

Sample Sizes and Response Rates

Documentation Report

HRS Staff

Survey Research Center
Institute for Social Research
University of Michigan
Ann Arbor, Michigan

Spring 2011

Funding

The Health and Retirement Study is funded by a grant from the National Institute on Aging (U01 AG009740) with supplemental support from the Social Security Administration. HRS is conducted by the University of Michigan.

Suggested Citation

HRS Staff. (2011). Sample Sizes and Response Rates. University of Michigan.

<https://hrs.isr.umich.edu/publications/biblio/9043>

Sample Sizes and Response Rates

1. Background

Since 1998, the objective of the HRS has been to provide information about the U.S. population over age 50 through biennial surveys with samples of that population. Prior to 1998, the target populations were more limited: the original HRS target population was limited to those born between 1931 and 1941, and that of the AHEAD study was limited to those born in 1923 or before. For practical reasons, the decision was made to add new cohorts every six years rather than at each wave of data collection. Therefore, in 1998 the target population was defined as those born in 1947 or before, and thus approximately those age 51 and older. Since new cohorts were not added in 2000 or 2002, the target populations were approximately 53 and older in 2000, and 55 and older in 2002. In 2004, a supplementary sample was added to make the total sample representative of those born in 1953 or before, and thus, once again, approximately age 51 and older. In the 2010 wave, the mid-baby boom cohort (born 1954-1959) will be added, and in 2016 the late baby boom cohort (born 1960-1965) is scheduled to be added..

Two of the five samples interviewed to date by HRS, and a majority of a third sample, came from a screening of 69,337 housing units conducted in 1992. That sample of housing units was generated using a multi-stage, clustered area probability frame. Of those housing units, 14% (9,419) were determined to be non-sample (unoccupied, or non-households). In all but 214 of the 59,918 identified households, the eligibility of the household members for inclusion in the HRS, AHEAD, or WB samples was determined, for a screening response rate of 99.6%.

The original HRS sample consists of individuals born between 1931 and 1941, inclusive. This sample came from the household screening described above. At the baseline data collection for the HRS sample in 1992, a total of 15,497 individuals were eligible for interviews. This total included persons identified in the household screening, plus their spouses or partners regardless of year of birth. Of those identified in this way, interviews were obtained with 12,652 respondents (7,704 households), for an overall response rate of 81.6%.

The second sample was generated for what began as a separate study: Asset and Health Dynamics among the Oldest Old (AHEAD). This sample consists of individuals born in 1923 or before. Those born between 1914 and 1923, and about half of those born in 1913 or before, were identified through the 1992 household screening operation. The other half of those born in 1913 or before were identified using the Medicare enrollment files maintained by the Health Care Financing Administration (HCFA, since renamed the Centers for Medicare, Medicaid Services, or CMS). For the AHEAD sample, interviews were obtained with 8,222 respondents (6,046 different households), a response rate of 80.4%.

The HRS and AHEAD studies were merged, with a single interview schedule in 1998. At the same time the third and fourth samples were added. The War Baby (WB) sample consists of those born between 1942 and 1947, inclusive, and was obtained from the same 1992 household screening. The Children of the Depression Age (CODA) sample consists of those born between 1924 and 1930 (the 'missing' birth cohorts between the HRS and AHEAD samples). These individuals were identified from the Medicare enrollment file. Since many members of these birth cohorts were already part of the study because they were current or former spouses and partners of those in the HRS and AHEAD cohorts, the new samples excluded those individuals with spouses or partners who were born in 1923 or before, or between 1931 and 1947. The baseline response rates for the CODA and WB samples in 1998 were 72.5% and 70%, respectively.

In 2004 a new sample cohort of individuals born in 1948-53 (age 51-56 in 2004) was introduced, which carries forward the steady state aspect of HRS. The Early Baby Boomer (EBB) sample was obtained through the screening of 38,385 households. Eligibility was determined in 91.3% of the screened households, and a total of 4,420 individuals in 2,755 households were found to be eligible. Interviews were completed with 3,330 individuals in 2,154 household for individual and household interview response rates of 75.3% and 78.2%, respectively. Factoring in the screening response rate yields overall baseline response rates of 68.7% for individuals and 71.4% for households.

The ground rules for following baseline respondents and their spouses in subsequent interview waves were as follows:

- Persons interviewed in the baseline data collection (i.e., in 1992 for the HRS sample, 1993-94 for the AHEAD sample, 1998 for the CODA and WB samples, and 2004 for the EBB sample) will be included in the reinterview sample, regardless of their year of birth and regardless of any change in marital status. The only exceptions are: a) those who had died and for whom exit interviews were obtained in a previous wave; and b) those who have been permanently removed from the sample at their insistence; there were 1,763 such cases (out of a total sample of about 31,000) at the end of the 2008 data collection period.
- Spouses and partners reported by respondents at any previous wave will be included in the reinterview sample, new spouses and partners reported at the time of a reinterview will be added to the sample, and interviews will be sought with them. This includes spouses and partners who declined to be interviewed at the baseline and/or follow-up data collections.
- For panel members who are found to have died, *exit* interviews will be sought with a proxy informant who is knowledgeable about the family and financial situation of the deceased. If the deceased is survived by a widow or widower, the exit interview will be sought with the former spouse.
- If an exit interview has been conducted at a previous wave but the information obtained in that interview was incomplete, generally because the estate had not been settled, a short *post-exit* interview will be sought to fill in the missing information.

