BRIGHAM HEALTH

BWH BRIGHAM AND
WOMEN'S HOSPITAL

MRCT MULTI-REGIONAL
CLINICAL TRIALS

THE MRCT CENTER of
BRIGHAM AND WOMEN'S HOSPITAL
and HARVARD

January 7, 2020

Francis S. Collins, MD, PhD
National Institutes of Health
6705 Rockledge Drive, Suite 750
Bethesda, MD 20892
Submitted electronically: https://osp.od.nih.gov/draft-data-sharing-and-management/

RE: DRAFT NIH Policy for Data Management and Sharing and Supplemental DRAFT Guidance

Dear Dr. Collins:

The Multi-Regional Clinical Trials Center of Brigham and Women's Hospital and Harvard (MRCT Center) appreciates the opportunity to comment on the National Institutes of Health (NIH) draft NIH "*Policy for Data Management and Sharing and Supplemental DRAFT Guidance"* (hereinafter the "Policy"), published in the Federal Register Vol. 84, No. 217 on November 8, 2019.

The MRCT Center is a research and policy center that addresses the ethics, conduct, oversight, and regulatory environment of international, multi-site clinical trials. Founded in 2009, it functions as a neutral convener to engage diverse stakeholders from industry, academia, patients and patient advocacy groups, non-profit organizations, and global regulatory agencies. The MRCT Center focuses on pre-competitive issues, to identify challenges and to deliver ethical, actionable, and practical solutions for the global clinical trial enterprise. Over the last five years, the MRCT Center has been intimately involved in data sharing, including (1) developing guidance for sharing aggregate plain language summaries for participants and the public, (2) developing guidance for sharing individual results with participants, (3) promoting principles of individual participant data (IPD) sharing including protections of patient/participant confidentiality and privacy and of confidential commercial information, (4) developing template data use agreements and data contributor agreements for IPD and other data sharing, (5) crafting informed consent language to promote participant understanding of the implications of sharing de-identified data, (6) launching Vivli, a platform for global data sharing of IPD data, and (7) furthering the establishment of credit for data sharing for those individuals who choose to share their data, among other efforts. Of note, the responsibility for the content of this document rests

with the leadership of the MRCT Center, not with the its collaborators, nor with the institutions affiliated with the authors.[1]

The MRCT Center strongly endorses the NIH draft policy and the importance that it places on data management and data sharing. This draft policy demonstrates an ongoing appreciation by the NIH of the utility and value of previously collected data and metadata not only for replication but for new discoveries. Further, proper stewardship of data is important, and the requirement for the submission of data management and data sharing plans prior to initiation of the research will be helpful in that regard. We are enthusiastic that NIH has taken this further step to include all scientific data (and metadata) as defined, of all data types and all sizes, and for all research funded by the NIH. We also understand that the NIH has outlined only the minimum expectations for NIH-wide Plans, and that the NIH ICOs may add additional requirements or expectations. We believe, however, that the NIH policy should be stronger, while nevertheless still permitting some flexibility.

We feel strongly that the **NIH should require data sharing, unless there is an ethical, scientific, or other defensible reason not to do so**. There should be a rebuttable presumption to share data; the burden should be on the investigator to provide cogent reasons that the data should not or cannot be shared. Subjective evaluations by investigators of potential data utility to the research community or the public should not be considered a sufficient reason not to share data.

There are risks to data sharing, including that of participant and patient privacy for studies that involve human participants and their data or biospecimens. Not all data need be downloadable and freely accessible: **measures to protect privacy and confidentiality** should be required. Those measures include de-identification, as mentioned in the draft policy, but also include other risk mitigation strategies: physical and technical security measures (e.g. data maintained in a repository, in a fit-for-purpose compute environment and not downloadable), controlled data access by qualified users, and other more novel methods (e.g. differential privacy, block chain technologies, etc.). We encourage the NIH to invest in the development and dissemination of these technologies to promote data sharing of sensitive data, and to issue appropriate guidance for their use. We further encourage the NIH to require disclosure of—and explanation of—data sharing plans to research participants during the informed consent process.

