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 secondary data analysis using data obtained from an existing repository. In this case, the Health and Retirement Study (HRS) is used as the example data of origin, but any publicly available data could be substituted. 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 work will not collect new data. This project will perform secondary analysis using publicly available data (accessible via the Health and Retirement Study) and data that are accessible through a data use agreement (Health and Retirement Study-linked Medicare 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.

The proposed work is a secondary data analysis of available data. No new data will be collected from this project.

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.

Analysis codes and user guides developed as part of the analysis will be shared. Data dictionaries and documentation will be provided for any new scales or variables generated from the secondary data as part of the metadata elements.

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.

Statistical analysis code will be generated using R and shared as part of the metadata elements.

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.

Data, documentation, and metadata for the primary data is available through the repository of origin. Analysis code, measures and scales created as part of the project using the data of origin will be shared as metadata elements.

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).

Analysis codes and guidebooks generated by the analysis will be deposited in [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.

Codes deposited on [Code Repository Name Here] will be indexed and searchable, preferably with a unique data object identifier (DOI).

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.

Analysis codes will be made available at either the end of the project period or at the time of publication and will be available on [Code Repository Name Here] 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 used for secondary analysis in this project are 1) publicly accessible (e.g., Health and Retirement Study), or 2) shared through a restricted use agreement (e.g., Health and Retirement Study-linked Medicare Data).

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 public-use data employed in this application is available through the repository of origin under their use guidelines.

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).

No new data will be collected from this project, only analysis of existing data. Data deidentification procedures applied to the data of origin ensures that no contact with original respondents is possible.

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.