2. Steady State Sample: Design Considerations

Implementing a sample of the U.S. population over age 50 is complicated by another objective of the study: to describe not just the characteristics of the individuals in the target population, but also the characteristics of the spending units of which those individuals are part. Specifically, spending units are defined as uncoupled individuals and of couples (whether or not married). To achieve that objective, interviews are sought at each wave not only with the sample of cohort-eligible individuals, but also with their current spouse or partner. Those spouses and partners may be of any age, that is, born in any year.

The first time the complication raised by this feature of the design was faced was at the first wave of the AHEAD study, in 1993. Recall that the eligibility for the original HRS sample in 1992 was that an individual either was born in the cohort range 1931-41, or was coupled with an individual born in that range; and that eligibility for the AHEAD sample was that an individual either was born in 1923 or before, or was coupled with an individual born in that range. Some couples consisted of one individual in the HRS cohort range and another born in the AHEAD cohort range. If all such couples were kept in the HRS sample, the AHEAD sample would not fully represent its target population. Therefore, a random subset of such couples were re-assigned to the AHEAD sample. To account for this random selection process, the weights of all such couples, in both the HRS and AHEAD samples, were multiplied by the inverse of the selection probability for inclusion in whichever sample they happened to be assigned.

After the 1998 merger of the HRS and AHEAD studies, the distinction between those assigned to one or the other of the two original samples became irrelevant. The addition of the CODA sample and the WB sample produced a sample that represented the entire household population born pre-1948. To implement this overall design, however, we again had to deal with the complication introduced by the fact that members of couples are not necessarily in the same birth cohort range. For the WB cohort, this was straightforward because of the fact that the sample of those in the WB cohort range was drawn from the same household screening operation that was used to identify the original

HRS and the AHEAD samples. That is, the screening operation identified households that included members born in 1942-47, as well as those born in 1931-41 and those born in 1923 or before. If a household member born in the WB range (1942-47) was coupled, in 1992, with someone born in 1931-41, or pre-1924, that individual was asked to participate, along with their spouse or partner, starting in 1992 or 1993. The remainder of those born in 1942-47 were not asked to participate in the study until 1998. Until 1998, the only individuals born in that age range were those coupled with older spouses or partners, and therefore could not be taken as representative of the entire population born in that age range, and so were given individual (respondent-level) weights of 0. However, starting in 1998, the combined sample (i.e., those coupled in 1992 with someone born in 1931-41, PLUS those coupled in 1992 with someone born pre-1924, PLUS those who were uncoupled in 1992, PLUS those who were coupled in 1992 with someone born in 1924-40 or after 1941) was representative of the entire population of those born in the target range of 1942-47 (less only those who migrated to the U.S. between 1992 and 1998).

The CODA sample was drawn from a separate sampling frame -- the Medicare enrollment data base -- but the principle we followed in selecting additional members of that cohort range was the same as that we used for the WB sample. Again, some individuals born in the CODA range (1924-30) were coupled with spouses or partners born in 1931-41, and therefore interviewed starting in 1992; others were coupled with spouses or partners born pre-1924, and therefore interviewed starting in 1993. In addition, some were coupled with individuals born in 1942-47, and therefore selected along with their WB spouse or partner starting in 1998. Therefore, to represent the complete population of those born in 1924-30, we selected from the Medicare frame anyone born in that range of years who were NOT coupled with someone born pre-1924, or 1931-41, or 1942-47.

The EBB sample is intended to represent those born in the years 1948-53, and was added to the existing HRS sample in 2004. But just as for the CODA and WB samples, part of the target EBB population was already part of the existing HRS sample: those coupled with spouses or partners born prior to 1948. To find this sample, we relied on a second household screening that was conducted in the winter and spring of 2004. Ideally, to avoid overlap with the existing HRS sample we would have wanted to exclude any individual who, though born in 1948-54, was coupled with someone born prior to 1948 at the time of the initial screening in 1992. This was impractical to implement, so instead we excluded those who were coupled with someone born prior to 1948 at the time of the 2004 screening. This may introduce a small distortion in the EBB sample. Specifically, those who were coupled with someone born prior to 1948 in 1992, but not in 2004, are overrepresented (they could have been selected either in 1992 or in 2004). (Those who were coupled with someone born prior to 1948 in 2004, but not in 1992, are represented by the spouses and partners of respondents in the four older samples, who are identified and asked to participate at each wave.)

For data analysts, the year in which an individual was first interviewed can generally be ignored. Respondents in the WB birth range have zero weights prior to 1998, because they represent only a part of their birth cohorts. Respondents in the WB birth range have non-zero weights starting in 1998, because they are now representative of their birth cohorts.