We encourage the NIH to provide **minimum expectations** for data management and scientific data, either within the policy or as additional guidance. The breadth of research and data acquisition supported by the NIH is expansive, covering different disciplines and including the spectrum of basic, translational, and clinical research. Guidance is needed to

---

[1] Brigham and Women's Hospital, Rope & Gray LLP, Harvard Medical School, Harvard University, and Yale Law School.

assist investigators and institutions, many unfamiliar with optimal data management and data sharing approaches.

Specific, **required elements** of the Plan should be developed, and an approximate (or "not to exceed") **time frame** regarding when the data will be made available should be stated. The completeness and sufficiency of the Plan will only be encouraged by written detail.

We appreciate the development of the Supplemental DRAFT Guidance: Elements of a NIH Data Management and Sharing Plan (Plan). While the descriptions of the specific data elements provide the reader with guidance on the development of a Plan, we encourage NIH to further complement this guidance with examples of (potential) comprehensive data sharing plans for different data types.

We also encourage NIH to provide **minimum expectations for data repositories and data sharing platforms** that meet requirements of the policy. We encourage NIH to develop and **maintain a database** that recognizes those repositories and platforms.

The policy states that "NIH may make Plans publicly available." We believe that the **NIH should affirm its commitment to make available to the public the Plans** of funded research proposals and contracts. Public visibility of the Plans will be informative and educational, permit tracking, and encourage compliance. ClinicalTrials.gov should be used to disseminate the Plans for registered clinical trials, and the Plans should be posted prior to study initiation. Additional repositories can be used for other types of research, or the NIH can simply publish the Plan as an additional field linked to or hosted on the NIH RePORTER.

Data holders and data contributors should be encouraged to apply **data tags (i.e. metadata) that describe how the data can be used**—and applicable restrictions to its use—to reflect any contractual terms (e.g. licensing, copyright), informed consent parameters, and institutional, state, and federal policies. Metadata that describe the terms of use will help ensure the appropriate and compliant use of the data in the future. Further, NIH should invest in developing a universal language or library for such data tags and tools to render such metadata machine-readable.

The burden of managing and sharing data does not rest solely on the data contributor but equally on the data scientists and researchers who have access to the data. **Strict policies with enforcement provisions should be communicated to those who access the data**, and data use agreements employed as appropriate. Data tagging as described above will make compliance both easier for the user and auditable if necessary.

The data management and sharing plan should be an important and determinative part of any NIH proposal, and the **Plan should be reviewed and scored by the study section** (or contracting entity). The Plan should not be relegated to a "Just-In-Time" submission but should affect whether a proposal is prioritized for funding. Consideration of data

BRIGHAM HEALTH
BWH BRIGHAM AND
WOMEN'S HOSPITAL

MRCT MULTI-REGIONAL
CLINICAL TRIALS
THE MRCT CENTER of
BRIGHAM AND WOMEN'S HOSPITAL
and HARVARD

management, integrity, and stewardship (and, later, sharing) is an integral part of study design and quality.

We believe further that **no two-page limitation** should be imposed on the Plan. The prospective description of data management and sharing of data and metadata should be be as long as necessary to describe all important details. To support its significance, the Plan should not "count" against the page limitations of the proposed science.

Finally, given that a principal goal of the NIH policy is to "serve the public," we believe strongly that this is a time when the **NIH should require return of aggregate study results** to participants, at least for the results of clinical research, and in plain language understandable to an individual. Absent a cogent reason, these aggregate results should be available to the public. While there are many issues with return of individual results to a participant that require consideration and analysis, summary results of clinical trials and clinical research should be widely available and understandable—and may help to promote public engagement and public trust in the research and scientific enterprise.

Thank you again for the opportunity to comment on this important issue. We believe that the NIH is in a unique position to harness the power of data sharing for the public good, but only if it uses this opportunity to advance the culture of, and infrastructure to support, data sharing.

We are available to discuss our comments with you if that would be helpful and would be happy to work with you on any of the aforementioned items. Please feel free to contact the MRCT Center at bbierer@bwh.harvard.edu, sawhite@bwh.harvard.edu, and mark.barnes@ropesgray.com.

Respectfully submitted,

Barbara E Bierer, MD
Sarah A White, MPH
Mark Barnes, JD, LLM