UNIVERSITY OF CALIFORNIA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

OFFICE OF THE VICE PRESIDENT – RESEARCH AND INNOVATION

OFFICE OF THE PRESIDENT
1111 Franklin Street, 11th Floor
Oakland, California 94607-5200

Submitted electronically via: https://osp.od.nih.gov/rfc-draft-supplemental-information-to-the-nih-policy-for-dms/

June 27, 2022

Office of the Director
NIH Office of Science Policy
9000 Rockville Pike
Bethesda, MD 20892

RE: UC Comments in Response to the NIH Request for Public Comments on DRAFT Supplemental Information to the NIH Policy for Data Management and Sharing: Protecting Privacy When Sharing Human Research Participant Data (NOT-OD-22-131)

To Whom It May Concern:

I write on behalf of the University of California (UC) system with regard to the NIH Request for Public Comments on Draft Supplemental Information to the NIH Policy for Data Management and Sharing: Protecting Privacy When Sharing Human Research Participant Data (NOT-OD-22-131) issued on May 12, 2022.

The UC system is comprised of ten research-intensive campuses, six medical schools, and three affiliated U.S. Department of Energy national laboratories. As a system, UC receives approximately $6 billion annually of extramural awards to support research conducted throughout all UC locations, and generally receives five to six percent of the NIH annual appropriations for research.[1]

UC understands the importance of finding the right balance between protecting research participant data and advancing research and innovation. We appreciate the NIH's work in developing a framework for protecting the privacy of research participants when sharing data under the NIH Data Management and Sharing Policy. We support NIH's effort on developing this framework as a set of considerations rather than requirements. Provided below, UC offers specific comments on each of the sections of the framework. In addition, we encourage NIH to continue to develop and support field-specific data repositories that would assist researchers in their data management and sharing efforts.

[1] University of California Accountability Report 2021: https://accountability.universityofcalifornia.edu/2021/chapters/chapter-9.html

Comments on Draft Operational Principles for Protecting Participant Privacy When Sharing Scientific Data

General Comment

Some of the principles listed in this section seem to assume that all data covered by the NIH Data Management and Sharing Policy are that for which informed consent has not been waived. Under the Federal Policy for the Protection of Human Subjects (i.e., the Common Rule), the IRB may approve an informed consent process that waives informed consent. While in UC's experience, most NIH-funded or supported studies do require that informed consents be obtained, there are some studies that meet the requirement to waive all or some elements of informed consent. We suggest that NIH consider the ways in which this may impact the draft operational principles. For example:

- Principle 1 describes the need to "protect the privacy and confidentiality of every participant *as described in the informed consent*" (emphasis added). We recommend adding an additional sentence at the end of this section, such as: "In cases where consent may not have been sought, researchers and institutions should consider the expectation of privacy associated with the original data and the sensitivity of the research."
- Principle 2 acknowledges that some data may be shared without informed consent. UC suggests simple revisions to this language to better address the bi-lateral considerations:
    - Researchers and institutions should proactively assess appropriate protections for sharing scientific data from participants, including determining whether sharing should be restricted through controlled access.[4] This is especially important where other requirements may not be applicable, such as ~~regardless of whether~~ where the data do not meet technical and/or legal definitions of "de-identified" ~~and~~ or can legally be shared without additional protections (e.g., because the research does not meet the definition of "human subjects research" under the Common Rule).
- Principle 3 encourages a robust consent process that speaks to future data use and sharing. The NIH may wish to recognize at the end of this section that there may be instances when the Common Rule allows the IRB to waive all or some elements of informed consent.

Comment on Draft Principle 1

Given the subsequent discussion of data sharing and use agreements, UC suggests mentioning them in Principle 1. For example, we recommend the following:

- NIH and the institutions it funds are obligated and required to protect the privacy and confidentiality of every participant as described in informed consent and in line with all applicable laws, policies, ~~and~~ regulations, contracts, and agreements."

