HRS Overview for Supporting Dementia and HCBS Analyses

Community Care Network for Dementia (CaN-D) Data Hub Resource

Prepared by Jessica Hayes and Jordan Harrison, RAND

Updated October 15, 2024

Background

The University of Michigan Health and Retirement Study (HRS) is a longitudinal panel study that surveys a representative sample of approximately 20,000 people aged 51 and over in America, conducted by the Institute for Social Research at the University of Michigan with funding from the National Institute on Aging and the Social Security Administration. The HRS's multidisciplinary data provide researchers the opportunity to investigate many different aspects related to population aging in the United States, including data on cognition, disability, dementia, and caregiving.

Data Access and Version

Public use files can be obtained from https://hrsdata.isr.umich.edu/data-products/public-survey-data

Public files used to create the provided code and dataset with relevant dementia and caregiving variables from HRS are listed as follows:

- **RAND HRS Longitudinal File 2020 (V2)** (https://hrsdata.isr.umich.edu/data-products/rand-hrs-longitudinal-file-2020)
- **RAND HRS Family Data 2018 (V2)** (https://hrsdata.isr.umich.edu/data-products/rand-hrs-family-data-2018)
- **RAND HRS Detailed Imputations 2020 (V2)** (https://hrsdata.isr.umich.edu/data-products/rand-hrs-detailed-imputations-file-2020)
- **Langa-Weir Classification of Cognitive Function (1995-2020)** (https://hrsdata.isr.umich.edu/data-products/langa-weir-classification-cognitive-function-1995-2020)

An account is needed to obtain the files. Sensitive and restricted files require additional application processes to obtain.

Documentation for each data file is provided on the download page for the relevant file.

HRS Documentation is also available here: https://hrs.isr.umich.edu/documentation

Note that as of May 2025, updated versions of the RAND HRS Longitudinal File and RAND HRS Detailed Imputations File are available on HRS's website. Users are encouraged to use the most updated versions when possible. The provided code and datasets here use the older versions listed above as that is what is currently available on the ENCLAVE.

HRS File Organization

The Core HRS survey has been fielded in 1992, 1993, 1994, 1995, and biennially 1996–2020. (2022 early release is not included in the program and output file mentioned below and 2024 is currently being fielded). Replenishment of the sample occurred in 1993 and then every six years starting in 1998.

Each round of the Core HRS survey includes multiple raw data files—one for each survey section along with additional preload and cover screen files.

HRS also provides several longitudinal files: (1) the tracker file, which includes the most up-to-date demographic, weight, and identifier variables; (2) various files that provide additional details on geographic variables, cognition, social security wealth, etc.

Additionally, HRS fields and provides exit interview data and off-year studies. They also provide researcher-contributed projects, including the Langa-Weir Classification of Cognitive Function.

The RAND HRS Data files create user-friendly files derived from all waves of the HRS. The RAND HRS Longitudinal File contains cleaned and processed variables with consistent and intuitive naming conventions, model-based imputations, and spousal counterparts of most individual-level variables. This file, in addition to the RAND HRS Detailed Imputations File and the RAND HRS Family Data File (respondent-level) are longitudinal files that use the HRS data as inputs to provide datasets that can be used with little to no additional processing by users. The RAND HRS data files are also publicly available on HRS's website.

Variables

Cleaned, derived, and processed variables taken from the public files listed above are provided in ***HRS_variables_description.xlsx***. The file contains the variable names and short descriptions of the variables along with information on which data files they were taken from. Please refer to the documentation for the relevant data file for more detailed information.

Dementia Classification

HRS collects the following information on dementia and cognition:

- A report by the respondent or a proxy respondent that a doctor told the respondent they had dementia or Alzheimer's disease. (Note: These are not directly used in the cognitive function classification described below)
- Proxy respondents are asked assessment of respondent memory and report on IADL limitations. Additionally, the interviewer is asked to assess the respondent's cognitive impairment (2000 forward). These form a 9-point scale in 1995-1998 and an 11-point scale 2000 forward **(completed by proxy respondent only)**
- Cognitive tests to evaluate the sample person's memory (immediate and delayed 10-word recall), working memory (serial sevens subtraction test), and mental processing (counting backwards). These form the 27-point cognition scale for all non-proxy respondents. Respondents 65 and older are additionally asked orientation questions (date, month, year, and day of the week; object naming; naming the President and Vice President). These are added to create a 35-point cognition scale.

