NATIONAL HEALTH AND AGING TRENDS STUDY (NHATS)
Development of Round 5 Survey Weights

December 1, 2016
Revised December 12, 2017

1. Introduction

The NHATS public use data originally supported weighted analysis of Medicare beneficiaries ages 65 and older living in the contiguous United States on September 30, 2010. The original cohort has been interviewed annually. Replenishment took place in Round 5 so that the sample could be used to study disability trends as well as individual trajectories. The replenishment sample was drawn as of September 30, 2014. For Round 5, separate sets of weights are provided for analyses pertaining to the original target population (the "2011 Cohort") and for analyses pertaining to the new target population (the "2015 Cohort"). The survey weights included with the Round 5 public use file account for differential probabilities of selection and adjust for potential bias related to unit nonresponse to the Round 1 through 5 interviews.

For Round 5 of NHATS, as for Rounds 1, 2, 3, and 4, for each of the two cohorts, two types of sampling weights have been produced: a tracker weight (on the Tracker file with the variable names w5trfinwgt0 and w5tr2011wgt0) and an analytic weight (on the Sample Person file with the variable names w5anfinwgt0 and w5an2011wgt0). For variance estimation (see Section 7), NHATS has also included replicate versions of these weights (w5trfinwgt1-w5trfinwgt56 and w5anfinwgt1-w5anfinwgt56 for the 2015 Cohort; w5tr2011wgt1- w5tr2011wgt56 and w5an2011wgt1- w5an2011wgt56 for the 2011 Cohort).

The methodology that was used to develop these weights and appropriate uses of each of these weights are discussed in the following sections. The next section provides an overview of how cases were classified for purposes of weight development. Sections 3 and 4 detail the creation of the tracker and analytic weights, respectively. Section 5 reports on the effect of weighting adjustments on the precision of NHATS survey estimates. Section 6 provides guidance on the use of the tracker and analytic weights. A final section provides information on the proper calculation of variance estimates to account for the complex design and estimation procedures used in NHATS.

2. Definition of Respondent

In the development of survey weights, an important first step is the classification of cases into groups based on eligibility and response status. For Round 5 of NHATS, Table 1 shows how the disposition codes map into respondent, ineligible, and nonrespondent statuses.

In the computation of the 2011 Cohort weights, only cases in the original sample were included. Both original sample and replenishment sample cases were included in the computation of the 2015 Cohort weights.

2015 Cohort Weights

For the 2015 Cohort Round 5 Tracker weight, only cases that were eligible as of September 30, 2014, and were classified in Round 5 as Respondents (including original sample cases for whom a Round 5 Last Month of Life (LML) interview was completed) or Ineligible (n = 8,799) are assigned a positive weight. Original sample cases for which at least one survey component is available (codes 60, 61, 62, 63 and 64) are considered respondents for purposes of the tracker weight.

Replenishment sample cases who became ineligible for the Round 5 interviews after they were selected, either due to death or due to moving outside the contiguous U.S., also have positive Round 5 tracker weights

For the analytic weight, only Respondents (codes 60, 61, 62, 63; n=8,155) are assigned a positive weight. For the SP interview, cases were required to have completed the self-reported disability protocol (through the section on Participation; PA) to be considered complete.

2011 Cohort Weights

For the 2011 Cohort Round 5 Tracker weight, only original sample cases classified as Respondents and Ineligible (n = 6,402) are assigned a positive weight. Original sample cases for which at least one survey component is available (codes 60, 61, 62, 63 and 64) are considered respondents for purposes of the tracker weight. Original sample cases who became ineligible for the Round 1 interview after they were selected, either because they died or moved out of the contiguous U.S. by the time of the fieldwork, have positive Round 5 tracker weights. Those who became ineligible for the Round 2 interview because they moved out of the contiguous U.S. by Round 2 or who completed a Round 2 Last Month of Life (LML) interview because they died between Rounds 1 and 2 also have positive tracker weights in Round 5, and the same is true for those who became ineligible for the Round 3 (or 4) interview because they moved out of the contiguous U.S. by Round 3 (or 4) and those for whom a Round 3 (or 4) LML interview was completed because they died between Rounds 2 and 3 (or between Rounds 3 and 4). Because a Last Month of Life (LML) interview was attempted for each SP who died between Rounds 4 and 5, deceased original sample SPs with a Round 5 LML interview completed by proxy (code 62) were also considered respondents and have a Round 5 tracker weight. Replenishment sample cases do not have positive 2011 Cohort tracker weights.

For the 2011 Cohort analytic weight, only original sample Respondents (codes 60, 61, 62, 63; n=4,026) are assigned a positive weight. For the SP interview, cases were required to have completed the self-reported disability protocol (through the section on Participation; PA) to be considered complete.

Table 1. Classification of Round 5 NHATS Sample for Weight Development Purposes

Disposition code	Original Sample			Replenishment Sample		
	N	Classification for Tracker Weight	Classification for Analytic Weight	N	Classification for Tracker Weight	Classification for Analytic Weight
60 Complete, community	3,314	Respondent	Respondent	3,756	Respondent	Respondent
60-Complete, NH or residential care	285			165		
61 Complete, NH facility	103	Respondent	Respondent	180	Respondent	Respondent
62 Complete, SP deceased, proxy interview	296	Deceased respondent[+]	Respondent[+]	0	N/A	N/A
63 Complete SP, FQ not complete	28	Respondent	Respondent	28	Respondent	Respondent
64 Complete FQ, SP not complete	126	Respondent	Nonrespondent	53	Respondent	Nonrespondent
75 Physically/mentally unable to participate, no proxy	12	Nonrespondent	Nonrespondent	56	Nonrespondent	Nonrespondent
76 Too ill to participate, no proxy	16	Nonrespondent	Nonrespondent	113	Nonrespondent	Nonrespondent
77 Refusal, Sample Person	91	Nonrespondent	Nonrespondent	1966	Nonrespondent	Nonrespondent
78 Language barrier	1	Nonrespondent	Nonrespondent	30	Nonrespondent	Nonrespondent
79 Unable to locate	2	Eligibility unknown[++]	Eligibility unknown[++]	170	Eligibility unknown[++]	Eligibility unknown[++]
80 Unavailable during field period	3	Nonrespondent	Nonrespondent	21	Nonrespondent	Nonrespondent
81-Final deceased, new sample only	0	N/A	N/A	419	Ineligible	Ineligible
82 Outside of Primary Sampling Unit	7	Nonrespondent	Nonrespondent	11	Nonrespondent	Nonrespondent
83 Ineligible (moved out of contiguous US)	3	Ineligible	Ineligible	43	Ineligible	Ineligible
85 Refusal, facility	7	Nonrespondent	Nonrespondent	14	Nonrespondent	Nonrespondent
86 Deceased, no proxy	13	Deceased nonrespondent[+]	Nonrespondent[+]	0	N/A	N/A
87 Refusal, proxy	21	Nonrespondent	Nonrespondent	82	Nonrespondent	Nonrespondent
88 Work stopped	3	Nonrespondent	Nonrespondent	11	Nonrespondent	Nonrespondent
89 Final other/specify*	2	Nonrespondent*	Nonrespondent*	1	Nonrespondent*	Nonrespondent*
Not attempted in Round 5						
Deceased in Round 1, 2, 3, or 4	2127	Ineligible	Ineligible	0	N/A	N/A
Other Round 1, 2, 3, or 4 ineligible	120	Ineligible	Ineligible	0	N/A	N/A
Round 1, 2, 3, or 4 nonrespondent	5,831	Nonrespondent**	Nonrespondent**	0	N/A	N/A
Total and Number Assigned Weight	12,411	6,580	4,026	7,119	7,119	4,129

[+] For the original sample, the weights of deceased SPs were adjusted separately from those of living SPs.

