Data Management and Sharing Plan

Element 1: Data Type:

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

Summarize the types and estimated amount of scientific data expected to be generated in the project.

The proposed study will generate demographic, clinical, MRI, 1H fMRS, and fMRI imaging data from 110 affected youth and 110 matched healthy controls. The estimated amount of data includes:
-Demographic data: age, sex, ethnicity, etc., for each participant.
-Clinical data: symptom severity scores, diagnostic assessments, and other relevant clinical information.
-MRI data: structural and functional images.
-1H fMRS data: spectroscopy data assessing metabolic activity in the brain.
-fMRI data: task-based and resting-state functional connectivity data.

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.

All collected demographic, clinical, MRI, 1H fMRS, and fMRI imaging data will be preserved and shared. The rationale is to contribute meaningfully to the understanding of mental health disorders in youth, facilitate secondary analyses, and enhance research collaboration and replication studies.

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 interpretation of the scientific data.

The following metadata and documentation will be made accessible:

- Study protocols.
- Data collection instruments (e.g., questionnaires, surveys).
- MRI, 1H fMRS, and fMRI imaging parameters.
- De-identification procedures to ensure participant privacy.

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.

Specialized tools and software needed for data analysis include:
- MRI analysis software (e.g., FSL, AFNI).
- Spectroscopy analysis software (e.g., LCModel, jMRUI).
- fMRI analysis software (e.g., SPM, FreeSurfer).

These tools can be accessed through institutional licenses or open-source repositories.

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.

The project will apply common data standards to enable interoperability, including:

- BIDS (Brain Imaging Data Structure) for MRI and fMRI data.

- ISMRM guidelines for 1H fMRS data acquisition and analysis.

- Standardized clinical assessment tools with established norms.

These standards ensure that the scientific data generated are compatible with existing datasets and resources in the field.

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.

Scientific data and metadata arising from this project will be archived in the NIMH Data Archive (NDA).

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.

Data will be made findable and identifiable through a persistent unique identifier provided by the NDA, allowing for easy access and citation.

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 the time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

The scientific data will be made available to other users no later than the time of an associated publication or end of the performance period, whichever comes first, and will remain available indefinitely through the NDA.

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.

Given that all participants consent for broad data sharing, there are no anticipated limitations on access, distribution, or reuse beyond standard protections for privacy and confidentiality.

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).

Access to the scientific data will not be controlled; once archived in the NDA, the data will be openly available to qualified researchers upon request.

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).

To protect participant privacy, all data will be de-identified prior to archiving. Participants are consented with an understanding that their de-identified data may be shared broadly with other researchers.

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).

Compliance with this Data Management and Sharing Plan will be monitored and managed by the Principal Investigator (PI) in collaboration with the institutional data management office. The PI, along with a designated data manager, will oversee data collection, storage, sharing, and compliance with NIH policies on an ongoing basis throughout the project period.