentives for academics to share data—the benefit being increased opportunity for publication and use of publication as promotion of the research and shared data, thereby rewarding and quantifying data sharing. Over time, such a metric may be offered for consideration in academic and career advancement.

We acknowledge that *Scientific Data's* initiatives to support and encourage data sharing is an important, impactful first step, and issues surrounding peer review criteria and implementation deserve attention in the near future.



The comments have been submitted on behalf of the Multi-Regional Clinical Trials (MRCT) Center of Harvard University and Brigham and Women's Hospital Data Sharing and Transparency workgroup. The opinions contained herein are those of the Multi-Regional Clinical Trials Center and Data Sharing and Transparency workgroup, and are not intended to represent the position of Harvard University or its schools.

Respectfully submitted:

Rebecca H. Li, PhD

Barbara Bierer, MD

Mark Barnes, JD

On behalf of the Multi-Regional Clinical Trials Center of Harvard and Brigham and Women's Hospital Data Sharing and Transparency Work Group