COGR

Council On Governmental Relations

An Association of Research Institutions

February 28, 2022

Submitted via Email to: https://osp.od.nih.gov/rfi-updating-the-nih-genomic-data-sharing-policy

Office of the Director
National Institutes of Health (NIH)

RE: Comments to Request for Information on Proposed Updates and Long-Term Considerations for the NIH Genomic Data Sharing (NOT-OD-22-029)

To Whom This May Concern:

The Council on Governmental Relations (COGR) is an association of 200 public and private U.S. research universities and affiliated academic medical centers and research institutes. COGR concerns itself with the impact of federal regulations, policies, and practices on the performance of research conducted at its member institutions.

COGR appreciates the opportunity to respond to the Request for Information on Proposed Updates and Long-Term Considerations for the NIH Genomic Data Sharing Policy ("GDS Policy"). COGR and its member institutions recognize the important role that timely and robust data sharing plays in both scientific advances and efforts to promote reproducibility and research integrity. Yet, we also appreciate that the sharing of data, particularly human genomic data, raises complex privacy and consent issues, and we commend NIH for seeing public input in this area.

We provide our comments on the specific items enumerated in the RFI below for your consideration. Our comments emphasize three major themes: (a) the need for consistency between the definitions and standards for identified and de-identified data that is subject to both the HIPAA Privacy Rule (45 CFR Parts 160 & 164) and the GDS Policy; (b) the importance of informing and educating stakeholders on advantages and risks of data set linkage; and (c) harmonizing certain major elements of the GDS Policy and the Data Management and Sharing Policy ("DMS Policy").

RFI Section I. Maximizing Data Sharing while Preserving Participant Privacy and Preferences

RFI Item 1: De-identification. *The risks and benefits of expanding de-identification options, including adding the expert determination described at 45 CFR 164.514 (b)(1) (the HIPAA Privacy Rule), as an acceptable method for de-identification under the GDS Policy, and whether other de-identification strategies exist that may be acceptable in lieu of HIPAA standards.*

Recommendation 1. COGR supports harmonizing the de-identification options specified in the GDS Policy with those currently available under the HIPAA Privacy Rule by adding the option to use the

expert determination described at 45 CFR 164.514(b)(1) to the GDS Policy. This de-identification option has been long been available under the HIPAA Privacy Rule and specifically requires the expert to consider how the information in question "could be used, alone or in combination with other reasonably available information." Moreover, we suggest that NIH fully harmonize the GDS Policy with the HIPAA Privacy Rule by tying the de-identification methods available for protected health information (PHI) under the GDS Policy directly to those available under the HIPAA Privacy Rule. Such an approach would eliminate confusion in this area and reduce compliance burden.

We support expanding options for de-identification to those not currently available under the HIPAA Privacy Rule for data not subject to the Rule (i.e., non-PHI) provided that the risk of re-identification from such additional method(s) remain(s) as low or lower than that posed by current options. To support harmonization of regulatory requirements, we suggest that if NIH develops additional de-identification methods, it considers coordinating with appropriate units within the Department of Health and Human Services to determine if those methods also could be included under the HIPAA Privacy Rule and any regulations/guidance issued under that Rule.

RFI Item 2: Use of potentially identifiable information. *The circumstances under which submission of data elements considered potentially identifiable to repositories under the GDS Policy would be acceptable, any additional protections (including for security) that would be warranted, and whether there is certain potentially identifiable information that would not be acceptable to submit.*

Recommendation 2. In concept we support the use of "potentially identifiable data elements," however we cannot clearly evaluate this recommendation without understanding the definition of this term. Thus, we recommend that NIH clearly define "potentially identifiable data elements" and cite examples that would distinguish such data elements from identified and de-identified information as set forth under the HIPPA Privacy Rule. Further, we note that consideration of such potentially identifiable data elements might better be addressed though the currently existing Common Rule requirement for agencies to periodically revisit the definition of "identifiable private information." [45 CFR § 46.102(e)(1)(ii)(7)(i)]. Finally, we urge NIH to ensure that any definition of potentially identifiable data elements does not undercut or contradict the definition of de-identified PHI subject to the HIPAA Privacy Rule or impose additional requirements on such information, or on non-PHI that is considered de-identified under any additional de-identification methodologies added under the GDS Policy.

RFI Item 3: Data linkage. Whether the GDS Policy should permit data linkage between datasets that meet GDS Policy expectations (e.g., data obtained with consent for research use and de-identification), and whether the GDS Policy should support such linkages to datasets that do not meet all GDS Policy expectations (e.g., data may have come from a clinical setting, may not have been collected with consent, may retain certain potentially identifiable information). Feedback is also requested on risks and benefits to any such approaches.

Recommendation 3. We support data linkage between datasets that meet GDS Policy expectations, either de-identified or with potential identifiers and consent for linking. We also believe that it may be appropriate to consider data linkage between datasets that do not meet all GDS Policy expectations (e.g., use of clinical data without consent), provided there is sufficient scientific justification and use of appropriate additional protections (e.g., review by data access committee, agreement not to use linked data to identify/contact individual participants, etc.). We also recommend the employment of data use agreements for the use of all linked data, whether obtained by consent or not.

RFI Item 4: Consent for data linkage. Whether data linkage should be addressed when obtaining consent for sharing and future use of data under the GDS Policy, as well as in IRB consideration of risks associated with submission of data to NIH genomic data repositories. And if so, how to ensure such consent is meaningful.

