EXAMPLE DATA MANAGEMENT AND SHARING PLAN (in compliance with SF-424 Forms H)

EXAMPLE FOR SECONDARY PHENOTYPIC AND CLINICAL DATA

ELEMENT 1: DATA TYPE

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

The data to be shared will include MRI images and clinical assessments from human research participants. This application is focused on secondary data analysis from existing data but will also deposit privately held data to a public repository. The existing data is available from the NIMH Data Archive (NDA) in collections 2134 (148 subjects) and 2433 (47 subjects). In addition, we have data from a previous study involving 155 research participants with major depressive disorder that have not yet been shared with the research community but will be uploaded to NDA during the second quarter of the first year of funding. As discussed in the application, structural MRI scans are available for time points before and after treatment along with relevant clinical data.

B. Scientific data that will be preserved and shared, and the rationale for doing so:
This is a secondary data analysis application, so new data is not being measured. Much of the data is already available through NDA. Clinical and imaging data from 155 new subjects will be shared.

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

Preparation for submitting existing data to NDA is largely complete. Within the first six months following the award, we will submit the Data Submission Agreement to NDA and will create the Data Expected list (see Standards section) in our new NDA Collection. The policies of our institution mandate that exact dates will not be shared (see Access section).

ELEMENT 2: RELATED TOOLS, SOFTWARE, AND/OR CODE

The basic statistical analyses described in the application will be done using R. We plan to use the MRI data analysis tools in the FMRIB Software Library (FSL) for multi-level modeling of group effects. BrainVoyager software will be used for anatomical segmentation to isolate regions of interest within individual subjects, and the AI-powered analyses described in the application will use custom code written with the PyTorch library for Python. R, FSL, Python, and PyTorch are all freely available to the research community. BrainVoyager is commercial software, with licenses available for purchase.

All R and Python code (including trained model weights) will be available on our lab Bitbucket page no later than when publications are submitted. The Bitbucket page is publicly assessable and will be hosted for at least 5 years after the grant award ends.

ELEMENT 3: STANDARDS

The data that will be used for some of the proposed secondary data analysis is already in NDA and is formatted using NDA data dictionaries. The new data we will deposit will also use existing NDA data dictionaries. Since the data set to be deposited into NDA was collected prior to the publication of NOT-MH-20-067, not all of the common data elements expected by NIMH are available. However, we will transform some existing demographic and clinical data into the formats expected for:

A. Age (ndar_subject01)
B. Sex at Birth (ndar_subject01)
C. Patient Health Questionnaire-9 (PHQ-9, cde_phq901 NDA data dictionary).

In addition, information from the Beck Depression Inventory will be deposited for all 155 research participants using the NDA bdi01 data dictionary. Deposited images will use the NDA image03 data

Commented [JB1]: *These example DMS Plans are provided for educational purposes to assist applicants with developing Plans but are not intended to be used as templates and **their use does not guarantee approval by NIH. Do not copy/paste this Plan without modifying it to reflect the types of data that are expected to be generated through your project.***

Note that the example DMS Plans may reflect additional expectations established by NIH or specific NIH Institutes, Centers, or Offices that go beyond the DMS Policy. Applicants will need to ensure that their Plan reflects any additional, applicable expectations (including from NIH policies, ICO policies, or as stated in the FOA).

In addition, these examples may reflect resources or policies that are in place at other institutions but that are not necessarily available at the University of Iowa. If needed, investigators can contact Research Data Services (lib-data@uiowa.edu) if they have questions regarding how to best complete their DMS Plan.

Commented [BJY2]: *Example from NIMH: https://www.nimh.nih.gov/sites/default/files/documents/funding/managing-your-grant/resource-sharing-docs/Secondary%20Data%20Analysis_Template_v2.docx*

dictionary. Data derived from the MRI images will be deposited into NDA using the imagingcollection01 data dictionary.

ELEMENT 4: DATA PRESERVATION, ACCESS, DISTRIBUTION AND ASSOCIATED TIMELINES

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

All previously unshared data will be deposited to NDA no later than 12 months after the award begins.

B. How scientific data will be findable and identifiable:

Data will be findable for the research community through the NDA collection that will be established when this application is funded. For all publications, an NDA study will be created, and the data relevant to that publication will be shared immediately. Each of those studies is assigned a digital object identifier (DOI). This data DOI will be referenced in the publication to allow the research community easy access to the exact data used in the publication.

C. When and how long the scientific data will be made available:
The research community will have access to the previously unshared data at the end of the grant award. Researchers will request data using the standard processes at NDA, and the NDA data access committee will decide which requests to grant. The standard NDA data access process allows access for one year and is renewable. Once the data are submitted to NDA, that archive will control the long-term persistence of the data set.

ELEMENT 5: ACCESS, DISTRIBUTION, OR REUSE CONSIDERATIONS

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

The two existing data sets from NDA used consents that allow broad data sharing. The new dataset to be uploaded to NDA also was collected using informed consent terms that allow broad data sharing. Access to data housed by the NDA requires the completion of a Data Use Certification, which prohibits any redistribution or attempts to re-identify research participants.

B. Whether access to scientific data will be controlled
To request access of the data, researchers will use the standard processes at NDA, and the NDA Data Access Committee will decide which requests to grant. The standard NDA data access process allows access for one year and is renewable. Once the data are submitted to NDA, that archive will control the long-term persistence of the data set. Currently, NDA has no process for deleting or retiring data sets.

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

The NDA GUID tool allows researchers to aggregate data from the same research participant without different laboratories having to share personally identifiable information about that research participant. The NDA data dictionaries do not permit personally identifiable information to be shared. NDA maintains a Certificate of Confidentiality.

For the 155 participants from our previous study, exact dates have been obscured via the Shift and Truncate method [1], which preserves within-case temporal relations.

ELEMENT 6: OVERSIGHT OF DATA MANAGEMENT AND SHARING

The following individuals will be responsible for data collection, management, storage, retention, and dissemination of project data, including updating and revising the Data Management and Sharing Plan as necessary each year at the time of the Research Performance Progress Report.

[Name, position title, host institution, ORCID, email]

Validation Schedule (this section is required by NIMH)

Since this is a secondary data analysis application, validation of newly collected data will not occur. The new data to be deposited to NDA will go through their validation tool when the data are initially uploaded.