[++] Due to the very low proportion of fielded cases in this category in Round 2 (0.46% of fielded cases), as well as the low proportion of Round 1 respondents that were ineligible for Round 2 (0.38%), in the original sample, these cases were treated as living nonrespondents in the computation of Round 2 weights. The same approach was used in the computation of Round 3, Round 4, and Round 5 weights. For the replenishment sample, these cases were treated as cases with unknown eligibility.

**These cases were previously adjusted for in the Round 1, Round 2, Round 3, or Round 4 nonresponse adjustment to the tracker weight; for original sample cases, the Round 4 nonresponse adjusted tracker weight was used as input to the Round 5 weighting process, so these cases are not included in the Round 5 nonresponse adjustment.

SP=Sample Person interview; FQ=Facility Questionnaire

3. Computation of Tracker Weights

2015 Cohort Weights

In computing the 2015 Cohort tracker weights, it was important to recognize that the target populations represented by the two samples (the original sample and the replenishment sample) contain substantial overlap (specifically, persons who were 65 or older, residing in the U.S., and on Medicare as of September 30, 2010 and who were still residing in the U.S. and on Medicare as of September 30, 2014). Compositing (a weighting approach that essentially averages the weights of two or more samples that represent the same population) was used to account for this.

The first step in the computation of the 2015 Cohort Round 5 tracker weight was to composite the weights of the two samples. The weights used in the compositing step are the Round 4 nonresponse adjusted tracker weight (prior to raking) for the original sample, and the base weight (which accounts for the probability of selection) for the replenishment sample. This Round 4 weight accounted for differential probabilities of selection and included adjustments for nonresponse to the Round 1, Round 2, Round 3, and Round 4 interviews but was not raked to the HISKEW[1]. See Montaquila et al. (2012) for details on the specific methodology used in computing and adjusting the Round 1 weights; also, refer to Montaquila et al. (2014, 2015a, 2015b) for information about the specific adjustments applied in Rounds 2 through 4, respectively.

In the compositing step, beneficiaries eligible for the sample in which they were selected but not eligible for the other sample (i.e., original sample cases who died or moved out of the U.S. by September 30, 2014, and replenishment sample cases who were not enrolled in Medicare as of September 30, 2010) retained their weights. For beneficiaries eligible for both samples, the weights were adjusted by the factor

$$\gamma = \frac{neff_S}{neff_{orig} + neff_{replen}}$$

where $neff_{orig}$ and $neff_{replen}$ are the effective sample sizes (accounting for unequal weighting design effects) for the original sample and the replenishment sample, respectively, based on the weights that were used as input to the compositing process, and $neff_S$ is the effective sample size for the sample into which the beneficiary was selected (either $neff_{orig}$ or $neff_{replen}$).

To produce the 2015 Cohort Round 5 tracker weight, two additional adjustments were made to the composited weight—an adjustment for Round 5 nonresponse and a raking adjustment to estimated population totals from the Medicare EDB.

The composited weights were adjusted for Round 5 nonresponse. Because response rates differed considerably between the two samples, and it was believed that response mechanisms were different for the two samples (since members of the original sample had been engaged in the study for several rounds, whereas Round 5 was the first contact and attempt at gaining cooperation with the replenishment sample), the two samples were adjusted separately for Round 5 nonresponse.

---

[1] The HISKEW file was a 20% sample of the Medicare enrollment database (as of Sept. 30, 2010) that served as the sampling frame for the original selection.

Potential variables for creating nonresponse cells for the 2015 Cohort Round 5 tracker weights came five sources:

- Beneficiary information from the sampling frame (the 20% HISKEW File for the original sample; the 20% extract of the Medicare Enrollment Database for the replenishment sample), including demographic characteristics of the beneficiary (e.g., age as of September 30, 2014, gender) and geographic information (e.g., census division, metro and micropolitan status) based on the beneficiary's address in CMS' Medicare Enrollment Database (EDB);
- County-level demographic information based on the 5% HISKEW file or the 5% extract of the Enrollment Database (e.g., percent of beneficiaries in the county who are Black; percent of beneficiaries in the county who are Hispanic) for the county linked to the beneficiary's address from the EDB;
- Census tract-level information based on the 2010-2014 5-year American Community Survey (e.g. tract-level demographic information), based on linkages to the beneficiary's address from the EDB;
- For the original sample, variables from the NHATS Rounds 1, 2, 3, and 4 interviews (race/ethnicity, highest education, and Rounds 1, 2, 3, and 4 residential settings); and
- For the replenishment sample, an indicator that the beneficiary's address from the EDB matches an address on a list of licensed assisted living facilities[2], and an indicator of whether the beneficiary could be considered a nursing home resident based on a match to records from the Minimum Data Set (MDS), which contains periodic assessments for all Medicare or Medicaid certified nursing homes. The latter indicator was based on an algorithm developed by Kasper, Edwards, and Freedman to identify beneficiaries who had a pattern of records in the MDS from January 1, 2015-December 31, 2015 consistent with a long-term resident rather than short-term skilled nursing stays. (See Appendix A of Montaquila, Freedman, Spillman, and Kasper, 2012 for further details.)

Appendix Table 1 provides weighted response rates (using the composited weights computed in the first step of the calculation of the 2015 Cohort Round 5 tracker weights) by categories of the various indicators. We used these variables as input to a classification tree analysis to determine which of these variables were associated with nonresponse. This approach uses a search algorithm to identify variables associated with response propensities. At each step in the process, chi-square tests were performed to determine the most significant predictor of response, given the set of conditions already specified in the particular "branch." We also set a minimum cell size of 50.[3]

---

[2] The list was compiled by the "Shaping Long Term Care in America Project" at Brown University funded in part by the National Institute on Aging (P01AG027296)."

[3] The classification tree analysis is designed to work with categorical predictor variables. Alternatives to this approach are propensity modeling based on logistic regression and Cartesian product cross-classification. The logistic regression approach uses a parametric model to identify predictors of response. When the pool of potential predictors includes continuous variables and categorizing the continuous variables would result in substantial losses of information, logistic regression modeling would be preferred over classification tree analysis. The Cartesian product cross-classification approach involves specifying a set of adjustment cell variables based on prior experience (generally, (1) prior analyses that identified predictors associated with response propensities; and/or (2) predictors associated with response and/or subject matter expertise that informs the choice of variables).

