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 collection of social and behavioral data. 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.

Data to be generated from this project include physical activity data from an accelerometer, responses to a survey characterizing caregiver well-being, and participant responses to semi-structured interviews captured via digital recordings. We will collect data from 250 participants across three timepoints.

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.

De-identified survey responses and raw accelerometer data will be shared. Individual-level digital recordings will not be shared publicly because of the difficulty masking the identity of participants. Digital recordings may be shared under a data-use agreement. Similarly, qualitative data from participant interviews will be de-identified prior to sharing under a data use agreement.

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 readme files, codebooks, survey instruments, analysis codes, and other supporting documentation.

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 shared in ascii files. No specialized software or resources are needed to access data files. Analysis codes will be created in R, which is an opensource software.

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.

Metadata will conform to DataCite, DDI, or other established metadata standards. Common data elements as defined by National Library of Medicine (NLM) will be used whenever practical.

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).

Data and metadata will be made available through [Repository Name Here].

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 and metadata will be assigned a unique Digital Object Identifier (DOI) and a formal study citation by the repository.

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 de-identified data will be available upon publication of related work or end of the project period, whichever comes first, and will remain available indefinitely.

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.

Data will be shared under a two-tier distribution system. All data that can be de-identified will be publicly shared. Data with specific confidentially risks, such as individual-level digital recordings will not be shared publicly because of the difficulty masking the identity of participants. Data with increased risk of reidentification will be shared under a data-use agreement that clearly establish use rules to facilitate respondent 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).

De-identified data will be freely available on the repository website. Restricted data (I.e., digital recordings which cannot be de-identified) will be available through a data use agreement. Data use agreements providing access to sensitive data will be managed by the repository of record.

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).

All direct participant identifiers will be removed from data prior to sharing. Restricted data (I.e., digital recordings which cannot be de-identified) will be available through a data use agreement. Study participants will be asked to consent to widespread data sharing with the research community based upon recommendations of the local IRB and approval of the consent document. 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. Communication with the study program officer will be made as needed to ensure compliance.