3. Response rates and numbers of interviews with living respondents

The sample sizes and response rates presented in this section pertain to living respondents (or those not known to have been deceased at the time of the designated wave). Sample sizes and response rates for exit interviews for deceased respondents are presented in Section 4. Baseline (Wave 1) response rates reflect the percent of all individuals who were determined to be eligible for HRS who completed a baseline interview. Response rates for follow-up waves (waves 2+) are based on the sample for which interviews were attempted in the designated wave. The base for these rates excludes individuals who requested to be permanently removed from the sample; those counts are shown in Table 6.

Information about the number of interviews and the number eligible at each wave is provided in Table 1.

Table 1: Overall interview response rates for each sample at each wave¹

Sample	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9
HRS									
# of cases	15,497	12,777	12,622	12,202	11,762	11,230	10,835	10,026	9,587
# Respondents	12,652	11,420	10,964	10,584	10,044	9,724	9,362	8,879	8,493
Response rate	81.6%	89.4%	86.9%	86.7%	85.4%	86.6%	86.4%	88.6%	88.6%
Year	1992	1994	1996	1998	2000	2002	2004	2006	2008
AHEAD									
# of cases	10,229	7,554	6,512	5,526	4,559	3,766	2,979	2,362	
# Respondents	8,222	7,027	5,951	5,000	4,107	3,365	2,700	2,142	
Response rate	80.4%	93.0%	91.4%	90.5%	90.1%	89.4%	90.6%	90.7%	
Year	1993	1995	1998	2000	2002	2004	2006	2008	
CODA									
# of cases	3,200	2,300	2,140	1,973	1,770	1,608			
# Respondents	2,320	2,124	1,951	1,777	1,618	1,454			
Response rate	72.5%	92.3%	91.2%	90.1%	91.4%	90.4%			
Year	1998	2000	2002	2004	2006	2008			
WB									
# of cases	3,619	2,652	2,630	2,612	2,539	2,488			
# Respondents	2,529	2,410	2,384	2,295	2,237	2,165			
Response rate	69.9%	90.9%	90.6%	87.9%	88.1%	87.0%			
Year	1998	2000	2002	2004	2006	2008			
EBB									
# of cases	4,420	3,461	3,433						
# Respondents	3,330	3,035	2,963						
Response rate	75.3%	87.7%	86.3%						
Year	2004	2006	2008						

The experience in obtaining the cooperation of sample members in follow-up waves is summarized in Table 2. The response rates range from a low of 85% to a high of 93%. Across samples, there was a drop in response rates of about 2 percentage points between 1994 and 2004 (from over 89.4% to under 87.6%), followed by a slight upward tick in 2006 and 2008.

Table 2: Overall response rate for each sample at each follow-up wave

Sample	Year(s) of Data Collection							
	1994	1995/96	1998	2000	2002	2004	2006	2008
HRS	89.4%		86.7%	85.4%	86.6%	86.4%	88.6%	88.6%
AHEAD		93.0%	91.4%	90.5%	90.1%	89.4%	90.6%	90.7%
CODA				92.3%	91.2%	90.1%	91.4%	90.4%
WB				90.9%	90.6%	87.9%	88.1%	87.0%
EBB							87.7%	86.3%
TOTAL (by year)	89.4%	89.2%	88.3%	88.0%	88.4%	87.6%	88.9%	88.4%

The overall response rate at any follow-up wave is a mixture of the response of three major types of individuals: those who participated in the prior wave (referred to as re-interview cases), those who were eligible to participate in

¹ Note: The denominator used in calculating response rates for the first wave includes sample members who were identified as eligible in the household screener or sample frame. The denominator used in calculating the response rates for the second and later waves includes only those who were themselves respondents at Wave 1, or whose spouse or partner was a respondent at Wave 1. That is, households in which no sample member was interviewed at the baseline are dropped from the sample in subsequent waves. In addition, individuals who have died or who have requested to be permanently removed from the sample are excluded from the denominators for the follow-up waves.

the prior wave but did not (referred to as re-contact cases), and new spouses who become eligible for the first time. A complete overview of response rates thus involves the baseline response rates, plus three wave-specific response rates for each follow-up wave. Details on these rates are given in Tables 3, 4 and 5, and described briefly here.

The reinterview and recontact response rates appear to show more stability than the baseline response rates discussed earlier. For the HRS sample, reinterview response rates track in the low to mid-90% range, with a slightly upward trend from the first reinterview (92.0% in 1994) to the eighth follow-up wave (95.9% in 2008). For the AHEAD sample, the reinterview response rates are slightly higher than for those for the HRS sample, and like the HRS sample show a slight upward trend from 93.8% in 1995 to 96.0% in 2008. The CODA sample reinterview rates track closely those for the AHEAD sample, while those for the WB sample are somewhat lower than those of the AHEAD or CODA samples, and somewhat higher than those for the HRS sample. The EBB sample had the lowest Wave 2 reinterview response rate of all of the cohorts, but their Wave 3 rate was slightly higher than that of the HRS sample.

