

Population Association of America Association of Population Centers

Office of Government and Public Affairs

8630 Fenton Street, Suite 722 • Silver Spring, MD 20910
www.populationassociation.org • www.popcenters.org • 301-565-6710 x 1006



Population Association of America President

Dr. John Casterline
Ohio State University

Vice President

Dr. Noreen Goldman
Princeton University

President-Elect

Dr. Eileen Crimmins
University of Southern California

Vice President-Elect

Dr. Sara Curran
University of Washington

Secretary-Treasurer

Dr. Bridget Gorman
Rice University

Past President

Dr. Wendy Manning
Bowling Green State University

Dr. David Bloom
Harvard University

Dr. Jennifer Dowd
King's College, London, UK

Dr. Pamela Herd
Harvard University

Dr. Emily Hannum
University of Pennsylvania

Dr. Jeffrey Morenoff
University of Michigan

Dr. Jenna Nobles
University of Wisconsin, Madison

Dr. Mary Beth Ofstedal
University of Michigan

Dr. Krista Pereira
University of North Carolina

Dr. Zhenchao Qian
Brown University

Dr. James Raymo
University of Wisconsin, Madison

Dr. Jenny Trinitapoli
University of Chicago

Dr. Kathryn M. Yount
Emory University

Association of Population Centers President

Dr. Kathleen Cagney
University of Chicago

Vice President

Dr. Debra Umberson
University of Texas at Austin

Treasurer

Dr. Andrew Foster
Brown University

Secretary

Dr. Jeffrey Morenoff
University of Michigan

January 10, 2020

The below comments are submitted on behalf of the over 3,000 members of the Population Association of America (PAA) (www.populationassociation.org) and the over 40 federally supported population research centers at U.S. based research institutions comprising the Association of Population Centers (APC) in response to 84FR60398, "Request for Public Comments on a DRAFT NIH Policy for Data Management and Sharing and Supplemental DRAFT Guidance."

Our organizations are gratified to see that several of the recommendations that we raised in our [December 7, 2018 letter](#) have been incorporated into the draft policy. To recap that letter, we recommended that the policy: (1) articulate a system for sharing and archiving data extracts; (2) address the management of paradata (i.e., data about the data collection process); (3) reward data collection and sharing by incentivizing data citations; (4) address the costs of data sharing; and, (5) specify a timeline for data sharing.

The draft policy goes a long way to addressing key issues relevant for NIH-funded population scientists and demographers. For example, our organizations were pleased with the way the latest policy articulates a system for sharing and archiving data extracts in such a way that offers flexibility across fields and gives NIH Institutes discretion on how to implement it. As the draft rightfully points out, various fields have developed data and metadata standards and that, where possible, these should be used. It also allows fields that have not developed such standards to name standards, which will likely spread best practices to fields currently without standards.

Our organizations were also happy to see the emphasis on the management of data, including the management and sharing of metadata. In general, the current draft is consistent with the FAIR (Findable, Accessible, Interoperable, and Re-usable) data principles.

We also support the draft plan and supplemental draft guidance to use established repositories when possible. These repositories are important institutional commitments to ensuring that data remain accessible even as the technology used for data storage and data finding changes. The investments that Institutes, such as NIA, NICHD and NIDA, have made in such repositories through the Inter-university Consortium for Political and Social Research (ICPSR) benefits the population sciences as well as many other disciplines. NIH grantees should be encouraged to use these repositories because they provide

avenues for data sharing in perpetuity and can provide secure access to restricted data, which is otherwise a major challenge.

We also were pleased that the dissemination of restricted data is mentioned explicitly, along with the need to safeguard respondents' identities and sensitive information. We encourage NIH to give greater emphasis to this matter in its final policy and state that the dissemination of restricted data, with appropriate protections of the privacy and confidentiality of respondents, should be standard practice, rather than an elective option.

We appreciate that the draft plan largely meets the needs of the research communities we represent and addresses the central challenge that resources and correct incentives are necessary to ensure good data stewardship and extensive data sharing occurs. However, we believe the policy could be expanded to include the following recommendations, which we included in our [December 2018 letter](#).

- In our earlier comments, the PAA and APC recommended rewarding data collection and sharing by incentivizing researchers to use citations to acknowledge the work that was done to curate and share data files. We continue to believe that the final policy should provide clear citation guidance, including recommendations for how to cite secondary data that are created and shared with the research community. Both primary and secondary data that are eligible for citation should receive an NIH data catalog record analogous to a PMID or PMCID (in addition to be cataloged using DOIs or other persistent identifiers). Such a mechanism aligns the incentives of academic rewards for principal investigators with the scientific community's data sharing needs.
- We also recommended that the final plan address the cost of data sharing. We recognize that the supplemental draft guidance on allowable costs for data management and sharing stipulates some expenses associated with data management and sharing could be budgeted in grant applications. We do not feel this goes far enough. Simply allowing these expenses does not ensure that researchers will budget sufficient resources for data sharing, especially when faced with funding caps that make it challenging to fully fund data collection and analysis. We continue to believe that funding for data management and data sharing needs to be provided outside the regular budget process, by separate supplements to cover data sharing costs and/or separate data sharing and archiving grants similar to a successful NICHD R03 program ([PAR-16-149](#)). An alternative would be to allow each NIH grant to have a separate budget for data management and sharing beyond any existing budget caps.
- The draft guidance is vague about what the expectations are regarding completeness and comprehensiveness of the data to be shared. There are

suggestions that investigators don't need to share every piece of data they collect and can make subjective decisions about what to include and exclude. There is a worry that investigators could withhold important data, either deliberately or inadvertently, and thus meet the letter but not the spirit of the policy. We suggest that NIH include language in its policy stipulating that all of the data and measures collected under its grants be shared with the broader research community, subject to standard restrictions related to protection of respondents' privacy and the confidentiality of the information they have provided.

Thank you for considering these comments. For more information, please contact Ms. Mary Jo Hoeksema, Director, PAA/APC Government and Public Affairs, at maryjo@popassoc.org.