



The Data Sharing Champion

De-identified Human Subject Data

A Short Guide for Sharing De-identified Human Subject
Data

ASAP Grants Team

De-identified Human Subject Data

The purpose of this document is to provide an overview of considerations for publishing de-identified, individual human subject data in preparation for publication.

Human subject data may be classified as sensitive data, which are data that contain personal information, such as protected health information, and any other data that is likely to negatively harm an individual or community if publicly released.

The [ASAP Open Science Policy](#) requires that sensitive data must be deposited to the extent allowed by the associated research ethics approval. If the data can be openly shared, then we require these data be deposited in a publicly accessible repository. If the data cannot be openly shared, but can be shared with restricted or controlled access, then it must be shared as restricted or controlled data, and come with instructions for how to request access to the data.

Please note that we require that grantees collecting data from human participants to provide evidence that there has been satisfactory review and approval of the plan to collect and/or share such data from the appropriate ethics committee(s) (or evidence that no such approval is required).

If you have additional questions, please email the Open Science Team at openscience@parkinsonsroadmap.org.

Table of Contents

- [Planning & Approval](#)
- [Data Submission](#)

Planning & Approval

In developing a study utilizing human subject data, please address the following:

1. Review your data management and sharing plan, required by the Institutional Review Board (IRB) / Independent Ethics Committee (IEC) or local equivalent.
 - a. Plan should include details on data access and considerations for distribution or reuse, as well as long-term data preservation and maintenance.
 - b. Ensure fidelity to the informed consent that has been obtained from study participants to broadly share de-identified data to researchers outside your institution.
2. [Select a data repository](#), paying close attention to recommended data formats and access restrictions.
 - a. Some institutions may require an additional review by the Privacy Office (or local equivalent) to share human subject data in an open repository like [Zenodo](#) (i.e. a repository without controlled access).

Data Submission

To share de-identified human subject data, please follow these steps:

1. Prepare a de-identified version of your dataset through pseudonymization and/or other appropriate de-identification methods to the greatest extent that maintains sufficient scientific utility.
 - a. Consider the regional and/or national standards for identifiability governing privacy at your institution. US privacy laws recommend de-identification occur through either expert determination or safe harbor methods (i.e. removal of personal identifiers). [1]
2. Gather all necessary data documentation and metadata to enable discovery, reuse, and broadest possible reuse and citation, using open file formats whenever possible.
 - a. Related study protocols (e.g. main study protocol, sample collection protocols) must all be deposited and cited in recognized repositories in parallel (or before).
 - b. Metadata should incorporate the use of [common data elements \(CDE\)](#) where appropriate to support standardization and re-use.
3. Upload your de-identified data to the selected repository and obtain a globally unique and persistent identifier ([DOI](#), [RRID](#)).

1. HHS (U.S. Department of Health and Human Services), 2012. Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. <https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html>