We fit separate classification trees for the original sample and the replenishment sample. For the original sample, separate trees were fit for all living non-nursing home cases (Figure 1), nursing home residents (Figure 2), and deceased SPs (Figure 3) because underlying nonresponse processes differed for these three groups. Likewise, for the replenishment sample, separate trees were fit for non-nursing home cases (Figure 4) and nursing home residents (Figure 5). For the original sample, unlike non-nursing home cases, nursing home residents include both Round 1 residents who were not required to complete an SP Interview and new Round 2, Round 3, or Round 4 nursing home residents who were eligible for the SP interview. Respondents to the LML interview conducted when the original sample SP was deceased were proxy respondents. We included all variables as input for each of the trees.

Appendix Table 1 indicates the variables used in the final non-response cells for the 2015 Cohort tracker weights; an "a" indicates variables retained in the non-nursing home tree for the original sample, a "b" indicates those retained in the nursing home tree for the original sample, a "c" indicates those retained in the deceased original sample tree, a "d" indicates those retained in the non-nursing home tree for the replenishment sample, and an "e" indicates those retained in the nursing home tree for the replenishment sample. For living SPs in the original sample who were living in the community and other residential settings (not nursing homes) in Round 4 and those in nursing homes in Round 4, final nonresponse cells included 14 indicators and 1 indicator, respectively. Combinations of these variables created 26 nonresponse cells among the original sample in the non-nursing home group and 2 nonresponse cells among the nursing home group (See Appendix Figures 1 and 2). For deceased SPs in the original sample, the total of 4 final nonresponse cells included 3 indicators (See Appendix Figure 3). For living SPs in the replenishment sample who were residing in the community and other residential settings (not nursing homes) and those identified as nursing home residents based on the information from the MDS (as described above), final nonresponse cells included 12 and 2 indicators, respectively. Combinations of these variables created 23 nonresponse cells among the replenishment sample non-nursing home residents and 4 nonresponse cells among the nursing home group (See Appendix Figures 4 and 5).

The final step in creating the 2015 Cohort tracker weight involved raking the nonresponse adjusted weights to control totals developed from the 5% EDB extract (of Medicare beneficiaries as of September 30, 2014) that was used for sampling. For consistency, the raking adjustment also included the ineligibles (primarily deaths), since the frame that served as the source of the control totals also includes beneficiaries who were ineligible for NHATS. In Round 5, weight trimming was done in conjunction with this raking adjustment, due to a few outlier weights; this is discussed further in section 5.

As in Rounds 1 through 4, four dimensions were used in this Round 5 raking adjustment[4]:

  (1) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by sex by race from the EDB (Black; non-Black);
  (2) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by Census region;
  (3) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by MSA status (from the HISKEW); and
  (4) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by a binary indicator of whether the SP was enrolled in Medicare prior to age 65.

---

[4] For purposes of raking, age categories refer to age at Round 5 sampling.

In addition, a fifth dimension—whether or not the beneficiary was eligible for selection into the original sample (i.e., age 65 or older and enrolled in Medicare as of September 30, 2010)—was used.

2011 Cohort Weights

The 2011 Cohort Round 5 tracker weight applies only to the original sample, and followed the approach used to compute the Rounds 1 through 4 tracker weights. This process began with the Round 4 nonresponse adjusted tracker weight (prior to raking). This Round 4 weight accounted for differential probabilities of selection and included adjustments for nonresponse to the Round 1, Round 2, Round 3, and Round 4 interviews but was not raked to the HISKEW[5]. See Montaquila et al. (2012) for details on the specific methodology used in computing and adjusting the Round 1 weights; also, refer to Montaquila et al. (2014, 2015a, 2015b) for information about the specific adjustments applied in Rounds 2 through 4, respectively.

To produce the 2011 Cohort Round 5 tracker weight, two adjustments were made to the Round 4 nonresponse adjusted tracker weight—an adjustment for Round 5 nonresponse and a raking adjustment to estimated population totals from the Medicare EDB. Potential variables for creating nonresponse cells for the 2011 Cohort Round 5 tracker weights came four sources:
- Beneficiary information from the sampling frame (the 20% HISKEW File for the original sample; the 20% extract of the Medicare Enrollment Database for the replenishment sample), including demographic characteristics of the beneficiary (e.g., age as of September 30, 2014, gender) and geographic information (e.g., census division, metro and micropolitan status) based on the beneficiary's address in CMS' Medicare Enrollment Database (EDB);
- County-level demographic information based on the 5% HISKEW file or the 5% extract of the Enrollment Database (e.g., percent of beneficiaries in the county who are Black; percent of beneficiaries in the county who are Hispanic) for the county linked to the beneficiary's address from the EDB;
- Census tract-level information based on the 2010-2014 5-year American Community Survey (e.g. tract-level demographic information), based on linkages to the beneficiary's address from the EDB; and
- Variables from the NHATS Rounds 1, 2, 3, and 4 interviews (race/ethnicity, highest education, and Rounds 1, 2, 3, and 4 residential settings).

Appendix Table 2 provides weighted response rates (using the Round 4 nonresponse adjusted tracker weights that were the basis for the 2011 Cohort Round 5 tracker weights) by categories of the various indicators. We used these variables as input to a classification tree analysis to determine which of these variables were associated with nonresponse. This approach uses a search algorithm to identify variables associated with response propensities. At each step in the process, chi-square tests were performed to determine the most significant predictor of response, given the set of conditions already specified in the particular "branch." We also set a minimum cell size of 50.[6]

---

[5] The HISKEW file was a 20% sample of the Medicare enrollment database (as of Sept. 30, 2010) that served as the sampling frame for the original selection.

[6] The classification tree analysis is designed to work with categorical predictor variables. Alternatives to this approach are propensity modeling based on logistic regression and Cartesian product cross-classification. The logistic regression approach uses a parametric model to identify predictors of response. When the pool of potential predictors includes continuous variables and categorizing the continuous variables would result in

Separate trees were fit for all living non-nursing home cases (Figure 6), nursing home residents (Figure 7), and deceased SPs (Figure 8) because underlying nonresponse processes differed for these three groups. For the original sample, nursing home residents include both Round 1 residents who were not required to complete an SP Interview and new Round 2, Round 3, or Round 4 nursing home residents who were eligible for the SP interview. Respondents to the LML interview conducted when the SP was deceased were proxy respondents. We included all variables as input for each of the trees.

Appendix Table 2 indicates the variables used in the final nonresponse cells for the 2011 Cohort tracker weights, with an "a" for the non-nursing home tree, a "b" for the Round 4 nursing home residents tree, and a "c" for the deceased SP tree. For living SPs who were living in the community and other residential settings (not nursing homes) in Round 4 and those living in nursing homes in Round 4, final nonresponse cells included 11 indicators and 1 indicator, respectively; combinations of these variables created 26 nonresponse cells among the non-nursing home group and 2 nonresponse cells among the Round 4 nursing home residents. For deceased SPs, final non-response cells included 3 indicators, resulting in 4 nonresponse cells (See Appendix Figures 6, 7, and 8).

