

HEALTH AND RETIREMENT STUDY  
2022 Tracker  
Early, Version 1.0  
(Partial Release: 2022 Panel Cases Only)

August 2024

Data Description and Usage

## Conditions of Use

By registering for access to HRS Public Release data, the User agrees to all of the following:

- Make no attempts to identify study participants.
- Not to transfer HRS Public Release data to any third party other than staff or students for whom you are directly responsible except as indicated below.
- Not to allow others to use your username and password to access this site.
- To certify the destruction of any downloaded Public Release data file as well as any data files derived from the downloaded file when requested to do so by the Health and Retirement Study.
- To include the following citation in any research reports, papers, or publications based on Public Release data:

*In text:*

"The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan."

*In references:*

"Health and Retirement Study, ([insert Product Name]) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, (year)."

- To include the following citation in any research reports, papers, or publications based on any Public Release data file tagged as "Early" or "Preliminary":

"This analysis uses data from the Health and Retirement Study, ([insert Product Name]), sponsored by the National Institute on Aging (grant number NIA U01AG009740) and conducted by the University of Michigan."

- Please note that although it is necessary to include the above citations to the HRS in any publications, it is not necessary to associate your publications with the HRS grant in [PubMed](#).
- Provide information regarding any research product (dissertation, thesis, journal article, book, book chapter, report, etc.) based on data obtained from the Health and Retirement Study by sending an electronic copy to [hrspublications@umich.edu](mailto:hrspublications@umich.edu).
- Report immediately to the Health and Retirement Study at [hqsquestions@umich.edu](mailto:hqsquestions@umich.edu) any disclosure of study participant identity as well as any discovery of flaws or errors in the data or documentation files.

Notify the Health and Retirement Study through use of the update function provided at this site or by electronic mail directed to [hqsquestions@umich.edu](mailto:hqsquestions@umich.edu) of changes in your electronic mail address, postal address, telephone number, organizational affiliation or organizational status.

## 1. Introduction

The HRS Tracker file is created to facilitate the use of HRS data within and across waves. The file contains one record for every person who was ever eligible to be interviewed in any wave. Each record contains basic demographic information, interview status, and if, when, and how an interview was conducted in each wave. Also included are cross-sectional weights and information on inter-respondent relationships, which are vital to almost all substantive analyses of the HRS data.

The 2022 HRS Tracker (Early, Version 1.0) is an "early partial" public release. It includes data for cases previously released and data for 2022 panel cases (only) as a result of the 2022 data collection effort. Data for the newly screened Early Generation X cohort will be available in a subsequent release.

This document, complementary to the Tracker file codebook, consists of eight sections. Section 1 introduces the document, describes the panel cohorts, and contains a list of data updates in the latest release of Tracker. Section 2 describes the basic structure of Tracker and lists each variable in the data file. Sections 3, 4 and 5 discuss the different types of data available in Tracker including: permanent characteristics, ancillary and supplement study indicators, and wave-specific variables. Sections 6 and 7 of the data description outline instructions for obtaining the data and contacting HRS. Finally, appendix A provides general instructions for merging the Tracker file with other HRS data and appendix B contains a list of all updates.

The current version of the Tracker file contains information for the following panel case cohorts:

- **(AHEAD)** Asset and Health Dynamics Among The Oldest Old born before 1924
- **(HRS)** The original Health and Retirement Study cohort born 1931 - 1941
- **(CODA)** Children of the Depression born 1924 - 1930
- **(WB)** War Babies born 1942 - 1947
- **(EBB)** Early Baby Boomers born 1948 - 1953
- **(MBB)** Mid Baby Boomers born 1954 - 1959
- **(LBB)** Late Baby Boomer born 1960 - 1965
- **(EGENX)** Early Generation X born 1966 - 1971

To report any problems regarding the Tracker file or this document, please contact us via email at [hrsquestions@umich.edu](mailto:hrsquestions@umich.edu).

### ***1A. Updates in Current Tracker Release (2022 EARLY, Version 1.0)***

New variables in this current Tracker release include:

- KLBCOMP - MODE COMPLETED 2006 LEAVE BEHIND QNAIRE
- LLBCOMP - MODE COMPLETED 2008 LEAVE BEHIND QNAIRE
- MLBCOMP - MODE COMPLETED 2020 LEAVE BEHIND QNAIRE
- NLBCOMP - MODE COMPLETED 2012 LEAVE BEHIND QNAIRE
- OLBCOMP - MODE COMPLETED 2014 LEAVE BEHIND QNAIRE
- PLBCOMP - MODE COMPLETED 2016 LEAVE BEHIND QNAIRE
- QLBCOMP - MODE COMPLETED 2018 LEAVE BEHIND QNAIRE
- RLBCOMP - MODE COMPLETED 2020 LEAVE BEHIND QNAIRE
- SAGE - AGE AT 2022 INTERVIEW
- SALIVE - 2022 VITAL STATUS
- SCOUPID - 2022 WAVE HOUSEHOLD ID
- SCOUPLE - 2022 WHETHER COUPLED OR PARTNERED
- SFAMR - 2022 WHETHER FAMILY RESPONDENT
- SFINR - 2022 WHETHER FINANCIAL RESPONDENT
- SINSAMP - 2022 SAMPLE STATUS
- SIWLANG - 2022 INTERVIEW LANGUAGE
- SIWMODE - 2022 INTERVIEW MODE
- SIWMONTH - 2022 INTERVIEW MONTH
- SIWTYPE - 2022 INTERVIEW TYPE
- SIWWAVE - 2022 WHETHER INTERVIEWED IN THE WAVE
- SIWYEAR - 2022 INTERVIEW YEAR
- SLBCOMP - 2022 MODE COMPLETED LEAVE BEHIND QNAIRE
- SLBWGTR - RESPONDENT WEIGHT FOR THE 2022 LEAVE BEHIND QNAIRE
- SLIVARR - 2022 LIVING ARRANGEMENT STATUS
- SMARST - 2022 MARITAL STATUS
- SNEWSP - 2022 NEW SPOUSE-PARTNER FLAG
- SNURSHM - 2022 NURSING HOME STATUS
- SPPN - 2022 SPOUSE-PARTNER PERSON NUMBER
- SPROXY - 2022 PROXY TYPE STATUS
- SRESCODE - 2022 RESULT CODE
- SSUBHH - 2022 SUB-HOUSHOLD IDENTIFIER
- SSUBHHIW - 2022 WHETHER SUBHH INTERVIEWED
- SWEB\_ELIG - 2022 WHETHER ELIGIBLE FOR WEB
- SWEB\_INVITE - 2022 WHETHER INVITED TO WEB
- SWGTR - 2022 WEIGHT: RESPONDENT LEVEL
- SWGTRNH - 2022 WEIGHT: NURSING HOME RESIDENT
- SWHY0HWT - 2022 WHY ZERO HOUSEHOLD LEVEL WEIGHT
- SWHY0RWT - 2022 WHY ZERO RESPONDENT LEVEL WEIGHT
- SWGTHH - 2022 WEIGHT: HOUSEHOLD LEVEL
- SPMWGTR - RESPONDENT WEIGHT FOR THE 2022 PHYSICAL MEASURES

Existing variables updated in this current Tracker release include:

- Updated death dates and last known alive dates: EXDEATHMO (n = 1,194), EXDEATHYR (n = 1,207), EXDODSOURCE (n = 1,207),

KNOWNDECEASEDMO (n = 1,457), KNOWNDECEASEDYR (n = 1,436),  
KNOWNDECEASEDSOURCE (n = 1,526), LASTALIVEMO (n = 13,860),  
LASTALIVEYR (n = 15,311), LASTALIVESOURCE (n = 1,795)

- Updated cases that were missing a value: (BIRTHMO (n = 15), BIRTHYR (n = 15), DEGREE (n = 29), FIRSTIW (n = 39), GENDER (n = 14), HISPANIC (n = 39), IMMGYEAR = (n = 1), RACE (n = 37), SCHLYRS (n = 39), USBORN (n = 39), WTCOHORT (n = 15)
- Updated CAMS21 for 031059/031 and 031059/032.
- Updated SFAMR & SFINR: Code 3 (2nd financial respondent, answer not retained) has been dropped and will no longer be used moving forward.

## **2. The Structure of Tracker**

### **2A. Variable Listing and Description**

Table 1 below lists all variables in the current version of Tracker in the order they appear in the data file and the codebook. The first two variables are primary identifiers, household identifier (HHID) and person number (PN), which together uniquely identifies each HRS panel member. Following the two identifiers is a set of permanent characteristics that do not vary across waves (except to be updated or corrected). These variables either provide some basic demographic information for a respondent, or describe their study membership.

It is important to note that starting in the HRS 2006 Core interview the question wording and format for the self-assessed race questions were changed. In short, respondents were able to identify themselves as part of more than one racial category. Every respondent was asked the new version of the question. Those who indicated multiple races were then asked, "Do you consider yourself primarily (first mention, second mention, etc.)?" For each new version of Tracker the most recent Core data is used to update RACE if this value is missing in the Tracker data. In cases where more than one race is mentioned in the Core interview, the follow-up question about the respondent's primary race is used to calculate RACE. If the primary race question is not answered, the first mentioned race is used.

