



December 19, 2015

Steven Hirschfeld, MD PhD

*Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD)

Sent via email to: [hirschfs@mail.nih.gov](mailto:hirschfs@mail.nih.gov)

Dear Dr. Hirschfeld,

On behalf of the Population Association of America ([www.populationassociation.org](http://www.populationassociation.org)) and Association of Population Centers ([www.populationcenters.org](http://www.populationcenters.org)), we are pleased to comment on RFI, HD-15-034, *Request for Information on Enhancing Timely data Sharing from NICHD*. PAA is the premier professional, scientific society for more than 3,000 behavioral and social scientists—including demographers, sociologists, economists, epidemiologists and statisticians—who study the implications of population change. Our members conduct research and train young scientists at U.S. universities and independent research organizations. The APC is composed of approximately 40 federally funded, interdisciplinary population research centers nationwide.

The population research community has long supported sharing data collected for research purposes. Therefore, our organizations applaud NICHD for broadening its data sharing expectations. Because the proposed policy may be reaching communities less familiar with the value and challenges of data sharing, we propose some clarifications for your consideration.

First, it is important to make explicit that data must be made available *along with relevant documentation*. Some documentation is at the meta level so that other researchers understand the underlying sampling and recruitment strategies. Documentation is also at the level of individual indicators. Without adequate documentation, just sharing the basic data is not useful.

Second, the presumption seems to be that other researchers would be able to do their own analyses with the data. We recommend stating more explicitly that this sharing does not imply that the original researcher can still control what analyses are done or control the publication of results. On the other hand, publications by researchers who have received such data should reference the source of the data. The RFI uses the phrase *freely available*, which we presume implies that it is free of constraints by the original authors. However, it can also refer to whether there can be a charge for the data. It would be useful if the guidance were explicit as to whether NICHD anticipated supporting the costs of data sharing either within individual projects (intramural or extramural), or providing financial support to data archives, or something else. At present, some data are shared with only a nominal data processing fee, but others are completely without charge.

Third, there are many ways to share data. There are advantages to having data submitted to a data archive that will ensure standardized practices of documentation, sharing agreements, and ensure that the data have been cleaned. Placement in an archive also helps ensure continuity

even after the original investigators have moved on or can no longer afford the necessary support. Data sharing is important and valuable, but it is not free. It was not clear whether NICHD was anticipating supporting a new data archive, but if so, then we encourage that this be viewed as a (valuable) supplement to existing data archives.

Fourth, while we would presume that demographic data is included under the rubric of clinical data, it would be useful for the guidance to be explicit about what data are covered by the policy. Data gathered under grants is clearly not the property of the federal government, and it would not be appropriate to require the data be held in a governmental repository. The goal should be effective sharing of data, not a single strategy or approach or archive.

Fifth, the timing of the release of data is plausible, but clearly there will be variability. Some small studies will be able to release study data in advance of the 12-month goal, whereas for other studies that timeline may be unrealistic. However, presenting the data 12- months from publication of the basic analyses is a good benchmark. If an investigator proposes to deviate from that approach, they must justify their own plan. The policy should also provide sufficient time for the grantees to go through the IRB review process given IRB approval will most likely be necessary to share the data. Further, subject confidentiality should be not sacrificed in order to meet a data sharing deadline.

Finally, we commend the Institute for explicating stating that sharing data tables is not sufficient. Sharing data must mean sharing data at the individual level. The demographic community has long supported sharing research data with proper protections for privacy and confidentiality. While some aspects of data sharing were outside the scope of the RFI, we encourage careful consideration of tools to protect research subjects/participants from deductive identification.

Thank you for considering our views. Please do not hesitate to contact us if our organizations can be of further assistance as you refine and implement the proposed policy.

Sincerely,



Steve Ruggles, Ph.D.  
President, Population Association of America



Lisa Berkman, Ph.D.  
President, Association of Population Centers