



PUBLIC RESPONSIBILITY IN
MEDICINE AND RESEARCH

Chair

Natalie L. Mays,
BA, LATG, CPIA

Vice Chair

Suzanne Rivera, PhD, MSW

Secretary

Martha Jones, MA, CIP

Treasurer

Owen Garrick, MD, MBA

Board of Directors

Albert J. Allen, MD, PhD

Elizabeth A. Buchanan, PhD

Holly Fernandez Lynch, JD, MBE

Bruce Gordon, MD

Mary L. Gray, PhD

F. Claire Hankenson,
DVM, MS, DACLAM

Karen M. Hansen

Megan Kasimatis Singleton,
JD, MBE, CIP

Jori Leszczynski, DVM, DACLAM

Vickie M. Mays, PhD, MSPH

Gianna McMillan, DBe

Robert Nobles, DrPH, MPH, CIP

Stephen Rosenfeld, MD, MBA

Ex Officio

Elisa A. Hurley, PhD
Executive Director

Submitted electronically at <https://osp.od.nih.gov/draft-data-sharing-and-management/>

Francis S. Collins, MD, PhD
National Institutes of Health
6705 Rockledge Drive, Suite 750
Bethesda, MD 20892

*RE: Draft NIH Policy for Data Management and Sharing and
Supplemental Draft Guidance*

Dear Dr. Collins:

Public Responsibility in Medicine and Research (PRIM&R) appreciates the opportunity to comment on the National Institutes of Health (NIH)'s Draft Policy for Data Management and Sharing and Supplemental Draft Guidance, published November 8, 2019.

PRIM&R is a nonprofit organization dedicated to advancing the highest ethical standards in the conduct of research. Since 1974, PRIM&R has served as a professional home and trusted thought leader for the research protections community, including members and staff of human research protection programs and institutional review boards (IRBs), investigators, and their institutions. Through educational programming, professional development opportunities, and public policy initiatives, PRIM&R seeks to ensure that all stakeholders in the research enterprise understand the central importance of ethics to the advancement of science.

PRIM&R strongly agrees with the NIH that sharing data resulting from taxpayer funded research enhances the value of that research, advances the pace of scientific discovery, and, in the case of human subjects research (which will be our focus), maximizes the contributions of human research subjects. We therefore appreciate the NIH's proposal to require NIH-funded researchers to provide a comprehensive plan describing how scientific data will be managed and shared before the launch of a study. **However, we note that NIH's draft Policy for Data Management and Sharing does not articulate a mandate to share such data. We strongly urge the NIH in the final policy to make a clear statement requiring researchers in**

both pre-clinical and clinical research to share their data, unless the agency determines that there is a compelling scientific, ethical, and/or logistical reason to not do so.

Evidence suggests that research subjects are eager to see their data shared and their contributions put to the best use.¹ Even individuals with rare diseases believe their research data should be made available to outside researchers,² despite the heightened privacy risks associated with being part of a smaller or more easily identified population. People participate in research in large part because they believe their contributions will advance science, which is more likely when more researchers are able to access and analyze their data. While there are of course ethical reasons not to share data in some cases—we explore some of these human subject research concerns below—**we believe there should be a rebuttable presumption that data will be shared**. The fact that data will be shared should, in turn, be disclosed in the informed consent process.

1. Review of data sharing plans for privacy and security issues

The NIH has an obligation to facilitate the ethical sharing of data. While we believe the NIH should require that data be shared, we also believe the agency has a simultaneous responsibility to **continue to revisit its practices and policies, in order to set appropriate expectations for the protection of research subjects' data by its grantees**. This should include vetting grantees' proposed data repositories and sharing platforms to ensure they support the secure and ethical sharing of data.

Deidentification is one privacy risk mitigation strategy currently discussed in the supplemental draft guidance. However, it is dangerous to think that deidentification will sufficiently protect research subjects' privacy interests, given that it is no longer possible to guarantee that data will remain permanently deidentified. At the very least, this fact should be appropriately communicated to grantees, oversight bodies, and other relevant stakeholders in both the final policy itself as well as any supplemental draft guidance the NIH develops. We also encourage the NIH to **think creatively about what additional risk mitigation strategies it might suggest**.

According to the current NIH proposal, data management and sharing plans would be required only once an application has gone through peer review and received a "fundable score." Review of submitted data management and sharing plans will, then, be done by individual program officers throughout the year, following the NIH grant cycle. **We urge the NIH to take additional steps to supplement and support this review process**. One option would be to convene a technical review group that includes individuals who are independent from the NIH and its grant recipients, which could more fully assess and

¹ [Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing](#). Mello, M., Lieou, V. & Goodman, S. (2018). *New England Journal of Medicine*.

² [Share and Protect Our Health Data: An Evidence Based Approach to Rare Disease Patients' Perspectives on Data Sharing and Data Protection - Quantitative Survey and Recommendations](#). Courbier, S., Dimond, R., & Bros-Facer, V. (2019). *Orphanet Journal of Rare Diseases*.

address data security and privacy issues.³ This technical review group could draft guidance documents that program officers could then use to review individual plans, thus standardizing reviews of data security and privacy issues across projects.

