

UC Office of the President

U.S. Department of Health and Human Services (HHS)

Title

UC Comments in Response to NOT-OD-20-013, "Request for Public Comments on a DRAFT NIH Policy for Data Management and Sharing and Supplemental DRAFT Guidance"

Permalink

<https://escholarship.org/uc/item/5wz509zr>

Journal

N/A

Author

UC Office of President

Publication Date

2020-01-10



OFFICE OF THE VICE PRESIDENT – RESEARCH AND INNOVATION

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

January 10, 2020

Dr. Andrea Jackson-Dipina
Director of the Division of Scientific Data Sharing Policy
National Institutes of Health
Office of Science Policy
6705 Rockledge Drive, Suite 750
Bethesda, MD 20892

RE: UC Comments in Response to NOT-OD-20-013, “Request for Public Comments on a DRAFT NIH Policy for Data Management and Sharing and Supplemental DRAFT Guidance”

Dear Dr. Jackson-Dipina:

I write on behalf of the University of California (UC) system with regard to the Request for Public Comments on a DRAFT NIH Policy for Data Management and Sharing and Supplemental DRAFT Guidance issued on November 11, 2019.

The UC system comprises 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. UC generally receives 5 to 6 percent of the NIH’s annual appropriations for research, making UC the largest single recipient of NIH funding for academic research.

The UC believes that the curation and sharing of research data offers benefits to the larger research community and advances public knowledge. UC supports the NIH’s effort to facilitate data sharing and appreciates the NIH’s recognition of the challenges that come with regulating data generated by a broad research community. There are many aspects of the NIH draft policy that we support and were pleased to see included in the policy such the incorporation of compliance checking and routine reporting of data management and sharing plans, the ability to request NIH funding to support data management, and the ability to make data management and sharing plans publicly available. While UC generally agrees with the NIH’s draft data management and sharing policy, however, we believe that in order to make this policy a success, the NIH should make developing templates, further resources and robust tools to better facilitate data sharing, particularly across scientific disciplines a high priority.

Comments on specific aspects of the draft policy are provided below.

I. Purpose Section

UC appreciates the NIH's philosophy that underlies the draft policy as discussed in the Purpose section. It serves as a helpful reminder that investigators are not conducting their work within a vacuum. Proper research data management will inevitably include not just researchers, but also various institutional research administrative units and departments, along with research funders, regulators and journals equally invested in promoting the dissemination of knowledge within the constraints of legal, ethical, technical, security, and privacy considerations. We suggest that the NIH include in this section a recognition of data preservation and sharing as part of the entire research lifecycle.

II. Definition Section

The definition of "scientific data" requires clarification. According to the draft policy, scientific data is "the recorded factual material commonly accepted in the scientific community as necessary to validate and replicate research findings regardless of whether the data are used to support scholarly publications." The definition is very broad, but at the same time, excludes specific sources of information. This ambiguity could lead to confusion about how best to interpret "scientific data" and implement proper data management and sharing plans. For example, the definition of scientific data incorporated into the draft policy explicitly excludes laboratory notebooks. Information contained in laboratory notebooks, however, might reasonably be seen as "necessary to validate and replicate research findings." Given its breadth, the proposed definition of scientific data seems open to the interpretation that almost all data collected for a study counts, and therefore should be shared. UC asks that NIH provide clarification on this in addition to examples of accepted data standards.

Furthermore, the NIH's expectation that reasonable efforts will be made to digitize all scientific data be digitized raises concerns to UC. Not all scientific records must be digital to be useful. Digitizing also imposes administrative burden upon institutions, and creates risk, both in loss or error in translation, as well as in of the disclosure or use of sensitive material, including medical information. If external parties are necessary to digitize such data, the risk of loss, error or exposure becomes more pronounced. Because of the potential for heightened risk, the decision to digitize data should be left to the principal investigator's discretion. UC recommends that this expectation be removed.

III. Effective Date

UC asks that the NIH choose a policy implementation date that will allow the research community to prepare sufficiently for the policy change and to consider a reasonable embargo period to advance intellectual property in its Supplemental Draft Guidance for data sharing plan elements (*Section 4, Data preservation, Access and Associated Timelines*). We recommend a minimum implementation date of one year after the release of the final policy, with a delay in enforcement actions for at least one year after the implementation deadline. Any determination of non-compliance should follow well-defined and transparent criteria.

IV. Requirements

While UC supports the requirement of a data management and sharing plan for NIH-funded or conducted research, we are concerned about varied supplementary information requirements requested by individual NIH Institutes, Centers, and Offices (ICOs). To minimize confusion and administrative burdens, we strongly urge trans-NIH coordination of these supplemental requests and listing ICO-specific requirements as part of centralized resources associated with the final data management and sharing policy.