The final step in creating the 2011 Cohort tracker weight involved raking the nonresponse adjusted weights to control totals developed from the 5% EDB extract (of Medicare beneficiaries as of September 30, 2010) that was used for sampling of the original sample. For consistency, the raking adjustment also included the ineligibles (primarily deaths), since the frame that served as the source of the control totals also includes beneficiaries who were ineligible for NHATS. In Round 5, weight trimming was done in conjunction with this raking adjustment, due to a few outlier weights; this is discussed further in section 5.

As in Rounds 1 through 4, four dimensions were used in this Round 5 raking adjustment[7]:

(5) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by sex by race from the EDB (Black; non-Black);
(6) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by Census region;
(7) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by MSA status (from the HISKEW); and
(8) Age category (65-69, 70-74, 75-79, 80-84, 85-89, 90+) by a binary indicator of whether the SP was enrolled in Medicare prior to age 65.

4. Computation of Analytic Weights

As with the tracker weights, separate analytic weights were computed for the 2015 Cohort (designed for analysis of the original and replenishment samples combined) and for the 2011 Cohort (designed for

substantial losses of information, logistic regression modeling would be preferred over classification tree analysis. The Cartesian product cross-classification approach involves specifying a set of adjustment cell variables based on prior experience (generally, (1) prior analyses that identified predictors associated with response propensities; and/or (2) predictors associated with response and/or subject matter expertise that informs the choice of variables).

[7] For purposes of raking, age categories refer to age at sampling.

analysis of the original sample alone). The computation of the analytic weights begins with the final Round 5 tracker weight for the respective cohort. A weighting class adjustment was developed for the class of nonrespondents who were eligible for but did not complete the SP interview: those living in nursing homes or non-nursing home residential care in Round 5 who had completed a facility interview but not a Sample Person interview (n=179 for the 2015 Cohort and n=126 for the 2011 Cohort; designated as code 64). (Round 5 nursing home residents who were nursing home residents at the time of their baseline interview (code 61) were not eligible for an SP interview in Round 5, thus are not part of the analytic weight nonresponse adjustment). The approach was designed to preserve the tracker weight distributions by Round 5 residence type (nursing home, non-nursing home). That is, we allowed the weights of residential care cases with both a completed FQ and a completed SP interview (n=450 for the 2015 Cohort and n=285 for the 2011 Cohort) to be adjusted to account for similar cases missing the SP Interview.

2015 Cohort Weights

Because it was believed that response mechanisms may be different for the two samples (since members of the original sample had been engaged in the study for several rounds, whereas Round 5 was the first contact and attempt at gaining cooperation with the replenishment sample), the two samples were adjusted separately for Round 5 analytic nonresponse. Since the sample size is much smaller for this nonresponse adjustment, only a subset of variables used in tracker weight classification tree analysis was considered for the analytic weight nonresponse adjustments; additionally, three variables that characterize the Round 5 nursing home status, non-nursing home residential care status, and area of the facility where the SP lives were included (see Appendix Table 3). In order to preserve the tracker weight distribution for each sample separately by Round 5 residence type, the first split in the tree for original sample cases was forced to be Round 5 nursing home status. (All subsequent splitting was based on response propensities.) For the original sample, 5 variables (designated with "o" in Appendix Table 3) were retained in the final classification tree, forming 7 cells (see Appendix Figure 9); for the replenishment sample, 2 variables designated with "r" in Appendix Table 3) were retained in the final classification tree, forming 3 cells (see Appendix Figure 10).

As a final step, we applied a raking procedure so that marginal totals based on the analytic weights would match the totals at replenishment sampling by: 5-year age groups, sex, race, region, micro/metropolitan status, and whether Medicare was received before age 65 (see footnote 2).

2011 Cohort Weights

As with the 2011 Cohort tracker weights, the 2011 Cohort Round 5 analytic weight applies only to the original sample. Since the sample size is much smaller for this nonresponse adjustment, only a subset of variables used in tracker weight classification tree analysis was considered for the analytic weight nonresponse adjustments; additionally, three variables that characterize the Round 5 nursing home status, non-nursing home residential care status, and area of the facility where the SP lives were included (see Appendix Table 4). In order to preserve the tracker weight distribution by Round 5 residence type, the first split was forced to be Round 5 nursing home status. (All subsequent splitting was based on response propensities.) Six variables (designated with "*" in Appendix Table 4) were retained in the final classification tree, forming 7 cells (see Appendix Figure 11).

As a final step, we applied a raking procedure so that marginal totals based on the analytic weights would match the totals at sampling by: 5-year age groups, sex, race, region, micro/metropolitan status, and whether Medicare was received before age 65 (see footnote 2).

5. Design Effects Related to Weighting

Although weighting adjustments are aimed at reducing bias, increased variation in weights generally increases the variances of survey estimates (Kish, 1965). Thus, in the development and implementation of the weighting methodology for NHATS, care was taken to balance the bias reductions against the potential increases in variance.

The estimated overall design effect due to variation in the Round 1 nonresponse adjusted tracker weights was 1.28. After applying Round 2 nonresponse adjustments within cells determined by the classification tree results, the estimated overall design effect due to unequal weighting increased to 1.33. Incorporating the Round 3 nonresponse adjustments, the estimated overall design effect due to unequal weighting was 1.35, and after Round 4 nonresponse adjustment this overall design effect was 1.34.

2015 Cohort Weights

The composited weights used in computing the 2015 Cohort tracker weights had an overall design effect (due to variation in the weights) of 1.34. After Round 5 nonresponse adjustment, the overall design effect was 1.55, with the increase being due to the extent of variation in response propensities between and within the two samples. In order to limit the variation in the weights, after the raking adjustment, trimming of the tracker weights was considered; however, no influential outlier weights were identified, so no weights were trimmed at this stage. After the raking adjustment, the design effect for the final 2015 Cohort Round 5 tracker weights was 1.54.

The additional steps involved in creating the analytic weight (nonresponse adjustment and raking) did not increase the estimated overall design effect. However, one case was identified as an influential outlier, and its analytic weight was trimmed; following trimming, the weights were re-raked. After the re-raking, the design effect for the final 2015 Cohort Round 5 analytic weights was 1.53 overall, and 1.51 for living SPs and 1.31 for deceased SPs.

2011 Cohort Weights

For the 2011 Cohort weights, after Round 5 nonresponse adjustment, the overall design effect was 1.33. In order to limit the variation in the weights, after the raking adjustment, the tracker weights were trimmed and then re-raked; three cases with extreme weights were trimmed at this point. After the raking adjustment and trimming, the design effect for the final 2011 Cohort Round 5 tracker weights was 1.35.