Table 3: Reinterview response rates for each sample

Sample	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9
HRS								
# of cases	12,295	11,171	10,696	10,234	9,645	9,431	8,971	8,484
# Respondents	11,317	10,402	10,012	9,612	9,129	8,993	8,564	8,139
Response rate	92.0%	93.1%	93.6%	93.9%	94.7%	95.4%	95.5%	95.9%
Year	1994	1996	1998	2000	2002	2004	2006	2008
AHEAD								
# of cases	7,409	6,063	5,065	4,160	3,403	2,719	2,140	
# Respondents	6,952	5,749	4,842	3,970	3,260	2,607	2,054	
Response rate	93.8%	94.8%	95.6%	95.4%	95.8%	95.9%	96.0%	
Year	1995	1998	2000	2002	2004	2006	2008	
CODA								
# of cases	2,220	1,977	1,804	1,633	1,467			
# Respondents	2,082	1,880	1,734	1,564	1,404			
Response rate	93.8%	95.1%	96.1%	95.8%	95.7%			
Year	2000	2002	2004	2006	2008			
WB								
# of cases	2,503	2,372	2,356	2,246	2,194			
# Respondents	2,323	2,246	2,217	2,140	2,095			
Response rate	92.8%	94.7%	94.1%	95.3%	95.5%			
Year	2000	2002	2004	2006	2008			
EBB								
# of cases	3,294	2,993						
# Respondents	2,948	2,801						
Response rate	89.5%	93.6%						
Year	2006	2008						

Recontact response rates are much lower than the reinterview response rates. The recontact rate at the first follow-up wave of the HRS sample was extremely low (8%), but this reflected a decision not to put any effort into trying to recruit the baseline non-respondents (i.e., the non-responding spouses and partners of baseline respondents) out of concern that this could jeopardize the cooperation of their spouses. At the second follow-up, that concern was overlooked and the recontact rate increased sharply, to 40%. There was a rather sharp downward trend across successive waves from 1995/96 to 2000: from 40% down to 25% for the HRS sample, and from 50% down to 31% for the AHEAD sample, followed by an improvement in 2002 for the HRS sample. There was a fairly sharp decline between 2002 and 2004 in the recontact rate for all four samples (Table 4). This reflects in part a policy decision to limit the effort to convert resistant cases because of concerns about losing them permanently. Since 2004, the recontact response rates have rebounded slightly for the HRS, AHEAD and CODA samples.

Table 4: Recontact response rates for each sample

Sample	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9
HRS								
# of cases	372	1,312	1,393	1,432	1,484	1,299	988	1,049
# Respondents	29	519	496	359	507	290	261	311
Response rate	7.8%	39.6%	35.6%	25.1%	34.2%	22.3%	26.4%	29.6%
Year	1994	1996	1998	2000	2002	2004	2006	2008
AHEAD								
# of cases	121	410	441	377	345	246	214	
# Respondents	61	171	145	117	90	80	81	
Response rate	50.4%	41.7%	32.9%	31.0%	26.1%	32.5%	37.9%	
Year	1995	1998	2000	2002	2004	2006	2008	
CODA								
# of cases	66	148	162	126	136			
# Respondents	29	59	39	43	46			
Response rate	43.9%	39.9%	24.1%	34.1%	33.8%			
Year	2000	2002	2004	2006	2008			
WB								
# of cases	119	222	229	263	274			
# Respondents	68	112	60	76	54			
Response rate	57.1%	50.5%	26.2%	28.9%	19.7%			
Year	2000	2002	2004	2006	2008			
EBB								
# of cases	113	391						
# Respondents	39	122						
Response rate	34.5%	31.2%						
Year	2006	2008						

Response rates for added spouses and partners are erratic, probably reflecting the rather small number of cases encountered at each wave, but the overall response rate across samples and waves is 71%, considerably lower than the baseline response rates obtained for the HRS and AHEAD samples, but comparable to the baseline response rates of the CODA and WB samples (Table 5).

As noted previously, respondents who refuse in a given wave are kept in the sample and attempts are made to interview them at each subsequent wave. An exception to this occurs when respondents are insistent that they do not wish to be contacted again. Through 2008, a total of 1,763 living respondents have been removed from the sample. This represents 5.7% of the total HRS sample of 31,022 as of 2008. Table 6 provides a breakdown of the number and percentage of living respondents who have been dropped from the sample at each wave by cohort. The rate of sample removal is more than twice as high for the original HRS cohort than for any other cohort.

Table 5: Response rates for new spouses and partners for each sample

Sample	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9
HRS								
# of cases	110	139	113	96	101	105	67	54
# Respondents	74	43	76	73	88	79	54	43
Response rate	67.3%	30.9%	67.3%	76.0%	87.1%	75.2%	80.6%	79.6%
Year	1994	1996	1998	2000	2002	2004	2006	2008
AHEAD								
# of cases	24	39	20	22	18	14	8	
# Respondents	14	31	13	20	15	13	7	
Response rate	58.3%	79.5%	65.0%	90.9%	83.3%	92.9%	87.5%	
Year	1995	1998	2000	2002	2004	2006	2008	
CODA								
# of cases	14	15	7	11	5			
# Respondents	13	12	4	11	4			
Response rate	92.9%	80.0%	57.1%	100.0%	80.0%			
Year	2000	2002	2004	2006	2008			
WB								
# of cases	30	36	27	30	20			
# Respondents	19	26	18	21	16			
Response rate	63.3%	72.2%	66.7%	70.0%	80.0%			
Year	2000	2002	2004	2006	2008			
EBB								
# of cases	54	49						
# Respondents	48	40						
Response rate	88.9%	81.6%						
Year	2006	2008						