After the primary identifiers and the permanent characteristics, are supplemental variables related to the HRS Ancillary studies including Off-Year Mail Surveys, Internet surveys, leave-behind questionnaires, ADAMS and other indicators.

The final group of variables includes sets of wave-specific variables, with one set representing each of the HRS survey waves to date. The waves are differentiated by the leading letter in variable names, from A to Q (see Table 2). These wave-specific variables document whether a panel member was in a particular sample/study, whether he or she provided an interview in a particular wave, and if

an interview was collected, when and how the interview was conducted. When there are two panel members in a sub-household, the wave-specific variables also describe the relationship between them. Moreover, the variables include cross-sectional weights that are often needed for both household level and respondent level analyses. VERSION is also included in this group, indicating the Tracker file version number.

**Table 1. Variables in the Current Tracker: Listing and Description**

**Group A. Primary Identifiers and Permanent Characteristics**

Variable	Type	Len	Label
HHID	Char	6	HOUSEHOLD IDENTIFIER
PN	Char	3	PERSON NUMBER
BIRTHMO	Num	8	BIRTHDATE: MONTH
BIRTHYR	Num	8	BIRTHDATE: YEAR
DEGREE	Num	8	HIGHEST DEGREE OF EDUCATION
EFTFASSIGN	Num	8	RESPONDENTS ASSIGNMENT FOR EFTF INTERVIEWS
EXDEATHMO	Num	8	MONTH OF DEATH
EXDEATHYR	Num	8	YEAR OF DEATH
EXDODSOURCE	Num	8	SOURCE OF EXIT DATE OF DEATH
FIRSTIW	Num	8	FIRST INTERVIEW: STUDY YEAR
GENDER	Num	8	GENDER
HISPANIC	Num	8	HISPANICITY TYPE
IMMGYEAR	Num	8	YEAR IMMIGRATED TO THE U.S.
KNOWNDECEASEDMO	Num	8	KNOWN DECEASED - MONTH
KNOWNDECEASEDYR	Num	8	KNOWN DECEASED - YEAR
KNOWNDECEASEDSOURCE	Num	8	KNOWN DECEASED - SOURCE
LASTALIVEMO	Num	8	LAST ALIVE - MONTH
LASTLIVEYR	Num	8	LAST ALIVE - YEAR
LASTLIVESOURCE	Num	8	LAST ALIVE - SOURCE
MOSFLAG	Num	8	2010 MINORITY OVERSAMPLE FLAG
OVHHID	Char	6	OVERLAP CASE: OLD HHID
OVPN	Char	3	OVERLAP CASE: OLD PN
OVRESULT	Num	8	OVERLAP CASE: RESULT CODE
OVYEAR	Num	8	OVERLAP CASE: FIRST YEAR OF NEW HHID AND PN
RACE	Num	8	RACE/ETHNICITY
SCHLYRS	Num	8	NUMBER OF YEARS IN SCHOOL
SECU	Num	8	SAMPLING ERROR COMPUTATION UNIT
STRATUM	Num	8	STRATUM ID
STUDY	Num	8	STUDY MEMBERSHIP
USBORN	Num	8	BORN IN THE U.S.
WTCOHORT	Num	8	BIRTH COHORT AT ENTRY
YRENTER	Num	8	YEAR RESPONDENT ENTERED STUDY
HHBASEWT	Num	8	HOUSEHOLD WEIGHT AT ENTRY

**Group B. Supplemental Variables**

<b>Variable</b>	<b>Type</b>	<b>Len</b>	<b>Label</b>
ADAMS1	Num	8	ADAMS1 SAMPLE STATUS
CAMS01	Num	8	2001 CAMS SAMPLE STATUS
CAMS03	Num	8	2003 CAMS SAMPLE STATUS
CAMS05	Num	8	2005 CAMS SAMPLE STATUS
CAMS07	Num	8	2007 CAMS SAMPLE STATUS
CAMS09	Num	8	2009 CAMS SAMPLE STATUS
CAMS11	Num	8	2011 CAMS SAMPLE STATUS
CAMS13	Num	8	2013 CAMS SAMPLE STATUS
CAMS15	Num	8	2015 CAMS SAMPLE STATUS
CAMS17	Num	8	2017 CAMS SAMPLE STATUS
CAMS19	Num	8	2019 CAMS SAMPLE STATUS
CAMS21	Num	8	2021 CAMS SAMPLE STATUS
CENSUS1940	Num	8	1940 CENSUS RECORD INDICATOR
CMSLINK	Num	8	CMS SAMPLE STATUS
DIAB03	Num	8	2003 DIABETES SAMPLE STATUS
DVSA07	Num	8	2007 DVS-A SAMPLE STATUS
DVSB07	Num	8	2007 DVS-B SAMPLE STATUS
GENETICS06	Num	8	GENETICS 2006
GENETICS08	Num	8	GENETICS 2008
GENETICS10	Num	8	GENETICS 2010
GENETICS12	Num	8	GENETICS 2012
HCMS11	Num	8	2011 HEALTH CARE MAIL SURVEY SAMPLE STATUS
HCNS13	Num	8	2013 HEALTH CARE & NUTRITION MAIL SURVEY SAMPLE STATUS
HCAP_SELECT	Num	8	HCAPSELECT
HCAP16ELIG	Num	8	2016 HCAP AGE-ELIGIBLE
HCAP16REL	Num	8	2016 HCAP RELEASE NUMBER
HCAP16RESP	Num	8	2016 HCAP RESPONSE
HSM19	Num	8	2019 HEALTH SURVEY MAILER SAMPLE STATUS
HWB09	Num	8	2009 HEALTH AND WELL-BEING SURVEY SAMPLE STATUS
HUMS01	Num	8	2001 HUMAN CAPITAL MAIL SURVEY STATUS
INTRNT03	Num	8	2003 INTERNET SURVEY SAMPLE STATUS
INTRNT06	Num	8	2006 INTERNET SURVEY SAMPLE STATUS
INTRNT07	Num	8	2007 INTERNET SURVEY SAMPLE STATUS
INTRNT09	Num	8	2009 INTERNET SURVEY SAMPLE STATUS
INTRNT11	Num	8	2011 INTERNET SURVEY SAMPLE STATUS
INTRNT13	Num	8	2013 INTERNET SURVEY SAMPLE STATUS
LHMS15	Num	8	2015 LHMS SAMPLE STATUS
LHMS17SPR	Num	8	2017 SPRING LHMS SAMPLE STATUS
LHMS17FALL	Num	8	2017 FALL LHMS SAMPLE STATUS
LHMS19SPR	Num	8	2019 SPRING LHMS SAMPLE STATUS
LHMS19FALL	Num	8	2019 FALL LHMS SAMPLE STATUS
MAIL99	Num	8	1999 MAIL OUT PILOT SURVEY STATUS
PDS05	Num	8	2005 PDS SAMPLE STATUS
PDS07	Num	8	2007 PDS SAMPLE STATUS
PHYMSR04	Num	8	SAMPLE INDICATOR FOR PHYSICAL MEASURES

TELOMERE08	Num	8	2008 TELOMERE SAMPLE STATUS
VA13	Num	8	2013 VA SAMPLE STATUS
VALINK	Num	8	RECORD IN VA ENROLLMENT DATA
VBS1618WGTR	Num	8	2016-2018 WEIGHT: VENOUS BLOOD STUDY
VBS16COMPLETE	Num	8	2016 VBS COMPLETION
VBS16CONSENT	Num	8	2016 VBS CONSENT
VBS16ELIG	Num	8	2016 VBS ELIGIBLE
VBS16SELECT	Num	8	2016 VBS SAMPLE SELECTION
VBS16VALID	Num	8	2016 VBS VALID RESULT
PVBSWGTR	Num	8	2016 VBS SAMPLE WEIGHT
VBSI16WGTRA	Num	8	2016 VBS WEIGHT FOR ANALYSES INCLUDING DNA METHYLATION AND EPIGENETIC CLOCKS, TELOMERES, AND HOMOCYSTEINE
VBSI16WGTRB	Num	8	2016 VBS WEIGHT FOR ANALYSES USING CLUSTERIN OR BRAIN-DERIVED NEUROTROPHIC FACTOR
VBSI16GRPA	Num	8	2016 VBS INNOVATION VALID RESULT: GROUP A
VBSI16GRPB	Num	8	2016 VBS INNOVATION VALID RESULT: GROUP B
VBS18COMPLETE	Num	8	2018 VBS COMPLETION
VBS18CONSENT	Num	8	2018 VBS CONSENT
VBS18ELIG	Num	8	2018 VBS ELIGIBLE
VBS18SELECT	Num	8	2018 VBS SAMPLE SELECTION
VBS18VALID	Num	8	2018 VBS VALID