V. Data Management and Sharing Plans

UC appreciates the proposal to collect data management and sharing plans as part of Just-in-Time (JIT) documentation for extramural awards. Requiring submission of the plan at the JIT phase rather than at the proposal stage minimizes administrative burden for both the applicant and peer reviewers. Shifting the review of plans to NIH staff members rather than volunteer reviewers will also make the process more uniform and streamlined, provided NIH staff are properly trained in reviewing data management plans. One potential drawback of submitting data management plans during JIT, however, is that accurately budgeting for data management costs at the time of proposal will be challenging as details of the plan remain to be finalized with the NIH Program staff at JIT. UC, therefore, recommends that the NIH allow additional data management costs to be added to the budget at JIT based on the final negotiated data management plan. We also recommend an option that allows grantees to appeal ICO-mandated data sharing requirements to the NIH Policy Office should the requirements be considered unreasonable or inappropriate by the grantee, without fear of reprisal.

Furthermore, UC strongly urges the NIH to provide data management and sharing templates. Rather than having each researcher develop their own plans in two pages or less, having a common template will streamline the process and reduce burden for both submitters and reviewers. It would also be useful if NIH published sample plans. Within these templates, we recommend that NIH provide minimum requirements for researchers to include in a data management and sharing plan, such as data type, standards and metadata, plans for data preservation, and projected data accessibility. In addition, the template could include the usage of persistent identifiers and tools to generate machine-readable and actionable data management plans that support compliance checking and will facilitate the interoperability of data between systems. At a minimum, it would be useful if the NIH policy required the use of ORCID ID's for all individuals mentioned in the data management and sharing plans and for the use of persistent identifiers for all research datasets.

Finally, UC recommends that the NIH provide guidance on appropriate ways to maintain sensitive data. Not only is this administratively burdensome, but it also introduces a dependence on the researcher (and their current contact information) that undermines the goal of long-term data accessibility. NIH should recommend restricted access repositories as well as other kinds of repositories that so that the researchers understand what qualifies as an "established repository." In this context, the NIH has the unique opportunity to lead the community by creating field-specific data repositories that have the added benefit of ensuring relevant security and privacy concerns are addressed.

VI. Compliance and Enforcement

UC appreciates that the NIH will allow researchers to make updates to their data management and sharing plan. In practice, data management plans may evolve over time as repositories, technologies, and data standards change. We are pleased to see some degree of flexibility built into the plans to address unforeseen issues.

The policy is not clear how adherence to a data management and sharing plan is verified at the end of a project. Requiring a final report on a data management and sharing plan would be an excellent way to enforce compliance and ensure that a researcher has followed through on plans for sharing and depositing data. Adding details how compliance is monitored would give the policy some enforcement capabilities and make sure it is taken seriously.

Additionally, we wondered who would be responsible for reviewing and approving data management and sharing plans during the initial review process. It may be useful to specify the required knowledge and expertise that will be required of reviewers to ensure that they have the necessary backgrounds to effectively analyze and comment upon data management and sharing plans.

VII. Supplemental Draft Guidance – Allowable Costs

UC appreciates the NIH's recognition of the costs associated with data management and sharing and applauds the inclusion of the supplemental guidance defining possible allowable costs. However, UC is concerned that the guidance only addresses those costs incurred during the term of the award but does not address costs associated with long-term data retention and accessibility, such infrastructure costs related to storing data. We strongly urge the NIH to allow researchers to budget for long-term data curation and preservation costs, or at a minimum clarify that grantee institutions may pre-pay from their awards these long-term costs. We would recommend that if these long-term costs are not permitted on a grant-by-grant basis, that the agency offers additional supplemental funding to institutions to develop infrastructure for data management and storage.

Increasing data management and sharing activities often requires significant support from personnel outside of the traditional laboratory environment, including librarians and data scientists, to provide the necessary expertise and guidance needed to comply with a data sharing policy and build good data management practices into an investigator's research process. NIH should strongly consider including these additional staff as part of the allowable costs. Again, if this is not doable, it will be necessary for the agency to provide supplemental funding to institutions in building up and maintaining services that support scientific data sharing.

VIII. Supplemental Draft Guidance – Elements of a NIH Data Management and Sharing Plan

UC suggests that the policy more clearly define acceptable timeframes for data sharing, as "timely manner" could be widely interpreted. These could be conveyed as ranges to preserve flexibility.

IX. Other Considerations Relevant to the Draft Policy Proposal

UC appreciates that the draft acknowledges the importance of human subject protections without including a mandate for IRBs to verify or otherwise be a gatekeeper for data sharing and management. The NIH Genomic Data Sharing Policy model of IRB involvement is not appropriate for a broader policy such as this. However, UC does suggest that future NIH guidance emphasize three points:

- It is important that PIs prepare protocols and IRB applications that are consistent with the data management and sharing plans.
- It is important that PIs prepare consent forms that are both consistent with the data management and sharing plans and provide adequate notice to subjects about plans to reuse data.
- Institutions and their IRBs should evaluate policies, instructions, application forms and other similar materials to ensure that the ability to comply with this policy is not inadvertently or unduly restricted.

Thank you for the opportunity to comment on this important issue and we look forward to continued engagement on this issue as the data sharing and management draft policy and other guidance is developed.

Sincerely,



Lourdes G. DeMattos
Associate Director
Research Policy Analysis & Coordination
Office of Research & Innovation



Chris Shaffer
University Librarian and Assistant Vice Chancellor
University of California San Francisco