Table 6: Number and percentage in each cohort dropped from the sample at each wave

Sample	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9	Total
HRS									
# dropped before deceased	11	239	162	180	144	396	68	68	1,268
% of eligible sample	0.09%	1.89%	1.33%	1.53%	1.28%	3.65%	0.68%	0.71%	9.26%
Year	1994	1996	1998	2000	2002	2004	2006	2008	
AHEAD									
# dropped before deceased	15	52	57	38	88	19	3		272
% of eligible sample	0.20%	0.80%	1.03%	0.83%	2.34%	0.64%	0.13%		3.18%
Year	1995	1998	2000	2002	2004	2006	2008		
CODA									
# dropped before deceased	14	12	43	6	5				80
% of eligible sample	0.61%	0.56%	2.18%	0.34%	0.31%				3.28%
Year	2000	2002	2004	2006	2008				
WB									
# dropped before deceased	17	13	45	24	16				115
% of eligible sample	0.64%	0.49%	1.72%	0.95%	0.64%				4.12%
Year	2000	2002	2004	2006	2008				
EBB									
# dropped before deceased	11	17							28
% of eligible sample	0.32%	0.50%							0.79%
Year	2006	2008							

Disposition of Cases across Waves. To date, nine waves of data have been collected from the HRS sample; eight waves from the AHEAD sample; six from the CODA and WB samples; and three waves from the EBB sample. In this section we describe the patterns of cooperation of members of the five samples across follow-up waves.

A total of 13,687 individuals are in the HRS sample since the baseline interviews in 1992. Over two-thirds (67.1%) of the respondents in this sample have complete interview histories from their initial entry through 2008. The remaining 32.9% have missed at least one interview: an average of 2.7 missed interviews (7.3 average attempts). The total AHEAD sample is 8,528 individuals. More than four out of five (81.8%) have been interviewed (self or proxy) at every wave they were eligible. The remaining (18.2%) missed an average of 2.0 interviews (5.5 average attempts). Complete interview histories have been provided by 80.8% of the 2,442 individuals in the CODA sample, and the remaining 19.2% have missed an average of 2.0 interviews (5.0 average attempts). For the 2,793 individuals in the WB sample, all requested interviews have been obtained from 74.0%, and the remaining cases missed an average of 2.1 interviews (5.4 average attempts). The EBB sample has a shorter history in the study and thus fewer opportunities to have refused to be interviewed. Complete interview histories have been provided by 81.3% of the 3,549 individuals in the EBB sample, and the remaining 18.7% have missed an average of 1.5 interviews (2.9 average attempts).

Another way of looking at panel retention is to calculate the percentage of survivors from the original sample who gave an interview in a specific followup wave. For example, of the original 12,652 persons interviewed in the 1992 HRS study, 2,909 died prior to 2008, 929 were removed from sample prior to 2008, and 7,843 gave an interview in 2008. The response rate among all survivors, including those removed from the sample, was 80.5%. Panel retention by duration in survey is generally similar for the other cohorts. The HRS cohort rate of 80.5% retention at 16 years of survey duration is slightly better than the National Longitudinal Surveys (NLS)-Older Men (76.3%) and Mature Women studies (73.1%), but somewhat below the record levels of the National Longitudinal Survey of Youth 1979 (NLSY79) cohort, which stood at 89% among survivors after 16 years and did not reach an 80.5% response rate of survivors until 2002—23 years after baseline.

4. Mortality ascertainment and exit interviews

The HRS typically learns of the death of a respondent when an interviewer attempts to reach the respondent for an interview during the main data collection period. The respondent's spouse or another close family member or friend is asked to provide a final interview on behalf of the respondent (the Exit interview). In some instances, HRS staff are notified of a respondent's death by a family member between waves. Such notification is not considered definitive. Interviewers are alerted to this in the Respondent Profile and the main data collection period is used to confirm the deceased status of the respondent and pursue an Exit interview with a spouse or another close family member or friend. There are also situations in which the respondent is in tracking because we were unable to locate him or her during a previous wave of data collection. Tracking resources for confirming the deceased status of respondents and locating an Exit proxy reporter include the Social Security Death Index (SSDI), which is accessed through the SSDI website and also through a paid subscription to Insight databases (available only to members of a national tracking team which conducts advanced tracking of respondents).

HRS has conducted linkages to the National Death Index following each wave since the 2000 wave. After every linkage, the public-release Tracker File is updated with the match status, month of death, and year of death for verified matches with the NDI database. Data on exact date of death and cause of death may be obtained as an HRS Restricted Data product (<http://hrsonline.isr.umich.edu/index.php?p=resdat>).

Exit Interviews. HRS has conducted Exit interviews for deceased respondents in every follow-up wave starting with the 1994 wave. The Exit interview was developed to find out about the status and activities of the respondent from the time of the last interview until his or her death and to find out about the circumstances of the death. HRS attempts to complete an Exit interview for all deceased respondents, with two exceptions. New spouses for whom a baseline interview was never obtained and respondents who had requested to be removed from the sample prior to their death are excluded from the Exit interview.