### Group C. Wave-Specific Variables

Variable	Type	Len	Label
XAGE	Num	8	AGE AT WAVE X INTERVIEW
XALIVE	Num	8	WAVE X VITAL STATUS
XNOCALL	Num	8	WAVE X NO CALL FLAG
XCOUPID	Char	6	WAVE X WAVE HOUSEHOLD ID
XCOUPLE	Num	8	WAVE X WHETHER COUPLED OR PARTNERED
XFAMR	Num	8	WAVE X WHETHER FAMILY RESPONDENT
XFINR	Num	8	WAVE X WHETHER FINANCIAL RESPONDENT
XINSAMP	Num	8	WAVE X SAMPLE STATUS
XIWLANG	Num	8	WAVE X INTERVIEW LANGUAGE
XIWMODE	Num	8	WAVE X INTERVIEW MODE
XIWMONTH	Num	8	WAVE X INTERVIEW MONTH
XIWTYPE	Num	8	WAVE X INTERVIEW TYPE
XIWWAVE	Num	8	WAVE X WHETHER INTERVIEWED IN THE WAVE
XIWYEAR	Num	8	WAVE X INTERVIEW YEAR
XLBWGTR	Num	8	RESPONDENT WEIGHT FOR THE WAVE X LEAVE BEHIND QNAIRE
XLIVARR	Num	8	WAVE X LIVING ARRANGEMENT STATUS
XMARST	Num	8	WAVE X MARITAL STATUS
XNEWSP	Num	8	WAVE X NEW SPOUSE-PARTNER FLAG
XNURSHM	Num	8	WAVE X NURSING HOME STATUS
XPMWGTR	Num	8	RESPONDENT WEIGHT FOR THE WAVE X PHYSICAL MEASURES SUBSAMPLE
XPPN	Char	3	WAVE X SPOUSE-PARTNER PERSON NUMBER
XPROXY	Num	8	WAVE X PROXY TYPE STATUS



XRESCODE	Num	8	WAVE X RESULT CODE
XSUBHH	Char	1	WAVE X SUB-HOUSHOLD IDENTIFIER
XSUBHHIW	Num	8	WAVE X WHETHER SUBHH INTERVIEWED
XWEBCONTROL	Num	8	WAVE X WEB CASE CONTROL GROUP FLAG
XLBCOMP	Num	8	WAVE X MODE COMPLETED LEAVE BEHIND QNAIRE
XWGTHH	Num	8	WAVE X WEIGHT: HOUSEHOLD LEVEL
XWGTR	Num	8	WAVE X WEIGHT: RESPONDENT LEVEL
XWGTRNH	Num	8	WAVE X WEIGHT: NURSING HOME RESIDENT
XWHY0HWT	Num	8	WAVE X WHY ZERO HOUSEHOLD LEVEL WEIGHT
XWHY0RWT	Num	8	WAVE X WHY ZERO RESPONDENT LEVEL WEIGHT
VERSION	Num	8	TRACKER FILE VERSION NUMBER

**2B. Naming Convention for the Wave-Specific Variables**

The wave-specific variables for each HRS wave begin with a unique letter per wave. For example, the wave-specific variables for HRS 1992 all begin with the letter A and subsequent waves follow sequentially through the alphabet, with the exception of the letter I which was not used. The specific HRS waves and their indicator letter are shown in the table below. All wave-specific variables share the same naming structures across waves. For example, ANEWSP through LNEWSP are new spouse or partner flags for HRS 1992 through HRS 2008 data waves, according to the table below, and the new spouse variable continues on in this way for each following wave thereafter. This feature has two advantages; first, it allows users to identify and link easily all the wave-specific variables, and second, it gives us the ability to name the variables specific to future HRS data waves in a predictable manner.

**Table 2. HRS Data Waves and Wave Indicators**

HRS Data Wave	Wave Indicator
HRS 1992	A
AHEAD 1993	B
HRS 1994	C
AHEAD 1995	D
HRS 1996	E
HRS 1998	F
HRS 2000	G
HRS 2002	H
HRS 2004	J
HRS 2006	K
HRS 2008	L
HRS 2010	M
HRS 2012	N

<b>HRS Data Wave</b>	<b>Wave Indicator</b>
HRS 2014	O
HRS 2016	P
HRS 2018	Q
HRS 2020	R
HRS 2022	S

## **2C. Variable Formats**

Following a protocol used in HRS data releases, the ID variables are stored in character format. In Tracker 2022, the ID variables are: HHID, PN, OVHHID (Inter-Study or Inter-Respondent Overlap HHID), OVPN (Inter-study or Inter-Respondent Overlap PN), xCOUPID (Wave-Household ID in Wave x), xPPN (Spouse or Partner Person Number in Wave x), and xSUBHH (Sub-Household ID in Wave x).

Non-ID variables (the remainder of the variables in the Tracker file) are stored in numeric format.

## **3. Permanent Characteristics**

### **3A. Demographic Information**

Basic demographic information, such as birth dates, race, hispanicity, gender, education, and immigration status, may come from a variety of sources. Most often the information was obtained initially through a respondent's answers when he or she was first interviewed. In other cases, the information may come from a spouse or partner or some other knowledgeable person if the sample person was not interviewed. In some cases, the information may have been revised, either by a spouse or partner in an exit interview, or by HRS staff, based on investigations into conflicting information for the respondent. Efforts have been made to resolve all such conflicts in a reasonable way, but it is possible that users may find some differences between the variables in the Tracker file and other HRS data. The Tracker data should be considered definitive.

### **3B. First Interview**

The variable FIRSTIW contains the year a sample member first provided an interview, including core, exit, self or proxy interviews. This variable differs from a sample member's entry cohort or study membership (STUDY) in that a sample member entering a given entry cohort or study might not have been interviewed the first time he or she was eligible for an interview. The variable YRENTER indicates the year that an individual entered the sample regardless of whether an interview was completed.

### **3C. Overlaps**

Overlaps refer to cases that have multiple primary IDs and require

special handling in constructing longitudinal files and in merging Tracker data to wave-specific files. The variables HHID and PN reflect the current status of the case, while overlap cases also have a former HHID and PN values from a previous wave. These former values are provided in the variables OVHHD and OVPN.

There are two types of overlap cases in the HRS data. First, there were a number of original HRS 1992 (Wave 1) households eligible to be either an HRS or AHEAD household. An interview was attempted for all of them in HRS 1992. Afterwards, a random subsampling was performed, resulting in 60% of the cases remaining in HRS and the remainder assigned to AHEAD. We refer to the 134 cases transferred to AHEAD as "HRS inter-study overlap" cases. Among those 134 cases assigned to AHEAD, 108 were actually interviewed in AHEAD, and 26 were not.

Second, there are a number of cases belonging to what we call "household merge overlap" resulting from intermarriage among respondents who entered the study in separate households. Below is a table that lists each of these type of overlaps. The OVHHID and OVPN columns (and corresponding variables in tracker) hold the individual's original HHID and PN values, the HHID and PN columns (and corresponding variables in tracker) hold their current ID values and the final column lists the first year in which their current HHID and PN values were used.

**Table 3. Household Merge Overlaps**

HHID	PN	OVHHID	OVPN	YEAR RESPONDENT ENTERED STUDY (YRENTER)	First year of new HHID/PN (OVYEAR)
205864	011	205906	010	1993	1998
208545	010	205399	010	1993	2000
525520	020	526934	010	2010	2012
520845	020	529766	010	2010	2012
540162	020	544849	010	2016	2018
906543	020	906126	010	2010	2018
541110	020	541986	010	2016	2020

Instructions on how to deal with the various types of overlaps when merging the Tracker file with the HRS Core data are provided in [Section 8](#).

**3D. Enhanced Face-to-Face assignments (EFTFASSIGN)**

Starting in 2004, HRS began administering physical measures and biomarker tests for our respondents. The sample is structured so that half of the respondents receive the enhanced interview each wave. The variable EFTFASSIGN holds the respondents' permanent assignment for enhanced face-to-face rotation from 2006 onward. (See also [40](#).)

[Physical Measures \(PHYMSR04\)](#) and [4P. Biomarker Data Sets \(GENETICS\)](#) below.)

The enhanced face-to-face interview includes a set of physical performance measures, collection of biomarkers, and a Leave-Behind Questionnaire on psychosocial topics. A random one-half of households were pre-selected for the enhanced face-to-face interview in 2006, with the other half of the sample selected for 2008. From that point on every household has repeated the enhanced face-to-face portion every other wave.

The specific physical measures include blood pressure, breathing test (peak flow), grip strength, timed walk (8 ft.), balance tests (semi-tandem, side-by-side, full tandem), height, weight, waist circumference, saliva (for which DNA was extracted and stored) and dry blood spots (analyzed for Hemoglobin Alc, total cholesterol and HDL cholesterol, and C-reactive protein and cystatin C).

**3E. Death dates and last known alive dates (EXDEATHMO, EXDEATHYR, EXDODSOURCE, KNOWNDECEASEDMO, KNOWNDECEASEDYR, KNOWNDECEASEDSOURCE, LASTALIVEMO, LASTALIVEYR, LASTALIVESOURCE)**

The Tracker file contains the death date variables EXDEATHMO and EXDEATHYR, which are obtained from interviews, and EXDODSOURCE, which indicates the interview source for these variables.

In addition, we provide six other variables: KNOWNDECEASEDMO, KNOWNDECEASEDYR, KNOWNDECEASEDSOURCE, LASTALIVEMO, LASTALIVEYR, and LASTALIVESOURCE. The KNOWNDECEASED variables reflect our best information about death date for deceased respondents. These variables will match EXDEATHYR and EXDEATHMO for respondents with an exit interview where a death date was provided, or whose spouse or partner has provided a death date during a core interview subsequent to the respondent's death. For other cases, this date reflects either a death reported by someone with knowledge of the respondent, or a date imputed from the last date we knew the respondent was alive and the date on which we learned the respondent was deceased.

