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 should be deleted 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:

The data to be shared will include software code from Stata SE version 17.0 produced in the analysis of the Health and Retirement Study data. This study is use on secondary data analysis from the Health and Retirement Survey which are publicly available through the University of Michigan Health and Retirement Study Data Portal https://hrsdata.isr.umich.edu/data-products/public-survey-data. No new data will be produced from this study. Data will be archived in OpenICPSR.

B. Scientific data that will be preserved and shared, and the rationale for doing so:

This is a secondary data analysis application, so new data is not being measured. All of the data is already available through the University of Michigan Health and Retirement Study Data Portal https://hrsdata.isr.umich.edu/data-products/public-survey-data. Software codes from Stata SE version 17.0 will be shared.

C. Metadata, other relevant data, and associated documentation:

Preparation for submitting existing data to NDA is largely complete. Within the first six months following the award, we will submit the Data Submission Agreement to NDA and will create the Data Expected list (see Standards section) in our new NDA Collection. The policies of our institution mandate that exact dates will not be shared (see Access section). This will include detailed records of how researchers can download data files from the University of Michigan Health and Retirement Data Portal. We will also include detailed documentation of our code made available through Stata SE version 17.0 do-files.

Element 2: Related Tools, Software and/or Code:

The basic statistical analyses described in the application will be done using Stata SE version 17.0. All Stata code saved in do-files will be available be uploaded to the openICPSR data repository and will remain assessable and will be hosted for a minimum of 7 years after the grant award ends.

Element 3: Standards:

This study uses publicly available secondary data analysis from the Health and Retirement Survey which are can be downloaded through the University of Michigan Health and Retirement Study Data Portal https://hrsdata.isr.umich.edu/data-products/public-survey-data. No new data will be produced from this study. Code in Stata SE version 17.0 which is used to transform variables and merge biennial years of the Health and Retirement Study will be uploaded to the openICPSR data repository.

Element 4: Data Preservation, Access, and Associated Timelines

A. Repository where scientific data and metadata will be archived:

The data and metadata from this project will be archived at the openICPSR. openICPSR is a CoreTrustSeal certified repository providing long-term access to and preservation of data packages curated by domain specialists.

B. How scientific data will be findable and identifiable:

Every ICPSR data collection receives a globally unique and persistent identifier, which are registered with DataCite (a global DOI provider) and included in the citation and metadata record of each ICPSR data collection. ICPSR creates rich study- and variable-level metadata records in the Data Documentation Initiative (DDI) disciplinary metadata format using information supplied by data depositors and other sources. Metadata are available for bulk export in a variety of metadata formats (Dublin Core, DDI, and MARCXML), as well as exportable from and embedded in dataset landing pages, including structured Schema.org data markup

indexed by leading search engines. Metadata are organized using standardized, well-established formats, templates, and vocabularies, and are released with a clear and accessible data usage license.

C. When and how long the scientific data will be made available

Data will be made available at the time of associated publication or end of project period, whichever is earlier. We will meet the data submission and release timeframes specified by the funding agency and policies, as described on NIH's data sharing webpage. Study data deposited in openICPSR will be available to the research community. ICPSR permanently archives deposited files, supporting the data through changing technologies, new media, and data formats.

Element 5: Access, Distribution, or Reuse Considerations

A. Factors affecting subsequent access, distribution, or reuse of scientific data:

Access to the Health and Retirement Study data requires the creation of an account and agreeing to conditions of use for Health and Retirement Study public release data through the University of Michigan Health and Retirement Study Data Portal. Access to data housed by the NDA requires the completion of a Data Use Certification (see the Get Data section of the NDA web site), which prohibits any redistribution or attempts to re-identify research participants.

B. Whether access to scientific data will be controlled:

To request access of the data, researchers will use the standard processes at NDA, and the NDA Data Access Committee will decide which requests to grant. The standard NDA data access process allows access for one year and is renewable. Once the data are submitted to NDA, that archive will control the long-term persistence of the data set. Currently, NDA has no process for deleting or retiring data sets.

C. Protections for privacy, rights, and confidentiality of human research participants:

All data in the Health and Retirement Study are deidentified with no available information on information including name, date of birth, social security numbers, or addresses. Accordingly, through the process of deidentification all participants in the Health and Retirement Study have protections of privacy and confidentiality.

Element 6: Oversight of Data Management and Sharing:

The Office of the Executive Vice President & Chief Academic Officer (EVP/CAO) and The Office of Data Science (ODS) at UTHealth Houston will provide joint institutional oversight for the DMS plan. Dataset(s) resulting from this research will be cataloged with essential project metadata and persistent identifiers in the institutional Data Ecosystem Portal for UTHealth Houston (DEPUT). DEPUT is the institutional research data portal supported by UTHealth Houston for DMS validation, data curation, and streamlined data transfer to appropriate data archives for long-term storage and accessibility. Project PI will update data status in DEPUT, and the institutional office of Sponsored Projects Administration (SPA) will perform annual validation according to the DMS plan. Validation results will be reported to EVP/CAO and ODS for review. Gaps, if any, will be identified with appropriate correcting measures implemented.