The additional steps involved in creating the analytic weight (nonresponse adjustment and raking) had minimal effect on the estimated overall design effect (1.33 overall; 1.32 for living SPs and 1.39 for deceased SPs) and did not introduce any influential outlier weights.

6. Use of the Tracker vs. Analytic Weight

When using the tracker weight from any round, respondents are weighted up to represent all Medicare beneficiaries ages 65 and older who were alive on as of the target date for the cohort (September 30, 2014 for the 2015 Cohort; September 30, 2010 for the 2011 Cohort) and residing in the contiguous United States. In contrast, the analytic weight at a given round reproduces only those alive and eligible for NHATS during the prior round fieldwork period (with the exception of a small number of persons from the prior round who are deemed ineligible in the current round because they relocated outside the contiguous U.S.). Thus, the Round 5 analytic weight reproduces those alive and eligible for NHATS during the Round 4 fieldwork period.

The only other difference between the two sets of weights is the treatment of respondents who live in residential care settings other than nursing homes. In cases where an FQ interview was completed but an (eligible) SP interview was not completed in Round 5, a positive Round 5 weight sits in the tracker file and a zero Round 5 weight in the analytic file. The analytic weights of individuals with both an SP and FQ interview have been adjusted to represent these cases (persons assigned both an SP and FQ interview but with only an FQ). For all other respondents (including cases with proxy responses to the LML interview) the analytic and tracker weights are equal.

Most often analyses will use the analytic weight. The tracker weight is appropriate for making national estimates using the FQ information (e.g. for services available to older adults living in residential care settings) and for investigating the role of mortality on Round 1 disability estimates and successive cross-sections.

Another important consideration is whether to use a Round 1, Round 2, Round 3, Round 4, or Round 5 weight and, for Round 5, whether to use the 2015 Cohort weight or the 2011 Cohort weight. A useful rule of thumb is to always consider the population to which an estimate is being generalized. To estimate, for example, the proportion of the population in Round 1 who has a particular characteristic in Round 2, 3, 4, or 5 (measured in the SP interview) or who was in a particular type of residential care in Round 2, 3, 4, or 5 (measured in the FQ interview), a Round 1 weight should be used. The former would use the Round 1 analytic weight and the latter the Round 1 tracker weight. To estimate characteristics of people ages 75 and older in Round 5, or the characteristics of those living in residential care settings in Round 5 as measured in the Round 5 FQ interview, the Round 5 weight should be used. The former would use the Round 5 analytic weight and the latter the Round 5 tracker weight. To estimate characteristics of people 65 and older in Round 5, the 2015 Cohort Round 5 weight should be used. To examine associations between a characteristics in Round 5 and a characteristic in Round 1 (or any round prior to Round 5), the 2011 Cohort Round 5 weight should be used.

7. Variance Estimation

Two broad classes of methods have been developed for computation of standard errors of estimates from complex sample surveys: (1) Taylor series linearization and (2) replication methods. The NHATS data files contain the information necessary for analysts to use either of these approaches to compute standard errors. The "stratum" and "cluster" variables that allow users to compute variance estimates using Taylor series linearization are provided on the NHATS tracker and SP files as the variables w5varstrat and w5varunit, respectively.

As discussed in Montaquila, Freedman, Spillman, and Kasper (2012), for NHATS, the replication approach that was used is the modified balanced repeated replication (BRR) method suggested by Fay (Judkins 1990). When estimating the variance of ratios of rare subsets, one problem that occasionally arises from standard BRR is that one or more replicate estimates will be undefined due to zero denominators. Instead of increasing the weights of one half-sample by 100 percent and decreasing the weights of the other half-sample to zero as in standard BRR, Fay's method perturbs the weights by ±100(1-K) percent where K is referred to as "Fay's factor." The perturbation factor for standard BRR is 100 percent, or K=0. For NHATS, K = 0.3 was used.

Nonresponse adjustment and raking were repeated for each of the replicates. The final tracker replicate weights are provided in the variables w5trfinwgt1-w5trfinwgt56 for the 2015 Cohort and w5tr2011wgt1-w5tr2011wgt56 for the 2011 Cohort, and the analytic replicate weights are provided in the variables w5anfinwgt1-w5anfinwgt56 for the 2015 Cohort and w5an2011wgt1- w5an2011wgt56 for the 2011 Cohort. Through the creation of person-level replicate weights, Fay's method approximately reflects the contribution of variance due to nonresponse adjustments, calibration adjustments (e.g., poststratification or raking), and other weight adjustment factors that are dependent on the observed sample.

References

Judkins DR. (1990). Fay's method for variance estimation. *Journal of Official Statistics, 6(3)*, 223-239.

Kish L. (1965). *Survey sampling*. New York: John Wiley and Sons.

Montaquila J, Freedman VA, Edwards, B, & Kasper JD. 2012. *National Health and Aging Trends Study Round 1 Sample Design and Selection*. *NHATS Technical Paper #1*. Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Montaquila, J, Freedman, VA, Spillman, B, & Kasper, JD. 2012. *National Health and Aging Trends Study Development of Round 1 Survey Weights. NHATS Technical Paper #2.* Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Montaquila, J, Freedman, VA, Spillman, B, & Kasper, JD. 2014. *National Health and Aging Trends Study Development of Round 2 Survey Weights. NHATS Technical Paper #6.* Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Montaquila, J, Freedman, VA, Spillman, B, & Kasper, JD. 2015a. *National Health and Aging Trends Study Development of Round 3 Survey Weights. NHATS Technical Paper #9.* Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Montaquila, J, Freedman, VA, Spillman, B, & Kasper, JD. 2015b. *National Health and Aging Trends Study Development of Round 4 Survey Weights. NHATS Technical Paper #9.* Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Appendix: Variables Used in Nonresponse Adjustment for Round 5 NHATS Weights

Appendix Table 1. Response Rates by Various Indicators: NHATS Round 5, 2015 Cohort