The variables LASTALIVEMO, LASTALIVEYR and LASTALIVESOURCE represent the date on which we last knew the respondent was alive. This date comes from the most recent of: the respondent's last core interview, the last core interview of the respondent's spouse/partner during which the respondent was reported to be alive, or the last date during field activity on which we talked to the respondent or someone with direct knowledge of the respondent.

**3F. Additional Variables with Permanent Characteristics**

The remainder of the permanent characteristics variables are SECU and STRATUM, which are used for analysis of sampling error, WTCOHORT, which provides the entry-wave birth year assignment that can be used to create household level and respondent level weights, and HHBASEWT,

which provides the initial household design weights and nonresponse adjustment factors and is the base starting point for wave-specific weights.

#### **4. Ancillary Studies and Supplemental Variables**

Beginning in 1999, questionnaires on various topics have been mailed to subsamples of the HRS every other year. Indicators for these surveys are in the Tracker file and are in parenthesis in the bulleted lists below. The data for these studies can be found on [the HRS website](#) and links to the [data descriptions](#) for each study are provided in the listing below.

##### **4A. Mail Surveys**

- [1999 Pilot Mail Survey](#) (MAIL99)
- [2001 Consumption and Activities Mail Survey](#) (CAMS01)
- [2001 Human Capital and Educational Expenses Mail Survey](#) (HUMS01)
- [2003 Consumption and Activities Mail Survey](#) (CAMS03)
- [2003 Diabetes Study](#) (DIAB03)
- [2005 Consumption and Activities Mail Survey](#) (CAMS05)
- [2005 Prescription Drug Study](#) (PDS05)
- [2007 Consumption and Activities Mail Survey](#) (CAMS07)
- [2007 Disability Vignette Survey](#) (DVSA07 and DVSB07)
- [2007 Prescription Drug Study](#) (PDS07)
- [2009 Consumption and Activities Mail Survey](#) (CAMS09)
- [2009 Health and Well-Being Study](#) (HWB09)
- [2011 Consumption and Activities Mail Survey](#) (CAMS11)
- [2011 Health Care Mail Survey](#) (HCMS11)
- [2013 Consumption and Activities Mail Survey](#) (CAMS13)
- [2013 Veterans Mail Survey](#) (VA13)
- [2013 Health Care and Nutrition Study](#) (HCNS13)
- [2015 Consumption and Activities Mail Survey](#) (CAMS15)
- [2015 Life History Mail Survey](#) (LHMS15)
- [2017 Life History Mail Survey](#) (LHMS17FALL)
- [2017 Fall-Supplement Life History Mail Survey](#)
- [2017 Consumption and Activities Mail Survey](#) (CAMS17)
- [2017 Spring Life History Mail Survey](#) (LHMS17SPR)
- [2019 Consumption and Activities Mail Survey](#) (CAMS19)
- [2019 Spring Life History Mail Survey](#) (LHMS19SPR)
- [2019 Fall Life History Mail Survey](#) (LHMS19FALL)
- [2019 Health Survey Mailer](#) (HSM19)
- [2021 Consumption and Activities Mail Survey](#) (CAMS21)
- 

##### **4B. Internet Surveys**

- [2003 Internet Survey](#) (INTRNT03)
- [2006 Internet Survey](#) (INTRNT06)
- [2007 Internet Survey](#) (INTRNT07)

- [2009 Internet Survey](#) (INTRNT09)
- [2011 Internet Survey](#) (INTRNT11)
- [2013 Internet Survey](#) (INTRNT13)

**4C. National Death Index (NDI) – No longer in Tracker**

The HRS contract agreement with the providers of the National Death Index has expired so relevant variables (NSCORE, NMONTH, NYEAR, and xNDIFLAG) were removed. We hope to update them in a subsequent release. Please see section 3E for alternative death date variables.

**4D. The Ageing, Demographics, and Memory Study**

- [The Aging, Demographics, and Memory Study](#) (ADAMS)

**4E. Physical Measures (PHYMSR04)**

In 2004, HRS administered a set of physical measures to a random subsample of about 3,300 respondents. In 2006, this effort was expanded in terms of both the size of the sample covered and the measurements conducted, as HRS initiated what is referred to as the Enhanced Face-to-Face interview. The variable PHYMSR04 in the Tracker file indicates whether a respondent was selected to participate in the physical measures portion of the interview in 2004, and if selected, whether they completed the measures. For 2006 and beyond, the variable **EFTFASSIGN** (see 3D above) holds the respondents' permanent assignment for enhanced face-to-face rotation, which includes the Physical Measures component.

**4F. Biomarker Data Sets**

- [2006 Biomarker Data](#) (GENETICS06)
- [2008 Biomarker Data](#) (GENETICS08)
- [2010 Biomarker Data](#) (GENETICS10)
- [2012 Biomarker Data](#) (GENETICS12)
- [2014 Biomarker Data](#) (GENETICS14)
- [2016 Biomarker Data](#) (GENETICS16)

**4G. Leave-Behind Questionnaires (DISAB04, PSYSOC04, PSYSOC06, PSYSOC08, PSYSOC10, PSYSOC12, PSYSOC14: NO LONGER IN TRACKER)**

The variables **DISAB04**, **PSYSOC04**, **PSYSOC06**, **PSYSOC08**, **PSYSOC10**, **PSYSOC12** and **PSYSOC14** have been removed from the Tracker file. For 2004, they are replaced by the variables JLBTYP and JV900 in Section LB of the core interview. For 2006 onward, they are replaced by xLBELIG and xLBCOMP in section LB of the core interview.

Please see the [Leave Behind Questionnaire documentation report](#) for more information about the Psychosocial measures.

**4H. Harmonized Cognitive Assessment Protocol**

- [2016 Harmonized Cognitive Assessment Protocol](#) (HCAP\_SELECT)

Sample indicators for the Harmonized Cognitive Assessment Protocol are included in Tracker 2018. In order to have a permanent sample pool for HCAP assessments the HRS randomly selected one-half of all HRS participants to be potential HCAP sample members. The sample selection was initially made prior to the 2016 wave for all in-sample cases and will be continuously updated for new entrants using the following process:

- 1) Randomly select 1/2 of uncoupled Respondents
- 2) Randomly select 1 Respondent from each coupled household
- 3) New spouses are given the opposite assignment of the HRS panel member they are partnered with
- 4) New cohort entrants will be assigned using 1) and 2)

HCAP selection status is indicated in the variable HCAP\_SELECT.

#### **HCAP 2016 Assessment**

The first HCAP assessment was conducted in 2016 and 2017. From the sample pool selected for HCAP, eligibility for the 2016 assessment was based on age and limited to those who completed an HRS interview in the 2016 wave. In order to have the HCAP assessments represent the population 65 and older at assessment, persons at least 64 and 10 months of age at their 2016 HRS interview were age-eligible for HCAP. This is indicated in the variable HCAP16ELIG.

In order to manage the HCAP fieldwork and maintain a short interval between the HRS interview and the HCAP assessment, eligible HCAP cases were released for HCAP fieldwork on a rolling basis. This is indicated in the variable HCAP16REL.

#### **4I. Venous Blood Study (VBS), 2016, 2018**

- [2016 Venous Blood Study](#) (VBS16SELECT)
- 2018 Venous Blood Study (VBS18SELECT)

#### **4J. 1940 Census**

Sample indicators for the 1940 Census are included in the Tracker. HRS attempted to locate in the 1940 Census every HRS sample member who was alive by the time of the decennial enumeration on April 1, 1940. This includes the bulk of the original HRS cohort, all of the AHEAD cohort, and all of the CODA cohort. In all we attempted to link 20,065 HRS respondents. More information about the 1940 Census can be found [here](#) on the HRS web site.

#### **4K. CMS Linkage**

This variable indicates the status of CMS Medicare and/or Medicaid MAX linkage, including whether it was linked to Medicare claims, Medicare and Medicaid MAX or Medicaid MAX only.

## **5. Wave-Specific Variables**

The wave-specific variables in the Tracker file may be divided into three groups: 1) those indicating a sample member's interview type, sample status, or study status in a given wave; 2) those indicating the relationships between the sample members in a household in a given wave; and 3) those describing other wave-specific information.

### **5A. Interview, Sample and Study Status**

#### **5A1. *xIWTYPE*, *xIWWAVE*, and *xINSAMP***

In most of the HRS data waves, a sample member could fall into one of seven categories within the *xIWTYPE* variable:

- (1) Core interview obtained
- (5) Core interview not obtained
- (11) Exit interview obtained
- (15) Exit interview not obtained
- (21) Post-exit interview obtained
- (25) Post-exit interview not obtained
- (99) Not in the sample this wave

This is the first variable a user should consult when checking a sample member's status in any given wave. By design, sample members in HRS 1992 or AHEAD 1993 can appear only in categories (1) and (5), and sample members in HRS 1994, AHEAD 1995 or HRS 1996 will not be in categories (21) and (25).