The HRS core data does not directly include an indicator for dementia but the Langa-Weir Classification of Cognitive Function (a researcher-contributed public use file) uses the measures listed above, conducts additional imputation for missing values, and classifies respondents as having normal cognition, cognitive impairment but no dementia (CIND), or having dementia (i.e., ***cogfunction***):

- Normal cognition
    - A value of 12-27 on the 27-point cognition scale for self-respondents or a value of 0-2 on the 9-point or 11-point cognition scale for proxy respondents.
- Cognitive Impairment, no dementia (CIND)
    - A value of 7-11 on the 27-point cognition scale for self-respondents or a value of 3-4 on the 9-point or 3-5 on the 11-point cognition scale for proxy respondents.
- (Probable) Dementia
    - A value of 0-6 on the 27-point cognition scale for self-respondents or a value of 5-9 on the 9-point or 6-11on the 11-point cognition scale for proxy respondents.

See relevant documentation from https://hrsdata.isr.umich.edu/ for more details:

- Cross-Wave Imputation of Cognitive Functioning Measures 1992-2020
    - https://hrsdata.isr.umich.edu/data-products/cross-wave-imputation-cognitive-functioning-measures-1992-2020
    - https://hrsdata.isr.umich.edu/sites/default/files/documentation/data-descriptions/1686605147/COGIMP9220_dd.pdf
- Langa-Weir Classification of Cognitive Function (1995-2020)
    - https://hrsdata.isr.umich.edu/data-products/langa-weir-classification-cognitive-function-1995-2020
    - https://hrsdata.isr.umich.edu/sites/default/files/documentation/data-descriptions/1695907706/Data_Description_Langa_Weir_Classifications2020_V2.pdf

Dementia Harmonization with Other Cohort Studies

This documentation report describes the 2016 Harmonized Cognitive Assessment Protocol (HCAP). The Harmonized Cognitive Assessment Protocol is a substudy within the ongoing Health and Retirement Study (HRS). It is part of the HCAP Network, an ongoing international research collaboration funded by the National Institute on Aging (NIA) that seeks to measure and understand dementia risk by collecting a carefully selected set of established cognitive and neuropsychological assessments and informant reports to better characterize cognitive function among older people.

Description source: https://hrsdata.isr.umich.edu/data-products/2016-hcap-respondent-informant-data

Documentation Report Citation:
Weir, D. R., Langa, K., M & Ryan, L. H. (2016). 2016 Harmonized Cognitive Assessment Protocol (HCAP) Study Protocol Summary. University of Michigan.
https://hrs.isr.umich.edu/publications/biblio/9950

HCBS Variables

HRS includes information on difficulty and help with activities of daily living (ADLs: walking across a room, dressing, bathing, eating, getting in and out of bed, toileting) and instrumental activities of daily living (IADLs: using a phone, managing money, taking medications, grocery shopping, preparing hot meals). For each task requiring assistance, a helper is identified, along with the following information about the helper: relationship to respondent, amount of help provided, whether care was paid/unpaid, and payment amount (excluding expenses paid by Medicaid or insurance). The helper information (number of helpers, amount of help, payment to helpers) is summarized and provided by helper type (all, paid/unpaid, family/non-family).

Additionally, information on difficulty with other functional activities (walking, jogging, sitting, getting up from a chair, climbing stairs, lifting, stooping, pushing, picking up a dime).

Stata Program and output

HRS_variables.do pulls relevant demographic, dementia, and caregiving variables from HRS, RAND HRS, and HRS-researcher-contributed data sets and combines them into one file and transforms from wide to long format.

Sources and download links for the input data files are listed at the top of the HRS_variables.do file (and in the **Data Access and Version** section of this document).

If rerunning the program, intermediate data files will be created and stored under the "data" folder global assignment that the user can edit.

The final data files, created by running HRS_variables.do, are:

- **hrs_cand_wide.dta**, which provides the relevant variables in wide format.
    - There will be one observation per HRS respondent.
- **hrs_cand_long.dta**, which provides the relevant variables in long format.
    - There will be one observation for each wave/year times the number of HRS respondents who responded that wave.