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
OVERALL	73.7%		
BENEFICIARY INDICATORS		**R1 HIGHEST EDUCATION[4][#]** (EL1HIGSTSCHL_R)	
Age[1][a][d] (H_AGECAT_R5)		0: Not applicable	93.9%
1: 65-69	64.8%	1: DK/RF	92.3%
2: 70-74	79.0%	2: Below high school	96.0%
3: 75-79	75.5%	3: High school	95.7%
4: 80-84	75.4%	4: Above High school	96.3%
5: 85- 89	80.9%		
6: 90+	85.1%	**TRACT-LEVEL INDICATORS (Quartiles)**	
Gender[1][a] (H_SEX)		**Household Income[3]** (C_AGG_HH_INC)	
1: Male	73.4%	1: 1st quartile	77.9%
2: Female	73.9%	2: 2nd quartile	74.9%
Census Region[2] (S_REGION)		3: 3rd quartile	73.5%
1: Northeast	70.2%	4: 4th quartile	71.3%
2: Midwest	76.2%	9: Missing	86.5%
3: South	74.7%	**Median Household Income[3][a][d]** (C_MED_HH_INC)	
4: West	72.4%	1: 1st quartile	79.3%
Census Division[2][a][b][c][d][e] (DIVISION)		2: 2nd quartile	73.9%
1: New England	70.9%	3: 3rd quartile	74.2%
2: Middle Atlantic	69.8%	4: 4th quartile	68.8%
3: East North Central	75.2%	9: Missing	86.5%
4: West North Central	77.7%	**Median Household Income 65+[3][a][d]**	
5: South Atlantic	74.6%	(C_MED_HH_INC_65)	
6: East South Central	74.7%	1: 1st quartile	78.5%
7: West South Central	74.8%	2: 2nd quartile	74.8%
8: Mountain	77.5%	3: 3rd quartile	72.9%
9: Pacific	71.7%	4: 4th quartile	69.4%
Census Metro/Micro Area Designation (2013)[2]		9: Missing	90.5%
(S_METMICRO)		**% Households with Adult 65+[3][a]** (C_PCT_HH_65)	
1: Metropolitan area	72.5%	1: 1st quartile	74.3%
2: Micropolitan area	79.4%	2: 2nd quartile	72.8%
3: Non-metro	77.1%	3: 3rd quartile	74.1%
Health Maintenance Organization Beneficiary[1][d][e]		4: 4th quartile	73.7%
(HMOTYPE)		**% Households in Poverty[3][d]** (C_PCT_HH_POV)	
0: Yes	75.4%	1: 1st quartile	70.8%
9: No	72.9%	2: 2nd quartile	72.9%
Age First Enrolled in Medicare[1] (MEDIC_BEG)		3: 3rd quartile	74.0%
1: Prior to age 65	73.7%	4: 4th quartile	78.5%
2: At or after age 65	73.7%	**% Households Reporting Public Assistance[3][a]**	
R1 RACE ETHNICITY[4][#] (RL1DRACEHISP_R)		(C_PCT_HH_PUBASST)	
1: White, non-Hispanic	96.3%	1: 1st quartile	73.6%
2: Black, non-Hispanic	95.4%	2: 2nd quartile	72.6%
3: Other, non-Hispanic	94.6%	3: 3rd quartile	73.8%
4: Hispanic	94.5%	4: 4th quartile	74.9%
5: DK/RF	90.6%	**% Households Reporting Retirement Income[3]**	
Enhanced Race Indicator[1][^][d] (H_ENHRACEETH)		(C_PCT_HH_RETIREINC)	
1: Non-Hispanic Black	71.2%	1: 1st quartile	75.2%
2: Hispanic	68.4%	2: 2nd quartile	74.5%
3: White/Other	62.3%	3: 3rd quartile	72.9%
		4: 4th quartile	73.0%

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
TRACT-LEVEL INDICATORS (Quartiles)		**COUNTY LEVEL INDICATORS**	
% Households Reporting Social Security[3][a] (C_PCT_HH_SOCSEC)		**% Black 65+ (deciles)[2][a][c][d]** (PCTBLK)	
1: 1st quartile	73.8%	0: 1st decile	80.5%
2: 2nd quartile	71.8%	1: 2nd decile	76.4%
3: 3rd quartile	73.7%	2: 3rd decile	71.3%
4: 4th quartile	75.1%	3: 4th decile	74.6%
% Households Reporting SSI[3][a] (C_PCT_HH_SSS)		4: 5th decile	73.6%
1: 1st quartile	73.0%	5: 6th decile	71.1%
2: 2nd quartile	71.4%	6: 7th decile	72.3%
3: 3rd quartile	74.4%	7: 8th decile	72.8%
4: 4th quartile	76.2%	8: 9th decile	71.1%
% Households Owning Their Home[3][a] (C_PCT_OWNHOME)		9: 10th decile	72.0%
1: 1st quartile	75.7%	**% Hispanic 65+ (deciles)[2][a][d]** (PCTHISP)	
2: 2nd quartile	73.9%	0: 1st decile	74.3%
3: 3rd quartile	73.2%	1: 2nd decile	76.1%
4: 4th quartile	72.8%	2: 3rd decile	76.0%
% Households 65+ Owning Their Home[3] (C_PCT_OWNHOME_65)		3: 4th decile	75.9%
1: 1st quartile	74.5%	4: 5th decile	78.7%
2: 2nd quartile	74.1%	5: 6th decile	73.4%
3: 3rd quartile	74.2%	6: 7th decile	68.1%
4: 4th quartile	72.1%	7: 8th decile	71.7%
% Households 65+ Below Poverty[3][d] (C_PCT_POV_65)		8: 9th decile	70.4%
1: 1st quartile	70.9%	9: 10th decile	72.9%
2: 2nd quartile	73.0%	**% Poverty (deciles)[2][a][d]** (PCTPOV)	
3: 3rd quartile	73.5%	0: 1st decile	71.7%
4: 4th quartile	76.6%	1: 2nd decile	73.4%
Per Capita Income[3][a][d] (C_PER_CAP_INC)		2: 3rd decile	73.0%
1: 1st quartile	78.8%	3: 4th decile	77.1%
2: 2nd quartile	75.6%	4: 5th decile	74.1%
3: 3rd quartile	72.2%	5: 6th decile	72.3%
4: 4th quartile	69.8%	6: 7th decile	73.1%
		7: 8th decile	73.2%
		8: 9th decile	74.2%
		9: 10th decile	75.8%

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
OTHER INDICATORS			
R4 RESIDENTIAL CARE STATUS[4][#][c] (R4DRESID)		**MDS Match Algorithm Indicator**[^] (MDSMATCH)	
1: R4 Community	96.0%	1: NH Resident	81.3%
2: R4 Residential Care Resident not nursing home (SP interview complete)	98.4%	2: Not NH Resident	63.1%
		Licensed Assisted Living Match Indicator[^] (ALADDRMATCH)	
3: R4 Residential Care Resident not nursing home (FQ only)	88.9%	1: AL Resident	94.3%
		0: Not AL Resident	63.1%
4: R4 nursing home (SP interview complete)	97.4%		
5: R4 nursing home (FQ only)	86.0%		
7: R1 to R3 Residential Care Resident not nursing home (FQ only)	92.0%		
8: R1 to R3 nursing home	95.6%		

[1]Based on Information either on the September 30, 2010 CMS 20% Health Insurance Skeleton Eligibility Write-Off (HISKEW) file if the case is in the original sample, or on the September 30, 2014 CMS 20% Enrollment Database (EDB) extract if the case is in the replenishment sample .

[2]Based on county-level information from the September 30, 2014 CMS 5% EDB extract linked to the beneficiary's EDB address.

[3]Based on tract-level information from the 2009-2013 5-year American Community Survey file linked to the beneficiary's EDB address.

[4]Based on responses to items in the Rounds 1 and 4 interviews.

[#]Response rates were computed only for the original sample.