*xIWWAVE* is a flag variable which collapses the interview type information in *xIWTYPE* into a single YES (1) or NO (0) dichotomy. *xIWWAVE* is equal to 1 if a sample member provided any type of interview; otherwise it is equal to 0 if the sample member was eligible but did not provide an interview. If a sample member was not eligible to be interviewed at a particular wave, *xIWWAVE* is blank.

While both *xIWTYPE* and *xIWWAVE* indicate whether an interview was obtained for a sample member in a given wave, *xINSAMP* indicates whether a sample member was supposed to be interviewed in the wave, and, if not, why. By design, some sample members should not be in certain data waves. Examples include:

1. The HRS members that are not HRS inter-study overlaps should not be in an AHEAD data wave, and vice versa
2. The CODA or WB respondents should not be in any data wave prior to HRS 1998
3. The EBB respondents should not be in any wave of data prior to HRS 2004, etc.

Tracker 2022 defines six types of *xINSAMP* statuses:

- (1) In the sample



- (4) Formally dropped from the sample: HRS-AHEAD overlap never appeared in AHEAD
- (5) No longer in sample because complete exit or post-exit interview has been obtained
- (6) Formally dropped from the sample before the wave: Not known to be deceased, termination per request by respondent, spouse, or other living proxy
- (7) Formally dropped from the sample before the wave: Deceased, unable to find exit proxy
- (8) Formally dropped from the sample: Other reasons

Which of these categories to include in non-response analyses will depend on the specific purpose, but in general, categories (4), (6), (7), and (8) should be regarded as eligible and therefore included in the denominator when calculating non-response and attrition rates.

The relationship among `xIWTYPE`, `xIWWAVE`, and `xINSAMP` is summarized below in Table 4.

**Table 4. The Relationship among `xIWTYPE`, `xIWWAVE`, and `xINSAMP`**

<code>xIWTYPE</code>	<code>xIWWAVE</code>	<code>xINSAMP</code>
1. Core interview provided 11. Exit interview provided 21. Post-exit interview provided	1. Yes	1. (Yes)
5. Core interview NOT provided 15. Exit interview NOT provided 25. Post-exit interview NOT provided	0. No	1. (Yes)
99. Not in the sample this wave	Blank.	4, 5, 6, 7, 8. (No)

**5A2. `xRESCODE`**

Part of the HRS tracking procedure is to note reasons for non-interviews. For each sample member eligible to be interviewed in a wave, the variable `xRESCODE` indicates the final disposition of the case. Of particular importance is the distinction between a non-interview for the current wave only (`xRESCODE` codeframes between 5000 and 8999) and a permanent removal from the study (`xRESCODE` codeframes beginning with 9000).

Each sample member eligible to be interviewed in a wave will have a non-missing value on `xRESCODE`. The sample members of wave X with missing values on `xRESCODE`, therefore, will be those with a value other than 1 on `xINSAMP`.

### **5A3. xSUBHH, and xSUBHHIW**

In HRS, unique sub-households are identified by combining the household ID (HHID) AND sub-household ID (xSUBHH), and are usually regarded as the unit of household-level analysis. An original household (xSUBHH = 0) may split off into a new xSUBHH for one of two reasons: dissolved marriage or partnership, or death of one of the partners. The sub-household ID for an eligible deceased sample member in a given wave is "3" or "4"; these values are never assigned to a sample member who is still alive. A respondent who is reported as deceased during the field period, but for whom we have not obtained an exit interview, will generally not be moved into xSUBHH of "3" or "4" until the exit interview is obtained, or until the surviving spouse / partner reports the death in Section A of their own core interview.

xSUBHHIW describes whether a unique sub-household is interviewed in a given wave. A sub-household is considered interviewed in a wave (xSUBHHIW=1) if any of the members in the sub-household provided an interview in the wave, regardless of whether it was a core, exit or post-exit interview; and if it was the interview was complete or partial. It is possible, therefore, for a sample member not interviewed in a wave to have a value of 1 on xSUBHHIW. On the other hand, when no members in a sub-household were part of the sample in a wave (xINSAMP>1), xSUBHHIW will be missing for all the respondents in the sub-household.

### **5B. Inter-Respondent Relationship**

#### **5B1. xCOUPLE, xCOUPID, xMARST, and xLIVARR**

When an eligible sample member in a given wave is coupled (married or living with a partner as if married), the variable **xCOUPLE** is equal to 1. Otherwise, it is 5. If one member of a couple is reported dead in a core or exit interview in wave x, the deceased respondent is assigned xSUBHH of "3" or "4", with xCOUPLE=1, while the surviving spouse or partner will be assigned xCOUPLE=5. (However, if the surviving spouse or partner has remarried or re-partnered, or was interviewed before the spouse / partner's death, xCOUPLE will equal 1).

Labeled as "xWave Household ID," **xCOUPID** is created to link the sample members who are married or partnered in a given wave. For sample members with xCOUPLE=1, xCOUPID is a concatenation of the person numbers of the sub-household members in ascending order. XCOUPID is also useful when a death has occurred. When one member of a couple is reported dead at wave x, and the partnership or marriage had not ended prior to the death, the deceased respondent is assigned xCOUPLE=1 and xCOUPID is established to describe the respondent's coupleness at death to facilitate linkage to the prior wave spouse or partner. The surviving spouse/partner is assigned 5 for xCOUPLE and xCOUPID is formed from his/her own PN followed by 000, unless he or

she has formed a new couple.

**xMARST** indicates whether a respondent is married, divorced/separated, widowed, or has never been married. This variable is available beginning with HRS 2004 Core data. It is important to note that xMARST may be different from the marital status assigned in the core released data. The marital status variable in the Tracker file (xMARST) was constructed by examining several data sources. In cases of disagreement the marital status variable in the Core interview was not changed, in order to preserve the flow through the instrument.

The basic procedure for generating xMARST is as follows. We first use core interview data to determine the marital status for those who provided core interviews at a given wave, thus creating xMARST. This variable is then modified to achieve cross-wave consistency based on one or more of the following: a) spouse/partner relationship, b) examination of cross sectional and longitudinal data, c) extrapolation based on the data, and sometimes d) common sense.

The "common sense" method stipulates that if a respondent was "never married" at wave t, then s/he was not married at any wave prior to t. Conversely, if a respondent was married, separated, or divorced at wave t, then s/he could not be "never married" at any wave after t. The "spouse/ partner relationship" method assumes that in a coupled household at wave t, if one respondent was married, the other should be married at the wave as well; and if one respondent was not married, the other should not be married either. The "extrapolation" method allows one to use a respondent's marital status at a given wave (t-1) to "approximate" his or her status at wave t when the latter information is not available.

**xLIVARR** describes a respondent's living arrangements and is available beginning with HRS 2002 Core data. For this variable, a respondent's marital status takes priority over any other living arrangements, so a married respondent who is separated and in another live-in relationship will be coded as "2 Married or partnered, not living with partner."

### **5B2. xFAMR and xFINR**

HRS respondents provide information at an individual level and, for some questions, at a household level. In particular, many questions about family or about income and wealth are asked of only one respondent in a household. The respondent providing family information in a given wave is called the Family Respondent (or Family R) in that wave, and the respondent providing financial data in a wave is called the Financial Respondent (or Financial R) in the wave. The family R and financial R designations are reflected in variables xFAMR and xFINR. For a single-person household, the respondent is designated to be both Family R and Financial R.

### **5B3. *xNEWSP***

When an original sample member takes a new spouse or partner following his or her baseline interview, an attempt is made to interview the new person. New sample members entering in this way are indicated with *xNEWSP*=1 for the wave in which an interview was first attempted, and *XNEWSP*=0 for that person in all other waves. There are no new spouses or partners in baseline waves. Thus, no sample members in HRS 1992 or AHEAD 1993 are considered new spouses or partners, and similarly, no sample members from the CODA, WB (1998), EBB (2004) and MBB (2010) entry cohorts are considered new spouses or partners.

### **5B4. *xPPN***

*xPPN* identifies the person number of the spouse or partner of a sample member if the sample member is part of a couple. For all the sample members still alive in a given wave, a concatenation of their own person numbers (PN) with *xPPN* in ascending order would generate *xCOUPID*. For deceased sample members, *xPPN* gives the person numbers of their spouses or partners in the last wave during which they were alive.

### **5B5. *xPROXY***

The HRS makes every effort to obtain core interviews directly with respondents themselves. In cases where the respondent is unavailable, unable, or unwilling to do the interview, a proxy is sought to provide an interview on behalf of the respondent (unless the respondent refuses to allow a proxy). Exit and post-exit interviews, on the other hand, are always conducted with proxies. The variable *xPROXY* indicates self-interviews, and classifies proxies into two types: spouse/partner, and any others.

## **5C. *Other Wave-Specific Variables***

### **5C1. *xALIVE***

This variable contains information about vital status based on core and exit interviews and HRS tracking information. It provides four categories:

- (1) Alive at this wave
- (2) Presumed alive as of this wave
- (5) Known deceased as of this wave
- (6) Known deceased as of prior wave
- (Blank) Not in the sample this wave

If the respondent gave an interview in the wave or was contacted directly by an interviewer during the wave, or was reported to be alive by a spouse or partner, they are considered definitely alive (*xALIVE*=1). Similarly, anyone who met those criteria at a subsequent wave will be classified as *xALIVE*=1 at all prior waves. If contact of a less definitive type was made, and there was no report of death,

we classify the respondent as presumed alive (xALIVE=2).