The Exit interview is administered to someone knowledgeable about the deceased respondent, preferably a surviving spouse or close family member, and is typically obtained during the field period when the death is first reported. Thus, for most respondents it occurs less than two years following the respondent's death. If it is not possible to obtain an Exit interview on the first attempt, the family member or proxy is recontacted during the next data collection period. If an Exit interview has been conducted at a previous wave but the information obtained in that

interview was incomplete, generally because the estate had not been settled, a short "post-exit" interview is sought to fill in the missing information.

Table 7 provides counts and response rates for the number of exit interviews in each year of data collection, by cohort. In 1994, exit interviews were attempted for the first of the decedents from the original HRS cohort. The sample size was small and the response rate somewhat lower than in subsequent waves. Since 1998, the number of exit cases has fluctuated around 1,500 and the response rate has ranged between 84% and 92%. As of 2008, exit interviews have been obtained for over 9,000 deceased respondents.

Table 7: Number of attempted and completed exit interviews at each wave, by birth cohort

Birth Cohort	Year(s) of Data Collection								
	1992	1993/94	1995/96	1998	2000	2002	2004	2006	2008
1890-1923									
# of cases	--	13	840	1104	1019	1012	758	703	646
# respondents	--	13	767	962	882	875	714	627	601
Response rate	--	100.0%	91.3%	87.1%	86.6%	86.5%	94.2	89.2%	93.0%
1924-30									
# of cases	--	29	55	90	208	276	270	289	301
# respondents	--	20	48	79	184	249	251	254	280
Response rate	--	69.0%	87.3%	87.8%	88.5%	90.2%	93.0%	87.9%	93.0%
1931-41									
# of cases	--	170	219	241	307	397	273	368	370
# respondents	--	129	178	200	247	317	225	318	334
Response rate	--	75.9%	81.3%	83.0%	80.5%	79.8%	82.4%	86.4%	90.3%
1942-47									
# of cases	--	16	13	9	32	61	43	83	67
# respondents	--	11	12	9	28	51	32	72	60
Response rate	--	68.8%	92.3%	100.0%	87.5%	83.6%	74.4%	86.7%	89.6%
1948-53									
# of cases	--	4	2	3	7	6	6	41	58
# respondents	--	3	2	2	6	6	3	34	48
Response rate	--	75.0%	100.0%	66.7%	85.7%	100.0%	50.0%	82.9%	82.8%
1954+									
# of cases	--	0	1	0	1	1	2	5	9
# respondents	--	0	1	0	1	1	1	4	7
Response rate	--	--	100.0%	--	100.0%	100.0%	50.0%	80.0%	77.8%
Unknown									
# of cases	--	0	1	3	1	1	1	1	0
# respondents	--	0	1	2	0	0	1	1	0
Response rate	--	--	100.0%	66.7%	0.0%	0.0%	100.0%	100.0%	--
TOTAL									
# of cases	--	232	1134	1497	1575	1754	1353	1490	1451
# respondents	--	176	1009	1254	1348	1499	1227	1310	1330
Response rate	--	75.9%	89.0%	83.8%	85.6%	85.5%	90.7%	87.9%	91.7%

5. Number of cases by birth cohort

Each of the five samples is defined in terms of a specific range of birth cohorts, but because interviews are sought with the spouses and partners of selected individuals, the actual range of birth cohorts observed in each sample is considerably broader. If those birth years fall within the range of another sample, then these spouses and partners form part of the sample of their own cohort and are given an appropriate sampling weight. If, however, a spouse or partner's year of birth lies outside the range of any sample, then he or she is given a zero sampling weight until their cohort is fully represented. Thus, for example, the spouses and partners of HRS sample members born in 1924-30 or in 1942-47 are assigned zero weights in 1992 through 1996, but are given non-zero weights starting in 1998 when their birth cohorts became fully represented with the addition of the CODA and WB samples.

Those in the original HRS birth cohorts (1931-41) constituted 78% of the respondents in 1992, while the remaining 22% were the spouses and partners of those individuals, born in earlier or later years. The number in the HRS birth cohorts dropped by about 900 for the 1993/94 data collection, and as a proportion of all respondents they dropped to just 45% because of the introduction of the AHEAD sample (those born prior to 1924 along with their spouses and partners). In 1998, the CODA (1924-30) and WB (1942-47) samples were added, though at a lower sampling rate than for the HRS sample, and the overall sample, when properly weighted, was representative of the U.S. population born prior to 1948. Table 8 provides details about the distribution of birth years for respondents at each data collection.

