The National Institute on Aging (NIA) Division of Social and Behavioral Research (BSR) provides the following sample Data Management and Sharing Plan for a project involving the collection of new data for an existing longitudinal study with plans to be shared under a restricted use agreement. Click here for more sample plans from NIA.

DATA MANAGEMENT AND SHARING PLAN

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommended not to exceed two pages. Text in italics are instructions and should be deleted before submitting the plan. There is no "form page" for the Data Management and Sharing Plan. The DMS Plan may be provided in the *format* shown below.

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 project involves administering an additional wave of data collection for a longitudinal study. Data to be generated include social, behavioral, and biological data collected through web-surveys, in-person interviews, and home health exams. Approximately 13,000 web-surveys and interviews and 200 home health exams will be completed.

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.

From the proposed additional wave of data collection, we will share social, behavioral, and biological data via a tiered access plan determined by sensitivity of data as described in Element 5.

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.

Metadata will include codebooks and user guides that provide variable descriptions, survey instruments and descriptions of collection processes, and guidance for use.

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.

Data will be provided in several user-friendly formats including ASCII, SAS, SPSS, R, and Stata. We provide examples that employ basic analysis codes using two survey software packages, Stata and SAS.

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.

Codebooks will be provided in Data Documentation Initiative (DDI) format. DDI is an internationally recognized metadata standard for documenting survey data used by hundreds of research organizations.

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).

Public-access data and metadata will be accessible via [Repository Name Here]. Restricted use and high-security restricted-use data will be provided by encrypted download through the repository virtual data portal. Virtual enclave-restricted data are available only through the data portal Virtual Data Enclave.

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 assigned a Digital Object Identifier (DOI) and will be searchable via [Repository Name Here] and other 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 time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

All data will be released by the time of publication or the end of the project period, whichever comes first, and will be available indefinitely through the repository websites.

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 examples of justifiable reasons for limiting sharing of data.

Due to the nature of the data that will be collected, and to ensure participant confidentiality, data will be made available under a tiered system based on data sensitivity. All tiers of data require a data-use agreement at a minimum, additional precautions for restricted-use, high restricted-use, and virtual enclave data will be included in the data request forms.

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).

Public-access data will be available to researchers through the [Repository Name Here] website. Researchers who wish to use the most sensitive data and/or link project data to other datasets must apply for access through a contract request to use the [Repository Name Here] virtual data enclave.

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).

Participants will be provided a broad consent document [if allowed under IRB review] to facilitate data sharing. Data will be de-identified, and the tiered-access plan will provide additional protections for the data based upon risk assessment. The institution will determine whether the de-identified individual-level data are subject to the NIH Certificate of Confidentiality policy, and if so, will protect accordingly and ensure that recipients of the data are aware of the protection.

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 PI will oversee archiving and sharing of the analysis codes generated from this project upon publication or at the end of the project period, whichever comes first. The final project report will summarize adherence to this data management and sharing plan.