No National Death Index information is used in constructing xALIVE. Both sources of information should be used to classify vital status according to the goals of any specific analysis.

### **5C2. xIWLANG, xIWMODE, xIWMONTH, xIWYEAR, and xAGE**

These variables indicate the language (English or Spanish) in which the interview was conducted, the mode (in-person, telephone, web), the month and year in which the interview began and the age the respondent was at a specific interview. If more than one birth date was provided in different waves, xAGE is constructed using a 'best source' method. A self-report was considered the most reliable, followed by a spouse report, and finally a non-spouse proxy report. The public data calculates age using the last date of birth given, regardless of the source. This can result in some discrepancies between xAGE and BIRTHYR. xAGE is calculated only for respondents who completed a core interview.

### **5C3. xNURSHM**

xNURSHM indicates whether the respondent was reported to be in a nursing home or other health care facility at the time of the respondent's interview or the spouse/partner's interview. This can be reported either by the respondent or a proxy during the respondent's interview at question A028, or by a spouse or partner describing the respondent's whereabouts. This definition of nursing home is broader than the Census definition of nursing home as a category of institutional group quarters. The HRS definition includes some people living in other facilities such as assisted living facilities. Beginning with the 2016 wave (PNURSHM), multiple data sources were used to make individual reassignments of residence type that are included in this Tracker file release.

HRS follows respondents longitudinally into nursing homes or other facilities. By design, HRS does not attempt a baseline interview with potential new respondents who are nursing home residents. Given the young age at recruitment of recent cohorts, this policy had significant impact only for the 1993 baseline wave of AHEAD, where the decision was that interviews would not be attempted for nursing home residents, even if they were part of a couple. We did decide to include these AHEAD nursing home spouses or partners in subsequent waves, however, and attempted to interview them. Current policy is to attempt a baseline interview with a nursing home resident who is part of a couple whose other member is part of the non-institutionalized population.

### **5C4. xNOCALL**

Beginning in 2018 as a cost-saving measure HRS selects a small number of households each wave to set aside to not be called for interviews. They receive advance letters and are given a number to

call if they wish to be interviewed. Households are selected based on prior history of high numbers of calls. Exemptions are made for minority respondents. By design, no household will be in this category two waves in a row. This protocol was stopped after the 2018 wave.

#### **5C5. *xWEBCONTROL***

Beginning in 2018 HRS offers some households a self-administered web interview as their initial option for the core survey. Households in the enhanced face-to-face group for the wave are not eligible for web interview that wave. To enable studies of the impact of web mode on response rates, data quality, and data values, about 40% of the households eligible for web were kept in a control group in their usual survey mode. For the 2022 wave, *xWEBCONTROL* was replaced with *xWEB\_ELIG* and *xWEB\_INVITE*

#### **5C6. *xWEB\_ELIG* and *xWEB\_INVITE***

In 2022, HRS is administering its third wave of Web as an alternate mode for data collection for the core biennial interview. Web was offered as an alternative for telephone and regular face-to-face respondents only; respondents in the half-sample scheduled for the enhanced face-to-face interview were not eligible for Web this wave. Households previously selected as Web eligible 2018 remain Web eligible as long as they continue to meet eligibility criteria based on a prior report of internet access along with other selection criteria (e.g., English speaking, self-respondent, non nursing-home resident in prior wave). In addition, the 2016 LBB new cohort households are also eligible for Web based on the same eligibility criteria.

In total, 3,317 eligible households were identified. By design, 60% of eligible households were to be assigned to receive a web invitation, with the remainder receiving their usual mode of either telephone or face-to-face. This resulted in 1,989 eligible households receiving the Web invitation and 1,328 eligible households not receiving the Web invitation. The intention was for this assignment to be random across cohorts and households. However, due to a programming error the designation was not random and should not be used to study mode effect in the 2022 wave.

### **5D. HRS Sample Weights**

#### **5D1. *xWGTHH*, *xWGTR*, *xWGTRNH*, *xWHYOWGT*, *xWHYOHWT*, and *xWHYORWT***

*xWGTHH* is the sampling weight for analysis at the household level for community-dwelling households. *xWGTR* is the sampling weight for

analysis at the respondent level for respondents living in the community. xWGTRNH is the sampling weight for analysis at the respondent level for respondents living in nursing homes. xWGTRNH has been constructed for each wave since 2000.

Since the HRS sample is a multi-stage area probability sample of the United States, with oversamples of Blacks and Hispanics, unbiased estimates of population parameters require the use of these sampling weights. Through 2002, the household level and respondent level weights (xWGTHH and xWGTR) are post-stratified to the March Current Population Survey (CPS) for the year of data collection. Starting in 2004, the much larger ACS is used in place of the CPS for post-stratification purposes. Although the HRS samples only community dwellers at baseline, it follows respondents longitudinally into nursing homes and beginning in 2000 can represent that population reasonably well. The nursing home sampling weight (xWGTRNH) is designed to post-stratify nursing home residents interviewed in a given wave (2000 and subsequent waves) to represent the national population of nursing home residents by race, sex, and age in that year. Population totals by age and sex, and by race, are based on Census data in 2000 and 2010, with intercensal interpolations based on nursing home populations from the Minimum Data Set (MDS). Only nursing home residents who are age 55 or older and interviewed in the wave have a non-zero value for xWGTRNH.

The (community dwelling) respondent level weight xWGTR is zero when a respondent is not cohort eligible in wave x, is residing in a nursing home, or is deceased. There are also a very few cases with zero xWGTR because their birth cohort was not ascertained. The variable xWHY0WGT summarizes which of these conditions is responsible for a zero respondent level weight for cases with a core interview in wave x (available for waves 1992 through 2002). For the 2004 and later waves, xWHY0HWT explains why a household has a zero weight and xWHY0RWT explains why a respondent has a zero weight.

Prior to 2004, HRS weights are generated based on WTCOHORT, not BIRTHYR. WTCOHORT is constructed using a "first-mention" or "entry-wave" birth year measure. This information may not be completely consistent with BIRTHYR, which takes into account corrections by the HRS staff to the birth year measure constructed in WTCOHORT. Starting in 2004, BIRTHYR has been used to generate the weights. Conversely, after considerable investigation and discussion with the NIA Data Monitoring Committee, race and ethnicity is now based exclusively on a "first-mention" basis. Any post-baseline reports of race are not used.

In the 2012 release of the Tracker file, the 2004-2010 weights were updated to incorporate several changes. First, the post-stratification method used to generate the household and respondent-level weights for the 2004 and later waves was modified to use ACS rather than CPS. Second, the updated weights are based on the first

mention of race and ethnicity, reversing some changes that were made to the race and ethnicity variables in some previous versions of Tracker. Finally, non-original sample members who are no longer coupled with an original sample member are now assigned non-zero weights, so long as they are cohort eligible. As a result of these updates, the core household and respondent-level weights for the 2004-2010 waves will differ slightly from those in earlier releases of the Tracker file.

For 2016 we provide the same three types of weight variables. However, the methods of construction differ from prior years in two key groups. The newly recruited Late Baby Boom (LBB) cohort had its initial sampling weights created with additional control variables in the non-response adjustment and in the final post-stratification. Secondly, nursing home residence was changed from self-report to be more consistent with the Census definition used in the ACS totals used for post-stratification.

A careful review has determined that a significant fraction of HRS respondents who report at question A028 living in nursing homes or other facilities, live in facilities such as assisted living facilities, that are not classified by the Census as nursing home institutional group quarters. Multiple data sources were used to make individual reassignments of residence type that are included in the 2020 Tracker file release and subsequent releases.

Starting in 2016, HRS post-stratifies community-dwelling respondents using the ACS 1-year Public Use Micro Sample (PUMS) data from the year corresponding to the survey year (e.g., ACS 2016 1-year data for HRS 2016). Because ACS PUMS does not include nursing home as a separate category, the nursing home resident post-stratification uses the ACS 5-year Summary File (SF) and intercensal estimates of nursing home residents.

Consistent with past years, households where the only member or households where both members were in nursing homes are assigned zero values for the household weight (xWGTHH).

HRS post-stratification is done at the household level as well as the respondent level. At the household level, population totals of household cohort, race/ethnicity of household members, coupleness status, household type and counts of financial units are used as done prior to 2016. In 2016, household post-stratification additionally controls for of region, education of household members and labor force status of household members for the LBB cohort households. Respondent level post-stratification differs by dwelling type. Community respondents are post-stratified to population totals of cohort, marital status, sex and race/ethnicity as done prior to 2016. Specifically for LBB respondents in newly recruited LBB households, population totals of region, education and labor force status are additionally used. Post-stratification for nursing home respondents controls for age, sex and race.



