

The National Institute on Aging (NIA) Division of Neuroscience (DN) provides the following sample Data Management and Sharing Plan for a hypothetical project proposing to conduct secondary analysis of data obtained previously from human subjects. Click [here](#) for more sample plans from NIA.

DATA MANAGEMENT AND SHARING PLAN

An example from an application focusing on secondary data analysis on data from human subjects.

Element 1: Data Type

A. Types and amount of scientific data expected to be generated in the project:

This is a secondary analysis of publicly available datasets comprising 1,000 subjects with Alzheimer's disease (AD) and 1,000 subjects without AD, aged 65-95 years. Primary data will be sourced from the Baltimore Longitudinal Study of Aging (BLSA) (clinical, cognitive data, MR imaging data) and from the OASIS-3 dataset (PET and MR imaging including multiple structural and functional sequences, along with related clinical, cognitive data, and APOE status), accessible through the Open Access Series of Imaging Studies (OASIS) database. This secondary analysis will yield novel insight into the functional properties of white matter (WM) circuits in normal aging and AD. This will involve examining local anisotropic correlation of the Blood oxygenation level-dependent (BOLD) signal on the voxel scale and templates for Diffusion Tensor Imaging (DTI) indices (fractional anisotropy, axial diffusivity, radial diffusivity, mean diffusivity) created using existing spatial normalization tools. For white matter/grey matter (WM/GM) connectivity, Fluorescence Optical Imaging (FOI) atlases in NIFTI (Neuroimaging Informatics Technology Initiative) format will be generated to delineate GM and WM connectivity, serving as normative data for future studies.

B. Scientific data that will be preserved and shared, and the rationale for doing so:

Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.

No new data will be collected for this study. All derived data, including atlases, codes, pipelines, and trajectories generated by the secondary analysis, will be shared through the respective repositories.

C. Metadata, other relevant data, and associated documentation:

Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate the interpretation of the scientific data.

We will provide all necessary documentation to reproduce our data analyses. The metadata will be shared with the derived data.

Element 2: Related Tools, Software and/or Code:

State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.

Pre-processing, post-processing MRI data analysis, and statistical analysis will be performed using code written in MATLAB, Python, or R. MATLAB is commercial software with licenses available for purchase, while Python and R are open-source software. DTI indices and FOI atlases will be generated using freely available tools such as SPM, AFNI, and FSL.

Element 3: Standards:

State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist.

Images from the OASIS and BLSA databases will be downloaded and inspected for acceptable quality (little motion, spikes, artifacts). The secondary data analysis will use the Brain Imaging Data Structure (BIDS) standard that provides methods to create BIDS-compliant datasets and to validate converted data structures.

Element 4: Data Preservation, Access, and Associated Timelines

A. Repository where scientific data and metadata will be archived:

Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived; see [Selecting a Data Repository](#)).

All secondary derived data, metadata, atlases, and pipelines will be returned to the BLSA and OASIS data systems for sharing through their established channels. All codes for the informatics framework and all analysis modules with image process routines and atlases will be stored in the NIH-supported Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC) infrastructure and GitHub.

B. How scientific data will be findable and identifiable:

Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.

The data generated in this study will be assigned a Digital Object Identifier (DOI) by BLSA and OASIS, where data will be shared, and referenced in the publication to provide the research community with easy access to the exact data generated in the study.

C. When and how long the scientific data will be made available:

Describe when the scientific data will be made available to other users (i.e., no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

All data and codes from this study will be available at the time of publication and no later than the end of the award. They will be accessible for as long as BLSA and OASIS make the data available.

Element 5: Access, Distribution, or Reuse Considerations

A. Factors affecting subsequent access, distribution, or reuse of scientific data:

NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing. See [Frequently Asked Questions](#) for examples of justifiable reasons for limiting sharing of data.

The primary datasets analyzed in this project are publicly available under controlled access and shared under the guidelines of their respective repositories (OASIS and BLSA). Individuals interested in accessing these datasets should contact the original repository. New variables generated from this secondary analysis of the existing dataset will be returned to the BLSA and OASIS data system for sharing through their established channels. We will adhere to the respective Data Use Agreements for OASIS and BLSA, which prohibit any attempts to re-identify research participants or redistribute primary and secondary derived data. We will provide proper attribution to these organizations in any publication or presentation of results derived from these datasets and submit the results for review prior to publication.

B. Whether access to scientific data will be controlled:

State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).

The existing primary datasets are publicly available and shared under controlled access according to their respective repositories' guidelines. We will not redistribute the primary data or attempt to identify research participants. All derived secondary data, including atlases and aging trajectory, will be stored in the primary data repositories (OASIS and BLSA) with controlled access. Codes and analysis modules for the secondary analyses will be uploaded to GitHub and NITRC where they will be publicly accessible.

C. Protections for privacy, rights, and confidentiality of human research participants:

If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

This study involves analyzing existing, de-identified data shared under the guidelines and restrictions of OASIS and BLSA repositories. No additional human subject deidentification or consent will be required.

Element 6: Oversight of Data Management and Sharing:

Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

The Principal Investigator (PI) will be responsible for the oversight of the DMS plan. The PI will have quarterly meetings with the research team of the project. The PI's Office of Sponsored Programs at their

institution will ensure that all data are shared and managed in accordance with NIH data-sharing policies as a part of the institution's compliance during the submission of the annual NIH Research Performance Project Report (RPPR).