Table 8: Counts and proportions of core interviews at each wave, by birth cohort

Birth Cohort	Year(s) of Data Collection								
	1992	1993/94	1995/96	1998	2000	2002	2004	2006	2008
1890-1923	213	7,573	6,415	5,357	4,454	3,561	2,862	2,235	1,715
1924-30	1,020	1,611	1,514	3,752	3,435	3,165	2,886	2,628	2,347
1931-41	9,817	8,917	8,537	8,240	7,777	7,530	7,229	6,859	6,545
1942-47	1,195	1,144	1,123	3,101	2,948	2,910	2,816	2,715	2,654
1948-53	295	285	280	675	677	682	3,370	3,108	3,022
1954+	111	111	122	259	286	317	966	924	934
Unknown	1	1	0	0	1	1	0	0	0
TOTAL	12,652	19,642	17,991	21,384	19,578	18,166	20,129	18,469	17,217
Unweighted proportions									
1890-1923	1.7%	38.6%	35.7%	25.0%	22.7%	19.6%	14.2%	12.1%	10.0%
1924-30	8.0%	8.2%	8.4%	17.6%	17.5%	17.4%	14.4%	14.2%	13.6%
1931-41	77.6%	45.4%	47.4%	38.5%	39.7%	41.5%	35.9%	37.1%	38.0%
1942-47	9.5%	5.8%	6.2%	14.5%	15.1%	16.0%	14.0%	14.7%	15.4%
1948-53	2.3%	1.5%	1.6%	3.2%	3.5%	3.7%	16.7%	16.8%	17.6%
1954+	0.9%	0.6%	0.7%	1.2%	1.5%	1.8%	4.8%	5.0%	5.4%
TOTAL	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Weighted proportions									
1890-1923	0.0%	48.3%	45.2%	22.4%	19.7%	17.2%	10.5%	8.5%	6.8%
1924-30	0.2%	0.2%	0.2%	18.3%	18.4%	18.2%	13.2%	11.9%	11.1%
1931-41	99.6%	51.4%	54.4%	32.7%	34.5%	35.6%	26.7%	26.0%	26.5%
1942-47	0.2%	0.1%	0.2%	26.4%	27.1%	28.6%	21.8%	22.6%	23.6%
1948-53	0.0%	0.0%	0.0%	0.1%	0.2%	0.2%	27.8%	30.9%	32.0%
1954+	0.0%	0.0%	0.0%	0.1%	0.1%	0.2%	0.0%	0.0%	0.0%
TOTAL	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

6. Self and proxy interviews

Although our goal has been to conduct interviews with the sampled individuals themselves whenever possible, we do permit a proxy interview to be conducted when an individual is unable to do so because of physical or cognitive limitations, and also occasionally when the individual is unwilling to take the time to be interviewed but consents to having someone else (almost always their spouse) be interviewed as their proxy. A substantial number of proxy interviews are in fact done when there is clearly no alternative, following the death of the sampled individual. Most proxy interviews are done because the interviewer finds that the sampled person clearly cannot complete the interview, or is given that message by a gatekeeper. A small number, starting with AHEAD 1995, have been triggered by the low score that a respondent obtains on a test of their cognitive abilities. When that occurs, the interviewer is encouraged (via an interviewer prompt) to either terminate the interview and start again with a proxy informant, or to continue the interview with the help of a care provider, if possible.

The proportion of interviews done by proxy informants is shown in Table 9 for each birth cohort and for each wave of data collection. The proportion of proxy interviews is higher among older respondents. For example, in the 2002 wave less than 9% of those born in 1931 or later had proxy interviews, compared to 19% of those born before 1924. The proportion of proxy interviews increased across the first decade of the study, from 5% in 1992 to 11% in 2002, before falling back to 9% in 2004. This rising pattern generally holds within birth cohorts as well, and likely reflects in part the increasing age of the sample members. The decline in proxy interviews starting in 2004 likely reflects the higher proportion of interviews done face-to-face rather than by telephone. Additionally, with the introduction of the enhanced face-to-face interview in 2006, interviewers were instructed to obtain a self-interview whenever possible in order to maximize the sample for the collection of physical measures and biomarkers.

Table 9: Proportion of core interviews done by proxy informants

Birth Cohort	Year(s) of Data Collection								
	1992	1993/94	1995/96	1998	2000	2002	2004	2006	2008
1890-1923	9.4%	10.6%	13.7%	16.0%	17.6%	19.1%	18.3%	17.2%	19.2%
1924-30	8.3%	8.9%	9.8%	8.1%	9.8%	11.5%	9.3%	7.7%	8.1%
1931-41	4.8%	6.1%	6.1%	7.7%	8.6%	9.0%	7.7%	5.5%	5.4%
1942-47	4.4%	5.2%	5.8%	6.0%	7.3%	8.3%	6.7%	4.6%	4.0%
1948-53	4.8%	5.3%	5.7%	7.0%	5.6%	7.0%	5.9%	4.2%	4.2%
1954+	2.7%	3.6%	4.9%	5.8%	8.4%	8.2%	6.9%	4.8%	3.8%
Total+	5.1%	8.0%	9.1%	9.6%	10.5%	11.2%	9.0%	6.8%	6.6%

7. Interviews with nursing home residents

The sample frame from which each cohort is initially drawn excludes the institutionalized population. However, respondents who move to nursing homes after the baseline wave are retained in the study and interviewed whenever possible. A majority of interviews with respondents in nursing homes are conducted with a proxy respondent, but in some cases interviewers are able to conduct a self-interview with the respondent in the nursing home. Information about the nursing home stay (including any stays that occurred since the last interview) is collected in the interview.