Alternatively, **the NIH could consider using such a technical review group as a centralized review entity that would weigh in on the merits of individual data management and sharing plans.** Such a group would be better equipped than individual program officers, or institutions' IRBs, to ensure that research subjects' privacy and security interests are protected. We acknowledge that the proposed policy and timing of review does not currently provide an opportunity for such a robust review process, but believe this approach would be of great benefit to both investigators who create data management and sharing plans, as well as the IRBs who review them.

The aforementioned guidance documents could also be made public and shared with the research oversight community which is struggling with the complexities of this new domain. While IRBs are increasingly aware of the privacy and security risks associated with the sharing, storage, and aggregation of scientific data, most do not have access to privacy and security experts who can advise them on the full range of issues or, most importantly, their mitigation. Ideally IRBs might work with computer scientists and engineers at their respective institutions to identify and respond to basic privacy and security trends; however, the current lack of funding support for such interdisciplinary collaboration makes this approach unlikely.⁴ Until there is a shift in funding incentives, or the field of experts in differential privacy grows, we encourage the to NIH lead the way by creating robust mechanisms for reviewing data management and sharing plans for security and privacy concerns.

2. Areas for further guidance

In response to the NIH's request for areas in which further guidance is needed, PRIM&R suggests the agency offer **specific guidance on the ethical issues involved in data sharing for the research oversight community, including IRBs.** Such guidance should help IRBs ensure that participants are adequately informed of the limits of deidentification and include clear recommendations for how both the facts about data sharing and its inherent risks should be conveyed during the informed consent process. The Common Rule now requires informed consent to include a statement when data collected during a research project will be deidentified for subsequent research use, including that further consent will not be sought for such use. We believe this statement is likely inadequate, given the limits of deidentification when data sets can be aggregated, and encourage the

³ Given the many shortcomings of deidentification, the NIH should consider the merits of differential privacy, which is currently the best option for eliminating any identifying features in a dataset, and consider incorporating differential privacy expertise. Because the number of differential privacy experts is small, we urge the agency to capitalize on their stature in the research field and contract with the few number of experts in this space accordingly.

⁴ [Credit Data Generators for Data Reuse](#), Pierce, H., Dev, A., Statham, E., & Bierer, B. (2019). *Nature*.

NIH's future guidance to better address how to communicate with prospective subjects about the realities and risks of data sharing.

Given the growing number and complexity of issues that institutional research oversight bodies will need to understand and monitor as data sharing efforts expand, we also urge the NIH to revisit the Supplemental Draft Guidance on Allowable Costs for Data Management and Sharing's language on facilities and administrative costs. Currently, the draft guidance states, "Budget estimates should not include infrastructure costs typically included in institutional overhead (e.g., Facilities and Administrative costs)." However, institutional overhead costs may rise with increased efforts to share data; as such, **current Facilities and Administrative allowances may be insufficient to cover increased institutional overhead costs.**

Relatedly, although the agency proposes that NIH budget requests may include costs tied to data curation, making data available in repositories, and local data management considerations, we note that many institutions and research investigators do not have the expertise needed for such efforts. At a minimum, the **NIH needs to provide potential grantees with examples of what kinds of costs they should make requests for**, e.g. what kinds of technology might need to be in place to ensure such efforts are successful. We are concerned that without an explicit, and more descriptive, acknowledgement of what these costs might look like, grantees might not make the appropriate requests for the funding needed for important privacy and security measures.

We also request more **clarification on whether grant funds may be requested in budgets or used for the costs associated with the continued storage and sharing of the data after the research has concluded.** It is presently unclear how researchers would be able to cover the annual costs of a data repository or the costs for deidentifying or processing data for sharing years after the grant is over. Relatedly, the **NIH should issue more guidance about how long they expect data to be available after the grant funding ends.**

3. Other issues

Given the number of complex matters we detail above, we again suggest that the proposed **two-page cap for data management and sharing plans is likely to be impracticable.**

Finally, it would be helpful to get some **clarification from the NIH about the relationship between this NIH-wide policy for Data Management and Sharing Policy and the policies that may be promulgated by specific NIH institutes, centers, and offices.** How much discretion will the separate institutes and centers have to create their own requirements, and how much can those requirements go beyond the NIH-wide policy? While there are no doubt good reasons to allow individual institutes to put in place additional rules, for instance, to protect special populations or particularly sensitive data, **we hope the NIH will consider the logistical difficulties and potential burdens of a**

dataset being subject to a number of different jurisdictions, and encourage harmonization of policies across the NIH as much as possible.

Thank you again for the opportunity to comment and for the NIH's continued work on this important issue. We greatly appreciate that the draft policy includes more language regarding the need to protect the rights and interests of research subjects than the 2018 RFI on the topic, and we hope our comments on the current draft policy will be useful in your next stage of policymaking in this area. PRIM&R stands ready to provide any further assistance or input that might be useful. Please feel free to contact me at 617.303.1872 or ehurley@primr.org.

Respectfully submitted,

A handwritten signature in black ink, reading "Elisa A. Hurley". The signature is written in a cursive, flowing style.

Elisa A. Hurley, PhD
Executive Director

cc: PRIM&R Public Policy Committee, PRIM&R Board of Directors