[^] Response rates were computed only for the replenishment sample.

a=retained in classification tree analysis for living SP non-nursing home branch of the original sample

b=retained in classification tree analysis for living SP nursing home branch of the original sample

c=retained in classification tree analysis for deceased SP branch of the original sample

d= retained in classification tree analysis for living SP non-nursing home branch (MDSMATCH=0) of the replenishment sample

e= retained in classification tree analysis for living SP nursing home branch (MDSMATCH=1) of the replenishment sample

N=10,817 (8,334 respondents and 2,483 non-respondents)

Variable names used in classification trees shown parenthetically.

Appendix Table 2. Response Rates by Various Indicators: NHATS Round 5, 2011 Cohort

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
OVERALL	96.0%	TRACT-LEVEL INDICATORS (Quartiles)	
BENEFICIARY INDICATORS		Household Income[3] (C_AGG_HH_INC)	
Age[1][a] (H_AGECAT)		1: 1st quartile	96.4%
1: 65-69	96.2%	2: 2nd quartile	95.8%
2: 70-74	96.5%	3: 3rd quartile	96.3%
3: 75-79	96.1%	4: 4th quartile	95.8%
4: 80-84	94.8%		
5: 85- 89	95.5%	Median Household Income[3][a] (C_MED_HH_INC)	
6: 90+	94.9%	1: 1st quartile	95.7%
Gender[1] (H_SEX)		2: 2nd quartile	96.9%
1: Male	96.2%	3: 3rd quartile	96.1%
2: Female	95.8%	4: 4th quartile	95.4%
Census Region[1][c] (S_REGION)			
1: Northeast	94.9%	Median Household Income 65+[3]	
2: Midwest	97.6%	(C_MED_HH_INC_65)	
3: South	96.5%	1: 1st quartile	96.1%
4: West	94.3%	2: 2nd quartile	96.8%
Census Division[1][a][b] (DIVISION)		3: 3rd quartile	95.4%
1: New England	93.3%	4: 4th quartile	95.8%
2: Middle Atlantic	95.7%	9: Missing	100%
3: East North Central	97.4%	% Households with Adult 65+[3] (C_PCT_HH_65)	
4: West North Central	97.9%	1: 1st quartile	95.6%
5: South Atlantic	96.6%	2: 2nd quartile	96.1%
6: East South Central	96.5%	3: 3rd quartile	96.1%
7: West South Central	96.3%	4: 4th quartile	96.0%
8: Mountain	92.8%	% Households in Poverty[3] (C_PCT_HH_POV)	
9: Pacific	94.5%	1: 1st quartile	95.6%
Census Metro/Micro Area Designation (2013)[2]		2: 2nd quartile	96.1%
(S_METMICRO)		3: 3rd quartile	96.7%
1: Metropolitan area	95.6%	4: 4th quartile	95.5%
2: Micropolitan area	97.2%	% Households Reporting Public Assistance[3][a]	
3: Non-metro	98.5%	(C_PCT_HH_PUBASST)	
Health Maintenance Organization Beneficiary[1]		1: 1st quartile	96.0%
(HMOTYPE)		2: 2nd quartile	96.0%
0: Yes	96.7%	3: 3rd quartile	96.1%
9: No	95.7%	4: 4th quartile	95.9%
Age First Enrolled in Medicare[1] (MEDIC_BEG)		% Households Reporting Retirement Income[3]	
1: Prior to age 65	95.9%	(C_PCT_HH_RETIREINC)	
2: At or after age 65	96.0%	1: 1st quartile	96.1%
R1 RACE ETHNICITY[4] (RL1DRACEHISP_R)		2: 2nd quartile	96.7%
1: White, non-Hispanic	96.3%	3: 3rd quartile	95.5%
2: Black, non-Hispanic	95.5%	4: 4th quartile	95.9%
3: Other, non-Hispanic	95.0%	% Households Reporting Social Security[3]	
4: Hispanic	94.6%	(C_PCT_HH_SOCSEC)	
5: DK/RF	90.7%	1: 1st quartile	96.2%
R1 HIGHEST EDUCATION[4] (EL1HIGSTSCHL_R)		2: 2nd quartile	95.8%
0: Not applicable	94.0%	3: 3rd quartile	95.3%
1: DK/RF	92.5%	4: 4th quartile	96.6%
2: Below high school	96.1%		
3: High school	95.6%		
4: Above High school	96.3%		

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
COUNTY LEVEL INDICATORS		**TRACT-LEVEL INDICATORS (Quartiles)**	
		% Households Reporting SSI[3][a] (C_PCT_HH_SSS)	
% Black 65+ (deciles)[2][a] (PCTBLK)		1: 1st quartile	96.6%
0: 1st decile	97.5%	2: 2nd quartile	95.5%
1: 2nd decile	97.4%	3: 3rd quartile	95.9%
2: 3rd decile	95.9%	4: 4th quartile	96.1%
3: 4th decile	95.2%	**% Households Owning Their Home[3][a]**	
4: 5th decile	94.0%	(C_PCT_OWNHOME)	
5: 6th decile	95.0%	1: 1st quartile	95.6%
6: 7th decile	95.7%	2: 2nd quartile	95.8%
7: 8th decile	96.4%	3: 3rd quartile	97.2%
8: 9th decile	97.5%	4: 4th quartile	95.3%
9: 10th decile	95.7%	**% Households 65+ Owning Their Home[3]**	
		(C_PCT_OWNHOME_65)	
% Hispanic 65+ (deciles)[2][a][c] (PCTHISP)		1: 1st quartile	96.1%
0: 1st decile	98.4%	2: 2nd quartile	95.4%
1: 2nd decile	95.7%	3: 3rd quartile	96.5%
2: 3rd decile	96.9%	4: 4th quartile	95.9%
3: 4th decile	97.1%	**% Households 65+ Below Poverty[3][a]**	
4: 5th decile	97.2%	(C_PCT_POV_65)	
5: 6th decile	96.4%	1: 1st quartile	95.7%
6: 7th decile	94.2%	2: 2nd quartile	96.4%
7: 8th decile	95.2%	3: 3rd quartile	96.1%
8: 9th decile	93.6%	4: 4th quartile	95.8%
9: 10th decile	95.1%	**Per Capita Income[3][a]** (C_PER_CAP_INC)	
		1: 1st quartile	95.7%
% Poverty (deciles)[2][a] (PCTPOV)		2: 2nd quartile	96.5%
0: 1st decile	96.7%	3: 3rd quartile	95.8%
1: 2nd decile	97.4%	4: 4th quartile	95.9%
2: 3rd decile	93.3%		
3: 4th decile	96.7%	**OTHER INDICATORS**	
4: 5th decile	96.6%	**R4 RESIDENTIAL CARE STATUS[4][c]** (R4DRESID)	
5: 6th decile	96.7%	1: R4 Community	96.0%
6: 7th decile	96.9%	2: R4 Residential Care Resident not nursing home (SP interview complete)	98.3%
7: 8th decile	94.3%	3: R4 Residential Care Resident not nursing home (FQ only)	89.3%
8: 9th decile	95.7%		
9: 10th decile	95.7%	4: R4 nursing home (SP interview complete)	97.6%
		5: R4 nursing home (FQ only)	86.5%
		7: R1-R3 Residential Care Resident not nursing home (FQ only)	92.4%
		8: R1- R3 nursing home	95.4%

[1]Based on Information on the September 30, 2010 CMS 20% Health Insurance Skeleton Eligibility Write-Off (HISKEW) file.
[2]Based on county-level information from the September 30, 2014 CMS 5% EDB extract linked to the beneficiary's EDB address.
[3]Based on tract-level information from the 2009-2013 5-year American Community Survey file linked to the beneficiary's EDB address.
[4]Based on responses to items in the Rounds 1 and 4 interviews.
a=retained in classification tree analysis for living SP non-nursing home branch
b=retained in classification tree analysis for living SP nursing home branch
c=retained in classification tree analysis for deceased SP branch
N=4,330 (4,152 respondents and 178 non-respondents)
Variable names used in classification trees shown parenthetically.