### **5D2. Sample Weights for 2004 Physical Measure and Leave-Behind Questionnaire (JPMSELWT, JPMWGTR, JWGTR\_PS AND JWGTR\_DB)**

Sample weights were developed for the 2004 Physical Measures and Leave-Behind components to account for differential probabilities of selection into the subsample and non-response to the specific component. These variables include:

- Physical Measures
  - PHYMSR04- sample selection indicator for 2004 physical measures
  - JPMSELWT - selection weight for the 2004 physical measures subsample
  - JPMWGTR - final respondent weight for 2004 physical measures subsample
- Psychosocial Leave-Behind
  - JWGTR\_PS - final respondent weight for 2004 psychosocial subsample
- Disability Leave-Behind
  - JWGTR\_DB - final respondent weight for 2004 disability subsample

A general description of the physical measures and leave-behind components can be found in the [2004 Core Data Description](#). Questionnaires for each are available on the HRS website:

- [Physical Measures](#)
- [Psychosocial Leave-Behind](#)
- [Disability Leave-Behind](#)

More detailed information on the selection indicators and sample weights is provided in the following documents:

- [Sample Weights and Sample Selection Indicator for the Physical Measures in HRS 2004](#)
- [Sample Weights, Sample Selection Indicators and Response Rates for the Psychosocial and](#)

### **5D3. Sample Weights for the 2006 and Subsequent Wave Physical Measures, Biomarkers, and Psychosocial Leave Behind Questionnaire**

Sample weights were developed for the physical measure and biomarker components separately. Respondents who completed at least one physical measure were assigned a physical measures weight and respondents with at least one valid biomarker result were assigned a biomarker weight. The weights were adjusted for the differential probabilities of participation by dividing the HRS 2006 sample weight by the predicted probability of response to each component. The

resulting interim weight was trimmed at the 5<sup>th</sup> and 95<sup>th</sup> percentiles and was then post stratified back to the entire 2006 HRS sample by age, sex, and race/ethnicity. The sample weights for each component are provided in the Tracker file. These include:

- EFTFASSIGN - enhanced face-to-face sample indicator
- KPMWGTR - final respondent weight for 2006 physical measures subsample
- KBIOWGTR - final respondent weight for 2006 biomarker subsample

More information on the physical measures component including details on sampling, consent, administration, and rationale is provided in the following [User Guide](#).

Information on the biomarker collection can be found in the [Data Description for the 2006 Biomarker Data file](#).

Separate respondent level survey weights were constructed to adjust for non-response to the Leave Behind component. The HRS 2006 sample weight was adjusted for non-response by dividing it by the predicted probability of response to the questionnaire. Age, sex, race/ethnicity, marital status, health status, depressive symptoms, low cognition, IADL/ADL difficulties, mobility, poor vision, work status, and religious attendance were used in the non-response model. The resulting interim weight was trimmed at the 5<sup>th</sup> and 95<sup>th</sup> percentiles and was then post stratified back to the entire 2006 HRS sample by age, sex, and race/ethnicity. The sample weights, sample selection and eligibility indicators for the psychosocial component include the following variables:

- EFTFASSIGN - enhanced face-to-face sample indicator
- KLBWGTR - final respondent weight for the 2006 leave behind questionnaire subsample

More information about the questionnaire can be found in the [User Guide](#).

Sample weights for subsequent Physical Measures, Biomarker, and Psychosocial Leave-Behind Questionnaire waves were developed using the same methodology as in 2006. The sample weights for each component are provided in the Tracker file. These include:

2008

- o LPMWGTR - final respondent weight for the 2008 physical measures subsample
- o LBIOWGTR - final respondent weight for the 2008 biomarker subsample
- o LLBWGTR - final respondent weight for the 2008 leave behind questionnaire subsample

2010

- o MPMWGTR - final respondent weight for the 2010 physical measures subsample
- o MBIOWGTR - final respondent weight for the 2010 biomarker subsample
- o MLBWGTR - final respondent weight for the 2010 leave behind questionnaire subsample

2012

- o NPMWGTR - final respondent weight for the 2012 physical measures subsample
- o NBIOWGTR - final respondent weight for the 2012 biomarker subsample
- o NLBWGTR - final respondent weight for the 2012 leave behind questionnaire subsample

2014

- o OPMWGTR - final respondent weight for the 2014 physical measures subsample
- o OBIOWGTR - final respondent weight for the 2014 biomarker subsample
- o OLBWGTR - final respondent weight for the 2014 leave behind questionnaire subsample

2016

- o PPMWGTR - final respondent weight for the 2016 physical measures subsample
- o PBIOWGTR - final respondent weight for the 2016 biomarker subsample
- o PLBWGTR - final respondent weight for the 2016 leave behind questionnaire subsample

2018

- o QPMWGTR - final respondent weight for the 2018 physical measures subsample
- o QBIOWGTR - final respondent weight for the 2018 biomarker subsample
- o QLBWGTR - final respondent weight for the 2018 leave behind questionnaire subsample

2020

- o RLBWGTR - final respondent weight for the 2020 leave behind questionnaire subsample

2022

- o SLBWGTR - respondent weight for the 2022 leave behind questionnaire subsample
- o SPMWGTR - respondent weight for the 2022 physical measure subsample

#### **5D4. Sample Weights for the 2016 Venous Blood Study - Full sample (PVBSWGTR) and Innovative Sub Sample (VBSI16WGTRA and VBSI16WGTRB)**

Respondents with at least one valid venous blood result (VBS16VALID) were assigned a VBS weight. The weights were adjusted for the differential probabilities of participation by dividing the HRS 2016 sample weight by the predicted probability of having a valid venous blood result among community-dwelling 2016 HRS respondents born prior to 1960, excluding all members of the LBB cohort. The resulting interim weight was trimmed at the 1st and 99th percentiles and was then post stratified back to the entire 2016 HRS sample born prior to 1960 by age, sex, and race/ethnicity. Two separate respondent-level weights were created for the VBS 2016 Innovative Sub Sample and should be used for analyses of data from that sample. VBSI16WGTRA should be used for analyses including DNA methylation and epigenetic clocks, telomeres, and homocysteine. VBSI16WGTRB should be used for analyses using clusterin or brain-derived neurotrophic factor.

More information on the 2016 Venous Blood Study, including details on sampling, consent, and administration, is provided in the VBS 2016 Data Description.

Note: PVBSWGTR replaces the preliminary VBS weight (vbswt2) that was released with the Venous Blood Study 2016 Early V1.0 data (Dec 2017).

#### **5D5. VBS 2016 and 2018 Combined Respondent Weight (VBS1618WGTR)**

The 2018 VBS data include respondents who were new sample members in 2016 (the LBB cohort), new spouse/partners in 2018, or who did not provide a blood sample in 2016 but were asked and provided a sample in 2018. As a result, the 2018 data cannot be used alone to represent the HRS sample. We recommend combining the 2016 and 2018 VBS samples for analyses of the entire HRS range of age-eligible respondents (51+ years). To facilitate this, we have created a combined weight for analysis. Age-eligible respondents with at least one valid venous blood result in 2016 or 2018 (VBS16VALID and VBS18VALID) were assigned a VBS weight. The 2016 and 2018 core weights were used as a starting point. The core weights were adjusted for the differential probabilities of participation by dividing the core HRS sample weight by the predicted probability of having a valid venous blood result among community-dwelling HRS 2016 and 2018 respondents. The resulting interim weight was trimmed at the 1st and 99th percentiles and was then post stratified back to the entire 2016 HRS sample born prior to 1966 by age, sex, and race/ethnicity. Analysts should use VBS1618WGTR to weight the data when using VBS 2016 together with VBS2018 data.

## **6. Obtaining the Data**

### **6A. Registration and Downloading the Data**

HRS data are available for free to researchers and analysts at the HRS website. In order to obtain public release data, you must first register at our [website](#). Once you have completed the registration process, your username and password will be sent to you via e-mail. Your username and password are required to download any data files.

By registering all users, we are able to document for our sponsors the size and diversity of our user community allowing us to continue to collect these important data. Registered users receive user support, information related to errors in the data, future releases, workshops, and publication lists. The information you provide will not be used for any commercial use, and will not be redistributed to third parties.

### **6B. Conditions of Use**

By registering, you agree to the [Conditions of Use](#) governing access to Health and Retirement public release data.

### **6C. Publications Based on Data**

As part of the data registration process, you agree to include specified citations and to inform HRS of any papers, publications, or presentations based on HRS data. Please send a copy of any publications you produce based on HRS data, with a bibliographical reference, if appropriate, to the address below.

Health and Retirement Study  
Attn: Papers and Publications  
The Institute for Social Research  
P.O. Box 1248  
Ann Arbor, MI (USA) 48106-1248

You may also contact us by e-mail at [hqsquestions@umich.edu](mailto:hqsquestions@umich.edu) with "Attn: Papers and Publications" in the subject line.

## **7. If You Need to Know More**

This document is intended to serve as a brief overview and to provide guidelines for using the Tracker 2022 data. If you have questions or concerns that are not adequately covered here or on our [website](#), or if you have any comments, please contact us. We will do our best to provide answers.

### **7A. HRS Internet Site**

Health and Retirement Study public release data and additional information about the study are available on the Internet. To access the data and other relevant information, please visit the HRS [website](#).