Recommendation 4. COGR understands the importance of educating research participants via the informed consent process about the way their data may be used, including the potential for data linkage, when obtaining consent for sharing and future use. We support addressing data linkage in the informed consent process, and to facilitate this process, we recommend that NIH further engage with and educate stakeholders (e.g., investigators, IRBs, focus groups) on the benefits and potential risks of data linkages and the best methods for communicating these concepts to research participants. Sample consent language along these lines developed by NIH should be provided for public review and comment prior to requiring any specific consent language/forms Further in this vein, we note that in its "Discussion of Public Comments on the Draft NIH Policy for Data Management and Sharing" in the Final NIH Policy for Data Management and Sharing (NOT-OD-21-013) NIH indicated that it "…intends to develop resources to help researchers and institutions in communicating the intent to share data with prospective research participants." Expanding this work to include information specific to data linkage in the context of genomic research would be appropriate.

RFI Section II. Expectations for Alternative NIH-Supported Genomic Data Management and Sharing Resources that Store Human Genomic Data

RFI Item 5: Data management and sharing principles for NIH-supported resources

a. Any aspect of the principles described for Data Submission.
b. Any aspect of the principles described for Data Access.
c. Any aspect of the principles described for Data Security.

Recommendation 5. We support these principles and have nothing further to add.

Section III. Policy Harmonization

RFI Item 6: Harmonizing GDS and DMS Policies. Any aspect of the approach to harmonize GDS and DMS Policies and Plans described above, including for non-human genomic data.

Recommendation 6. We note the generality of this item and the resulting difficulty in evaluating its impact given the substantial differences among the various types of data covered by GDS and DMS. Nonetheless, we see significant benefit to harmonizing GDS and DMS Policies with respect to the following components: required elements of a data management plan, submission of a single plan, identification of sensitive data at time of proposal or application, and NIH process for reviewing submitted plans.

RFI Item 7. GDS and DMS data sharing timelines. Whether the continued use of earlier submission expectations for human genomic data in the GDS Policy (e.g., submission of human data within three months of data generation) is needed, or whether timelines should be harmonized with the DMS Policy expectations (i.e., sharing of data no later than the time of publication or at the end of the performance period, whichever comes first), as described in the proposal above.

Recommendation 7. We recommend that the timelines be harmonized with the DMS Policy expectations with a caveat that the researcher can share data earlier at his/her own discretion. Further, we note individual institutes and centers currently may require data sharing on specific schedules, thus permitting tailoring to unique circumstances. We ask that NIH harmonize institute/center policies with the Final NIH Data Management and Sharing Policy to the maximum extent practicable.

RFI Section IV. Long-Term Consideration of the Scope of GDS Policy

RFI Item 8. Types of research covered by the GDS Policy.

a. Whether there are other types of research and/or data beyond the current scope of the GDS Policy that should be considered sensitive or warrant the type of protections afforded by the GDS Policy (e.g., with consent for future use and to be shared broadly, as well as IRB review of risks associated with submitting data to NIH), even when data are de-identified.

Recommendation 8.a**:** COGR believes that NIH should consider extending the GDS Policy to other types of "omics" research (e.g., proteomics, transcriptomics, etc.); however, our previously expressed concerns regarding identification, de-identification, privacy, and consent also would apply to any extension of the policy in these areas.

b. Whether small scale studies (e.g., studies of fewer than 100 participants) and those involving other data types (e.g., microbiomic, proteomic) should be covered under the GDS Policy, and if training and development awards (e.g., F, K, and T awards) should be covered by the GDS Policy ("Implementation of the NIH Genomic Data Sharing Policy for NIH Grant Applications and Awards," NOT-OD-14-111).

Recommendation 8.b**:** Given the greater risks of inadvertent identification and possible stigmatization for participants in smaller studies, COGR believes that data from smaller scale studies should not be covered by the GDS Policy unless consent is obtained from all participants for contemplated data sharing and data linkage. If NIH determines that it will include data from all small-scale studies within the scope of Policy, COGR believes that such data should be subject to all GDS Policy requirements and that linkage to data sets that do not meet GDS Policy requirements should be prohibited absent consent.

Additionally, COGR recommends that the funding of an individual through an F, T or K award should not automatically trigger the application of the GDS Policy. Neither F or T awards include research dollars, and although K awards may include some research dollars, the primary aim of the grant is to fund an individual. Thus, it seems inappropriate for the GDS Policy apply to "reach through" and apply to research projects that are funded through other sources.

c. Whether NIH-funded research that generates large-scale genomic data but where NIH's funding does not directly support the sequencing itself should be covered by the GDS Policy.

Recommendation 8.c.: For consistency purposes and because the privacy and consent issues for human genomic data remain the same, we support including within the GDS Policy's scope NIH-funded research

that generates large-scale genomic data even though NIH funding does not directly support the sequencing in such studies.

RFI Item 9. Data sharing expectations under the GDS Policy. Whether there are other types of research and/or data that warrant the data processing level and timeline expectations established by the GDS Policy (e.g., sharing lower levels of processed data, not just those of sufficient quality to validate and replicate findings as in the DMS Policy). We have no comments regarding this RFI Item.

Conclusion

COGR appreciates the opportunity to respond to the RFI. We value our partnership with NIH on this important topic and look forward to future opportunities to provide input. If you have any questions concerning our comments, please contact Kris West, Director of Research Ethics & Compliance, at kwest@cogr.edu.

Sincerely,

Wendy D. Streitz
President