Table 10 presents the number and proportion of core interviews at each wave that are conducted with respondents who reside in a nursing home, by birth cohort. As expected, the vast majority of nursing home residents are in the AHEAD birth cohort (born before 1924). For the sample as a whole, the percentage of respondents who reside in nursing homes has ranged between 2.0% and 2.6%. This percentage is much higher for the AHEAD birth cohort, which reached about 14% in 2008.

By the 2000 wave (7 years after the baseline wave for the AHEAD cohort), the HRS sample is fairly well representative of the nursing home population. We have produced sample weights for the nursing home population for the 2000 and 2002 waves and plan to add them for other waves as data to do so become available.

Table 10: Number and proportion of core interviews conducted with nursing home residents*

Birth Cohort	Year(s) of Data Collection								
	1992	1993/94	1995/96	1998	2000	2002	2004	2006	2008
# respondents									
1890-1923	--	1	268	381	378	364	324	274	235
1924-30	--	2	8	12	27	47	60	84	87
1931-41	--	5	19	32	45	46	68	74	107
1942-47	--	0	3	4	6	3	6	4	5
1948-53	--	0	0	0	1	0	1	1	7
1954+	--	0	0	0	0	0	1	1	0
TOTAL	--	8	298	429	457	460	460	438	441
Proportion									
1890-1923	--	0.8%	4.2%	7.1%	8.5%	10.2%	11.3%	12.3%	13.7%
1924-30	--	0.2%	0.5%	0.3%	0.8%	1.5%	2.1%	3.2%	3.7%
1931-41	--	0.1%	0.2%	0.4%	0.6%	0.6%	0.9%	1.1%	1.6%
1942-47	--	0.0%	0.3%	0.1%	0.2%	0.1%	0.2%	0.2%	0.2%
1948-53	--	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.2%
1954+	--	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%
TOTAL	--	0.1%	1.7%	2.0%	2.3%	2.5%	2.3%	2.4%	2.6%

*Interviews may have been completed by the respondent or a proxy

8. Minority sample

An important goal of the HRS is to support research on racial and ethnic disparities. To achieve this, HRS has oversampled Black and Hispanic populations in most of the recruitment cohorts. In the original HRS, AHEAD and EBB cohorts, Blacks and Hispanics are oversampled at a rate of about 2 to 1 relative to their distribution in their respective age groups in the population. Oversampling of minority groups was not done for the War Baby and CODA cohorts for cost reasons. The number of interviews with core (living) respondents in each wave is shown by race/ethnicity and birth cohort in Table 11.

HRS has been successful at maintaining high baseline and followup response rates for Black and Hispanic sample members.² However, because the baby boom cohorts were sampled at a lower rate than the original HRS and AHEAD cohorts (about 60 percent of the original rate), the sample sizes for the minority groups in the baby boom cohorts are considerably smaller than they were when HRS began. To address this issue, HRS will supplement the minority sample in the EBB and MBB cohorts in the 2010 wave, boosting the combined number of Hispanic and African American respondents in these cohorts by over 2,000.

² M.B. Ofstedal and D.R. Weir. The recruitment and retention of minority participants in the Health and Retirement Study. *The Gerontologist*, 51: S8-S20, 2011.

Table 11: Number of core interviews at each wave by race/ethnicity³

Birth Cohort	Year(s) of Data Collection								
	1992	1993/94	1995/96	1998	2000	2002	2004	2006	2008
1890-1923									
Hispanic	29	433	357	309	277	215	177	136	113
Black	40	1041	850	707	567	440	346	271	205
White	137	6007	5134	4281	3554	2868	2312	1809	1381
Other/unknown	7	92	74	60	56	38	27	19	16
1924-30									
Hispanic	79	119	114	246	229	209	203	178	154
Black	122	183	166	377	334	304	273	246	222
White	806	1285	1213	3064	2818	2602	2367	2171	1936
Other/unknown	13	24	21	65	54	50	43	33	35
1931-41									
Hispanic	912	773	758	735	691	694	682	621	602
Black	1688	1504	1401	1334	1248	1197	1153	1047	1013
White	7048	6486	6245	6045	5716	5531	5291	5098	4839
Other/unknown	169	154	133	126	122	108	103	93	91
1942-47									
Hispanic	105	96	98	245	236	227	227	221	215
Black	164	150	146	462	426	429	420	397	393
White	899	874	854	2336	2232	2197	2117	2047	1994
Other/unknown	27	24	25	58	54	57	52	50	52
1948-53									
Hispanic	39	41	41	82	79	79	476	406	405
Black	49	45	43	93	88	85	563	512	494
White	196	188	185	481	492	503	2232	2099	2045
Other/unknown	11	11	11	19	18	15	99	91	78
1954+									
Hispanic	11	10	14	35	38	46	162	147	164
Black	17	12	13	32	43	49	124	127	126
White	78	84	90	185	192	207	633	611	605
Other/unknown	5	5	5	7	13	15	47	39	39
TOTAL									
Hispanic	1175	1472	1382	1652	1550	1470	1927	1709	1653
Black	2081	2936	2619	3005	2706	2504	2879	2600	2453
White	9164	14924	13721	16392	15005	13909	14952	13835	12800
Other/unknown	232	310	269	335	317	283	371	325	311

³ Black, White and Other categories exclude Hispanics.