Data sharing in biological anthropology: Guiding principles and best practices

Trudy R. Turner | Connie J. Mulligan

1Department of Anthropology, University of Wisconsin – Milwaukee, Milwaukee, Wisconsin
2Department of Anthropology, University of Florida, Gainesville, Florida
3Genetics Institute, University of Florida, Gainesville, Florida

Correspondence
Trudy R. Turner, Department of Anthropology, University of Wisconsin – Milwaukee, Milwaukee, WI.
Email: ajpa@uwm.edu

Funding information
National Science Foundation, Grant/Award Number: 1826885

In 2017 the AAPA established an ad hoc committee on data access and data sharing; committee members were Trudy Turner (Chair), Connie Mulligan (Co-Chair), Doug Boyer, Eric Delson, and William Leonard. With support from a NSF grant to the association, the committee led a workshop on data sharing in biological anthropology on February 8–9, 2019 in Milwaukee, WI. Forty participants representing all aspects of the field were able to reach consensus on data sharing. The goal of the workshop was to formalize a set of guiding principles and best practices in order to increase and normalize data sharing within biological anthropology. Here we present the results of the workshop and we invite comments through Letters to the Editor of AJPA from the biological anthropology community.

1 | GUIDING PRINCIPLES FOR DATA SHARING IN BIOLOGICAL ANTHROPOLOGY

We believe data sharing is essential for the advancement of the field of biological anthropology. Data used for research in our field are collected from physical materials and phenomena that are temporally or geographically limited and may not be able to be collected again. These data often represent unique evolutionary, ecological and cultural histories. We have a responsibility to maintain and share these resources for purposes of scientific integrity, world heritage, and advancing knowledge.

Data sharing improves research transparency and promotes replicability, reproducibility, and correction. We believe data sharing leads to more equitable access to resources and responsible use of research materials and funded information. Public sharing of data also brings the discipline into compliance with requirements from funders and publishers. As a field, we should strive to make our research data “findable, accessible, interoperable, and reusable”, or FAIR, as outlined by the H2020 programme guidelines on FAIR Data Management in Horizon 2020, July 26, 2016 (http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf).

We maintain that open and public access of data is the ultimate goal, but limitations exist with respect to specific data sets, institutional policies, cultural heritage considerations, and international conventions. Thus, data access should be “as open as possible, as closed as necessary” (H2020 Programme, 2016).

2 | BEST PRACTICES FOR DATA SHARING IN BIOLOGICAL ANTHROPOLOGY

- Data refer to facts or pieces of information and do not include collected samples. Metadata refer to the information that is necessary to understand, analyze, and interpret the primary data, such as accompanying documentary, behavioral, demographic, or historical data.
- At a minimum, data that are used in a publication should be made publicly available no later than time of publication.
  o Ideally, data should be made publicly available in time for the data set identifier (e.g., accession numbers or DOIs) to appear in the publication to ensure easy access to the data.
  o All data required for replicating statistical analyses and underlying summary statistics, including but not limited to raw data, metadata, and a codebook of variables, should be shared whenever possible. Sharing summary statistics is not sufficient.
All digital resources used to generate quantitative data required for replicating statistical analyses, including but not limited to digital imagery, custom software programs, scripts and annotated code, should be shared whenever possible.

Data should be posted in a “trusted” data repository—see https://docs.google.com/document/d/11wj0Y8l6H7OfvGmtvwsX3bwGqs0Durtv8n9ldnIlwdw/edit for a list of recommended repositories. An accessible, enduring, digital file format should be used that is appropriate to each data type or data set.

Authors may embargo access to archived data from a published paper that are referenced by a DOI or other persistent link for up to 6 months post-publication date, when justified.

Ideally, all data (raw, summary, metadata) from a project, that is, not just data used in a publication, should be made publicly available.

Due to issues of participant confidentiality, and ethical and legal considerations, it may not be possible to make all data publicly available at the time of publication, but efforts should be made to make as many data sets publicly available and as open as possible.

The following information should also be shared: the presence and identity of other stakeholders, the conditions of data collection, and the limits these factors may place on archiving and sharing of each data set used or collected in a grant or publication.

Project design should include a clear data management and sharing plan that is in place prior to the start of the project. Data sharing should be viewed over a time horizon related to the length of the research project, such that different parts of a data set may be shared at different times. For example, timelines in a grant proposal might include specific target dates for making particular data available (e.g., metadata, raw data, etc.).

In working with living or descendant communities, consultation with those communities is expected in order to establish the conditions under which data sharing is acceptable. Part of the consenting process should include a conversation about possible sharing of the data and/or any collected samples.

All data sharing plans should conform to local, institutional, national, and international policies and expectations.

Data should be deposited in a public website that is an established and trusted repository that adheres to the FAIR principals (Wilkinson, et al. 2016, Scientific Data, https://www.nature.com/articles/sdata201618). Public websites are preferable to personal websites because they contain multiple data sets that are searchable and include features such as embargoes.

Graduate programs in biological anthropology should include education and training in data management and data sharing strategies and plans. Training should include consideration of ethical issues such as responsible use and sharing of data.

- Responsible use of shared data includes appropriate reference and acknowledgement of source of data and funding agencies, and consultation with representative communities where appropriate. Data integrity is maximized when directly acquired from the data repository rather than second-hand sources.

### 3 | BEST PRACTICES FOR DIGITAL ARCHIVING IN BIOLOGICAL ANTHROPOLOGY

Data archiving is considered separate from data sharing because all data should be archived without exception and stakeholder or institutional or confidentiality restrictions should not affect archiving practices.

- Data should digitized whenever possible (physical objects and documents should be retained)
- All data should be archived whether they were published or supported by public or private funds, according to resources available.
- Plans for archiving should be incorporated into the research design at the inception of the project in accordance with available metadata standards.
- All data and annotated code required for replicating statistical analyses and underlying summary statistics of published research should be archived in a “trusted” data repository.
- All digital resources, such as digital imagery and custom software, used to generate quantitative data required for replicating statistical analyses should be archived with permanent links.
- Data archiving should prioritize data collected using public funds or public resources.

**ORCID**

Trudy R. Turner [https://orcid.org/0000-0001-8615-442X](https://orcid.org/0000-0001-8615-442X)
Connie J. Mulligan [https://orcid.org/0000-0002-4360-2402](https://orcid.org/0000-0002-4360-2402)

---

**How to cite this article:** Turner TR, Mulligan CJ. Data sharing in biological anthropology: Guiding principles and best practices. Am J Phys Anthropol. 2019;170:3–4. [https://doi.org/10.1002/ajpa.23909](https://doi.org/10.1002/ajpa.23909)