Comment on Draft Principle 5

Draft Principle 5 as written speaks to data collection, rather than data sharing and may be outside of the scope of the NIH Data Management and Sharing Policy. UC suggests either incorporating some of the information from Principle 5 with Principle 2 or revising the information so it addresses

management and sharing of data from non-traditional research setting. For example, if NIH incorporates this information with Principle 2, the following can be stated:

- Researchers and institutions should proactively assess appropriate protections for sharing scientific data from participants, including data from non-traditional research settings, such as wearable technology, health apps, social media, consumer reports, and public health surveillance. ~~determining whether sharing should be restricted through controlled access.~~[4] This is especially important where other requirements may not be applicable, such as ~~regardless of whether~~ where the data do not meet technical and/or legal definitions of "de-identified" ~~and~~ or can legally be shared without additional protections (e.g., because the research does not meet the definition of "human subjects research" under the Common Rule). Assessments should include determining whether sharing should be restricted through controlled access.[4]

Comment on Draft Principle 6

Draft Principle 6 recognizes the need for exceptions to sharing data. UC supports this consideration and encourages individual NIH institutes, centers, and offices (ICO) to harmonize around a consistent process for evaluating and permitting such exceptions.

We recognize that the NIH Data Management and Sharing Policy allows researchers the opportunity to update their Data Management and Sharing Plans (Plans) throughout the course of their awards. However, given that researchers are required to submit their Plans to the NIH during the proposal phase of their projects, campus research compliance offices, such as the Institutional Review Boards (IRBs), may have not conducted their reviews prior to proposal submission. It is possible, therefore, that the institutional research compliance offices may not agree to a Plan that was included in a proposal and already submitted to NIH. We recommend that NIH emphasize in Principle 6 the importance of researchers working with their institution to align Plans with research protocols, and that Plan revisions may be necessary due to institutional review requirements.

Comments on Draft Best Practices for Protecting Participant Privacy When Sharing Scientific Data

Comment on Best Practice 1: Ensure Appropriate De-identification

The 2018 revision to the Common Rule outlined the role federal agencies play with respect to the "identifiability" of information. Specifically, agencies implementing the Common Rule are required to:

- Reexamine the meaning of identifiable private information and identifiable biospecimens with consultation with appropriate experts (including experts in data matching and re-identification), within 1 year and at least every 4 years thereafter.
- Assess whether there are analytic technologies or techniques that should be considered by investigators to generate identifiable private information or identifiable biospecimen, with consultation with appropriate experts, within 1 year and at least every 4 years thereafter. Any such technologies or techniques will be included on a list of technologies or techniques

> that produce identifiable private information or identifiable biospecimens that will be published in the Federal Register after notice and an opportunity for public comment. [2]

We request that NIH revise Best Practice 1 to address federal agencies' role in the reexamination of "identifiability" and how the results of such reexamination will shape any requirement to use advanced de-identification methods. Specifically, we recommend that NIH convene an expert panel task force to examine common methods of de-identification, and common new technologies that may produce identifiable data and make the determinations required by 45 C.F.R. § 46.102(e)(7). That process will provide uniformity and expert informed guidance to institutions.

Additionally, regarding the first bullet listed in Best Practice 1, we recommend that the Safe Harbor HIPAA de-identification method not be cited as other methods of de-identification may be a more effective way to de-identify information.

Moreover, regarding the third bullet in Best Practice 1, it is unclear who would make the determination that the need for scientific utility outweighs a data subject's privacy rights. We suggest replacing the last sentence of this bullet point to the following:

- In some cases, scientific utility may be lost if shared data are de-identified. It may consequently be justifiable in certain cases to share scientific data under the DMS Policy that meet a legal or regulatory standard for identifiability.[10] ~~In those cases, data sharing may be subject to particular rules, and researchers should also consider whether other relevant protections should be employed.~~ In those cases, researchers should consult with their institution's research compliance office (such as the privacy office or Institutional Review Board (IRB)), legal counsel, or other appropriate institutional office, to conduct a privacy balancing analysis to evaluate the associated privacy risks to data subjects that would result if identifiable data is used or shared and to make recommendations aimed at mitigating these risks. If, following the privacy balancing analysis, a determination is made to move forward with sharing identifiable data, the researchers should seek to mitigate the privacy impact to data subjects, to the extent possible using other controls and protections.

