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.

This project will generate the following types and estimated amounts of scientific data:
- Compound Protocols: Documentation of design, production, characterization, and purification protocols for 30 analogs of Compound A.
- In Vitro and Cell-Based Assay Data: Quantitative and qualitative results from testing all 30 Compound A analogs in various vitro and cell-based assays (e.g., cell viability, target binding, potency).
- Substance Stability and Toxicology Data: Data on stability and in vitro/in vivo toxicology for 10 selected compounds, including dose-response data from 6 wildtype (WT) mice per dose across 5 dosing levels.
- Pharmacokinetics (ADME) Data: Absorption, distribution, metabolism, and excretion data for 3-4 prioritized compounds, using 10 5xFAD mice per dose across 4 doses.
- Preclinical PET Imaging Data: PET imaging datasets (raw and analyzed) for 2-3 compounds, from up to 15 WT or 5xFAD mice per compound.
The total estimated data volume is approximately 20-30 GB, primarily consisting of assay result spreadsheets (CSV, XLSX), imaging files (DICOM, TIFF), protocol documents (PDF, DOCX), and associated metadata files.

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 key scientific data supporting published findings will be preserved and shared, including:
 - Protocols for compound design, production, and characterization.
 - Raw and processed assay data for all tested compounds.
 - Toxicology and stability data for 10 compounds.
 - Pharmacokinetics (ADME) datasets for 3-4 compounds.
 - Preclinical PET imaging data (raw DICOM files and processed images, along with analysis results).
These data are preserved and shared to ensure reproducibility of findings, enable secondary analyses by the community, and accelerate therapeutic development in aging-related disorders.

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:

- Detailed protocols for compound synthesis, purification, and characterization.

- Assay protocols and standard operating procedures (SOPs).

 - Experimental design metadata (e.g., mouse strain, age, sex, dosing regimens, randomization).

- Data dictionaries and codebooks for datasets.

 - Imaging acquisition and analysis parameters.

- Quality control and validation reports.

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.

Most data (spreadsheets, PDF protocols) can be accessed with standard office software (e.g., Microsoft Excel, Adobe Reader). PET imaging data will be provided in DICOM format, compatible with open-source viewers such as MicroView, OsiriX, or ImageJ/Fiji (all freely available). Example analysis scripts (e.g., R or Python code) used for data processing will be shared as supplementary files and hosted on GitHub (public repository). All tools and software are either open-source or widely available at no cost.

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 employ the following data standards:

- Data Formats: CSV for tabular data, DICOM for imaging, PDF/DOCX for protocols.

- Metadata Standards: Minimum Information About a Microarray Experiment (MIAME) for in vitro assays where applicable, and Digital Imaging and Communications in Medicine (DICOM) standards for imaging data.

- Identifiers: Unique compound and sample identifiers will be assigned and referenced across all datasets.

- Ontology/Controlled Vocabularies: Common ontologies such as MeSH and ChEBI will be used for compound and assay annotation.

Where applicable, standards recommended by repositories (e.g., Figshare, OpenNeuro) and the NIH will be followed to maximize interoperability.

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 will be archived in [Figshare](https://figshare.com/) (generalist NIH-recommended repository) for protocols and tabular data, and [OpenNeuro](https://openneuro.org/) for PET imaging data. Supplementary scripts will be deposited in a public [GitHub](https://github.com/) repository.

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.

All datasets will be assigned persistent Digital Object Identifiers (DOIs) by Figshare and OpenNeuro, ensuring long-term findability and citability. Metadata will be indexed and searchable via repository search tools and major search engines.

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.

Data will be made available no later than the first publication of results or the end of the funding period, whichever occurs first. Data will remain accessible in the repositories for at least 10 years after publication, in accordance with repository policies.

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.

All data derive from non-human (murine) studies and in vitro assays; no human subject data are involved. There are no anticipated restrictions on access, distribution, or reuse. Data will be shared under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, permitting broad reuse with attribution.

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 be open and unrestricted. No controlled access is required.

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).

Not applicable; no human participants or human-derived data are involved in this project.

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 overseen by the Principal Investigator (PI), in coordination with the project's Data Steward and Institutional Research Data Management Office. Data management practices, archiving, and sharing will be reviewed quarterly during internal project meetings. The Data Steward will be responsible for ensuring timely deposition, completeness of metadata, and adherence to standards. Annual compliance reports will be submitted to the PI and institutional leadership.