**7B. Contact Information**

Internet: Help Desk at our website

E-mail: hrsquestions@umich.edu

Postal service:

Health and Retirement Study  
The Institute for Social Research  
The University of Michigan  
P.O. Box 1248  
Ann Arbor, MI 48106-1248

FAX: (734) 647-1186

## APPENDIX A

### **A. Merging the Tracker with other HRS Data**

When merging the Tracker data with other HRS data, you should remember that the Tracker contains more records than other data releases for any given wave. The first step is, therefore, to subset the records from the Tracker that are present in the other data file. `xIWTYPE` provides necessary information for performing this task. In addition, since the Tracker contains only one record per sample member, the four respondents with overlap identifiers (as described in 3C Overlaps) will not match with other HRS data without some manipulations before merging. In fact, dealing with the respondents with multiple identities across waves proves to be the most complicated task when integrating the Tracker data with other HRS data.

In what follows, we provide some coding examples for merging the Tracker with various HRS data, covering most of the issues that users may encounter when performing similar merges of their own. For simplicity, we assume that all the data, including the Tracker and other HRS data, are stored in the same directory (or SAS data library).

#### **A.1. HRS 1992 Merges**

Merging the Tracker with the HRS 1992 data is complex because of the HRS-AHEAD overlaps explained in Section 3C. Two different strategies are available for working with HRS 1992 data. First, one may consider the entire, original HRS sample as the data source of the analysis. In this strategy, one wants to retain all the data records as released in the HRS 1992 Core data, that is,  $N=12,652$ . Alternatively, one may consider only the non-overlap cases as the HRS respondents, leaving the HRS overlaps to AHEAD. In this strategy, the information in the HRS 1992 Core data for the overlap cases will be dropped, and  $N=12,521$ .

While we have no intention to force users to take one strategy over the other, it is suggested that the second strategy be used whenever the HRS 1992 data are involved in a longitudinal analysis. The overlap cases are no longer treated as part of the HRS sample in the later waves.

The following examples illustrate these two strategies when merging with a dataset called HEALTH

##### **A.1.1 Retaining the Entire, Original HRS Sample**

```
(i) SAS Code
    data tracker;
    set dat.trk2022;
    if ovresult=1 then do; /* Convert AHEAD IDs into Original HRS
IDs */
```

```

        hhid=ovhhid;
        pn=ovpn;
    end;
run;

proc sort data=tracker;
    by hhid pn;
run;

proc sort data=dat.health out=health(keep=hhid pn v1 );
    by hhid pn;
run;

data healtht;
    merge tracker health(in=inh);
    by hhid pn;
    if inh;
run;

```

(ii) Stata Code

```

use trk2022
replace HHID=OVHHID if OVRESULT==1
replace PN=OVPN if OVRESULT==1
sort HHID PN
save tracker, replace

use health
sort HHID PN

merge HHID PN using tracker
keep if _m==3
save healtht, replace

```

(iii) SPSS Code

```

GET FILE "c:/temp/health.sav".
SORT CASES BY hhid pn.
execute.
save outfile='C:/temp/health2.sav'.

GET FILE "C:\temp\trk2022.sav".
* Convert AHEAD IDs into Original HRS IDs.
if (ovresult =1)hhid=ovhhid.
if (ovresult =1)pn=ovpn.
SORT CASES BY hhid pn.
execute.
save outfile='C:/temp/tracker.sav'.

MATCH FILES /FILE="C:/temp/health2.sav"
/IN=h

```



```

    /FILE='C:\temp\tracker.sav'
    /BY hhid pn
select if h=1.
EXECUTE.

```

```
save outfile='C:/temp/healtht.sav'.
```

### **A.1.2 Keeping Non-Overlaps Only**

(i) SAS Code

```

proc sort data=dat.trk2022 out=tracker;
  by hhid pn;
run;

```

```

proc sort data=dat.health out=health(keep=hhid pn v1);
  by hhid pn;
run;

```

```

data healtht;
  merge tracker(in=int) health(in=inh);
  by hhid pn;
  if ovresult=0 and inh;          /* Keep only Non-Overlap
Cases */
run;

```

(ii) Stata Code

```

use trk2022
sort HHID PN
save tracker, replace

```

```

use health
sort HHID PN

```

```

merge HHID PN using tracker
keep if _m==3 & OVRESULT==0    /* Keep Only Non-Overlap
Cases */
save healtht, replace

```

(iii) SPSS Code

```

GET FILE "c:/temp/health.sav".
SORT CASES BY hhid pn.
execute.
save outfile='C:/temp/health3.sav'.

```

```

GET FILE "C:\temp\trk2022.sav".
SORT CASES BY hhid pn.
execute.
save outfile='C:/temp/tracker2.sav'.

```

```

MATCH FILES /FILE="C:/temp/health3.sav"
  /IN=h
  /FILE='C:\temp\tracker2.sav'
  /BY hhid pn.
select if h=1 and ovresult=0.
EXECUTE.

save outfile='c:/temp/healtht.sav'.

```

## **A.2 HRS 1998 Merges**

In this example, we merge the tracker file with dataset h98a\_r. In this case, there is one respondent identified in the 1998 data by their OVHHID and OVPN, so we revert the identifiers in tracker to match the released data before merging.

### (i) SAS Code

```

data tracker;
  set dat.trk2022;
  if ovresult=8 and hhid eq "205864" then hhid=ovhhid;
  if ovresult=8 and hhid eq "205864" then pn=ovpn;
  /* Convert AHEAD Inter-Respondent Overlap ID into Original */
  run;

proc sort data=tracker;
  by hhid pn;
  run;

proc sort data=dat.h98a_r out=h98a_r (keep=hhid pn f968 f970);
  by hhid pn;
  run;

data h98a_rt;
  merge tracker h98a_r(in=inh);
  by hhid pn;
  if inh;
  run;

```

### (ii) Stata Code

```

use trk2022
replace HHID=OVHHID if OVRESULT==8 & HHID~="205864"
replace PN=OVPN if OVRESULT==8 & HHID~="205864"

sort HHID PN
save track, replace

use h98a_r
keep HHID PN F968 F970

```

```

sort HHID PN

merge HHID PN using track
keep if _m==3
save h98a_rt, replace

```

(iii) SPSS Code

```

GET FILE "C:\temp\trk2022.sav".

* Convert AHEAD Inter-Respondent Overlap ID to original.
  if (ovresult =8) and (HHID <> '205864') hhid=ovhhid.
if (ovresult =8) and (HHID <> '205864') pn=ovpn.
execute.
SORT CASES BY hhid pn.
execute.
save outfile='C:/temp/tracker.sav'.

GET FILE "C:\temp\h98a_r.sav"/keep hhid pn f968 f970.
SORT CASES BY hhid.
execute.
save outfile='C:\temp\h98a_r2.sav'.

MATCH FILES /FILE="C:/temp/h98a_r2.sav"
  /IN=h
  /FILE='C:\temp\tracker.sav'
  /BY hhid pn.
select if h=1.
EXECUTE.
save outfile='C:/temp/h98a_rt.sav'.

```

### **A.3 HRS 2010 Merges for 2012 Overlaps**

As described in Section 3C, there are two respondents who appear in 2010 released data under their OVHHID and OVPN. In this example, we convert their identities in the core 2010 data to match both tracker and the 2012 data release.

(i) SAS Code

```

data tracker;
  set dat.trk2022;
  run;

proc sort data=tracker;
  by hhid pn;
  run;

data h10a_r;
  set dat.h10a_r;

```

```
if hhid = '526934' and pn = '010' then do;
    hhid = '525520'; pn = '020';
end;

else if hhid = '529766' and pn = '010' then do;
    hhid = '520845'; pn = '020';
end;

run;

proc sort data=h10a_r (keep=hhid pnma019);
    by hhid pn;
run;

data h10a_rt;
    merge tracker h10a_r(in=inh);
    by hhid pn;
    if inh;
run;
```

**APPENDIX B**

**Table 5. List of Tracker Updates**

<b>Date updated</b>	<b>Primary Changes</b>
<b>8/16/24</b>	<ul style="list-style-type: none"> <li>• Added Wave S specific variables: (SAGE, SALIVE, SCOUPID, SCOUPLE, SFAMR, SFINR, SINSAMP, SIWLANG, SIWMODE, SIWMONTH, SIWTYPE, SIWWAVE, SIWYEAR, SLBCOMP, SLBWGTR, SLIVARR, SMARST, SNEWSP, SNURSHM, SPPN, SPROXY, SRESCODE, SSUBHH, SSUBHHIW, SWGTR, SWGTRNH, SWHY0HWT, SWHY0RWT, SWGTHH, SPMWGTR)</li> <li>• Added mode completed leave behind questionnaire variables for wave K - S: (KLBCOMP, LLBCOMP, MLBCOMP, NLBCOMP, OLBCOMP, PLBCOMP, QLBCOMP, RLBCOMP, SLBCOMP)</li> <li>• Added variables to replace xWEBCONTROL: (SWEB_ELIG, SWEB_INVITE)</li> <li>• Updated cases that were missing a value: (BIRTHMO (n = 15), BIRTHYR (n = 15), DEGREE (n = 29), FIRSTIW (n = 39), GENDER (n = 14), HISPANIC (n = 39), IMMGYEAR = (n = 1), RACE (n = 37), SCHLYRS (n = 39), USBORN (n = 39), WTCOHORT (n = 15))</li> <li>• Updated CAMS21 for 031059/031 and 031059/032.</li> <li>• Updated SFAMR &amp; SFINR: Code 3 (2nd financial respondent, answer not retained) has been dropped and will no longer be used moving forward.</li> </ul>