Comment on Best Practice 2: Establish Scientific Data Sharing and Use Agreements

UC appreciates NIH's recommending the use of data sharing and use agreements and listing key elements to be included in such agreements. We are concerned, however, about the concept of including an institutional certification ensuring that scientific data have been appropriately de-identified. A blanket recommendation for institutional certification in all cases does not account for the wide variety of circumstances. We encourage NIH to recognize that certifications and similar statements (e.g., warranties and representations) are the product of negotiations among the parties to a data use agreement and vary depending on the data, uses, and parties involved (e.g., public entities, private institutions, etc.).

In addition, while a certification may add an additional safeguard to the data sharing workflow, it seems more appropriate as an internal control at the institutional level rather than as part of the data

[2] 45 C.F.R. § 46.102.

sharing and use agreement with a third party. Perhaps this protection can be achieved under Best Practice 1 (first sub-bullet) by way of including language that institutions develop procedures for certification of data de-identification prior to sharing, for example, through an honest broker, data steward, or other individual adequately trained in data de-identification in accordance with the applicable identifiability standards.

Last, UC recommends that the scope of this section be expanded to acknowledge data shared beyond repositories. It may also be beneficial to include language on why the use of standardized data sharing and use agreements are important. For example, UC recommends the following edits:

- **Establish Scientific Data Sharing and Use Agreements.** NIH recommends the use of scientific data sharing and/or use agreements, preferably standardized to promote common understanding of requirements and expectations, when sharing data from participants ~~with and from repositories~~.

Comment on Best Practice 3: Understand Legal Protections Against Disclosure and Misuse

Best Practice 3 is entitled "Understand Legal Protections Against Disclosure and Misuse" but only describes the NIH Certificate of Confidentiality (CoC) Policy. There are a myriad of different legal protections against disclosure and misuse of data (e.g., Health Insurance Portability and Accountability Act (HIPAA), Family Educational Rights and Privacy Act (FERPA), General Data Protection Regulation (GDPR), and various state laws among other rules). Rather than pointing to CoCs as the only available tool for legal protection (a protection which at this point has not been tested in the courts), we suggest that NIH provide a brief overview or compendium of applicable legal protections. Alternatively, if NIH only wishes to focus on the protections afforded by the NIH CoC Policy, then it should rename the title of this best practice.

Comments on Draft Points to Consider for Designating Scientific Data for Controlled Access

We recommend that NIH explain that de-identification alone is not a sufficient method for ensuring participant privacy concerning sensitive data, even if participants have agreed to secondary uses of the data. Sharing of sensitive data must also include an emphasis on establishing agreements and emphasizing legal consequences of disclosure and misuse.

Furthermore, we suggest the following revisions to the provided language in this section:

- Retitle this section to "Points to Consider for ~~Designating Scientific Data for Controlled Access~~ Protecting Participant Privacy When Sharing Scientific Data." This change aligns with the titles of the other two sections. A new introductory paragraph to this section to explain important considerations around sensitive data and that de-identification alone is not a sufficient method for ensuring participant privacy may be helpful.
- Given potential for downstream misuse of data that is shared without restriction, we recommend requiring affirmative consent:
    - In cases where participants explicitly and affirmatively consent to share scientific data without restrictions, it may be appropriate to share data without access controls.

- In number 3, we recommend strengthening the language by replacing "may be" with "often are" as seen below:
    - Cannot be de-identified to established standards or cannot sufficiently reduce the possibility of re-identification. Access controls, among other measures, ~~may be~~ often are appropriate to further mitigate the risk of re-identification.[16]

Other Considerations

In a previous comment letter responding to the NIH Draft Data Sharing Policy (NOT-OD-20-013), UC underscored the unique opportunity NIH has to lead the community by creating field-specific data repositories. Such repositories help the scientific community in ensuring relevant security and privacy concerns are addressed. We encourage NIH to focus on the development, establishment, and funding of additional data repositories into which shared data can be deposited. Such efforts will significantly assist institutions in their data sharing and management efforts.

Thank you for the opportunity to comment. We welcome further discussions on this important topic. If you have any questions concerning these comments, please contact Agnes Balla, Research Policy Manager, at agnes.balla@ucop.edu.

Sincerely,

Deborah Motton, Ph.D.
Executive Director
Research Policy Analysis and Coordination
University of California, Office of the President