Appendix Table 3. Sampled Person Interview Response Rates Among Cases with Completed Facility Questionnaires, by Various Indicators: NHATS Round 5, 2015 Cohort

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
OVERALL	71.8%	COUNTY LEVEL INDICATORS	
BENEFICIARY INDICATORS		% Black 65+ (deciles)[2][o]	
Age[1] (H_AGECAT_R5)		(PCTBLK)	
1: 65-69	86.5%	0: 1st decile	65.2%
2: 70-74	76.8%	1: 2nd decile	72.6%
3: 75-79	63.1%	2: 3rd decile	68.4%
4: 80-84	76.6%	3: 4th decile	75.3%
5: 85- 89	68.0%	4: 5th decile	73.5%
6: 90+	71.8%	5: 6th decile	68.1%
R1 Race Ethnicity[4][#] (RL1DRACEHISP_R)		6: 7th decile	70.5%
1: White, non-Hispanic	70.4%	7: 8th decile	72.7%
2: Black, non-Hispanic	68.3%	8: 9th decile	80.4%
3: Other, non-Hispanic	61.7%	9: 10th decile	78.9%
4: Hispanic	40.5%		
5: DK/RF	50.9%	% Hispanic 65+ (deciles)[2] (PCTHISP)	
Enhanced Race Indicator[1][^]		0: 1st decile	76.9%
(H_ENHRACEETH)		1: 2nd decile	73.2%
1: Non-Hispanic Black	89.0%	2: 3rd decile	73.6%
2: Hispanic	53.1%	3: 4th decile	73.2%
3: White/Other	74.2%	4: 5th decile	78.8%
Gender[1] (H_SEX)		5: 6th decile	76.5%
1: Male	78.2%	6: 7th decile	64.5%
2: Female	69.1%	7: 8th decile	59.9%
Census Region[1] (S_REGION)		8: 9th decile	73.9%
1: Northeast	64.2%	9: 10th decile	67.4%
2: Midwest	69.4%		
3: South	80.9%	% Poverty (deciles)[2] (PCTPOV)	
4: West	68.6%	0: 1st decile	58.8%
Census Division[1][o][r] (DIVISION)		1: 2nd decile	68.4%
1: New England	74.6%	2: 3rd decile	77.1%
2: Middle Atlantic	58.7%	3: 4th decile	74.6%
3: East North Central	70.6%	4: 5th decile	70.8%
4: West North Central	68.4%	5: 6th decile	71.6%
5: South Atlantic	78.0%	6: 7th decile	60.6%
6: East South Central	89.7%	7: 8th decile	75.6%
7: West South Central	84.3%	8: 9th decile	81.3%
8: Mountain	64.4%	9: 10th decile	87.4%
9: Pacific	69.6%		
Census Metro/Micro Area Designation (2013)[1]		OTHER INDICATORS	
(S_METMICRO)		Facility Type Indicator[3][o][r] (FQ5DLOCSP)	
1: Metropolitan area	72.8%	1: Independent living/other	79.0%
2: Micropolitan area	59.2%	2: Assisted Living	71.3%
3: Non-metro	79.0%	3: Special care/memory care/Alzheimers unit	53.7%
Health Maintenance Organization Beneficiary[1]		4: Nursing home	60.6%
(HMOTYPE)		8: Facility type not reported	36.3%
0: Yes	68.0%	R1 RESIDENTIAL CARE STATUS[4][#][o]	
9: No	73.3%	(R1DRESID_R)	
Age First Enrolled in Medicare[1]		1: Community	78.9%
(MEDIC_BEG)		2: Residential Care Resident not nursing home	56.4%
1: Prior to age 65	67.1%		
2: At or after age 65	72.4%		

Variable & Values		Weighted Response Rate	Variable & Values		Weighted Response Rate
OTHER INDICATORS			**R2 RESIDENTIAL CARE STATUS[5 #]**	(R2DRESID_R)	
R2 NURSING HOME STATUS[5 #]	(R2NH)		1: Community in R2		77.5%
1: Yes		69.1%	2: Residential care in R2		61.4%
2: No		69.0%	3: Nursing home in R2		69.1%
R3 NURSING HOME STATUS[6 #]	(R3NH)		**R3 RESIDENTIAL CARE STATUS[6 #]**	(R3DRESID_R)	
1: Yes		52.8%	1: Community in R3		79.6%
2: No		71.1%	2: Residential care in R3		65.9%
R4 NURSING HOME STATUS[7 #]	(R4NH)		3: Nursing home in R3		52.8%
1: Yes		54.0%	**R4 RESIDENTIAL CARE STATUS[7 #]**	(R4DRESID_R)	
2: No		72.5%	1: Community in R4		82.3%
R5 NURSING HOME STATUS[8 # o]	(R5NH)		2: Residential care in R4		70.1%
1: Yes		58.7%	3: Nursing home in R4		54.0%
2: No		73.2%	**R5 RESIDENTIAL CARE STATUS[8]**	(R5DRESID_R)	
			2: Residential care in R5		74.0%
			3: Nursing home in R5		58.7%

[1]Based on Information either on the September 30, 2010 CMS 20% Health Insurance Skeleton Eligibility Write-Off (HISKEW) file if the case is in the original sample, or on the September 30, 2014 CMS 20% Enrollment Database (EDB) extract if the case is in the replenishment sample .

[2]Based on county-level information from the September 30, 2014 CMS 5% EDB extract linked to the beneficiary's EDB address.

[3]Based on the responses to two items on the type of facility from the FQ, FQ6 (fq5facdescri; including answers from FQ6A) and FQ10 (fq5faaretype).

[4]Based on responses to items in the Round 1 interview or interview process.

[5]Based on responses to items in the Round 2 interview or interview process.

[6]Based on responses to items in the Round 3 interview or interview process.

[7]Based on responses to items in the Round 4 interview or interview process.

[8]Based on responses to items in the Round 5 interview or interview process.

[#]Response rates were computed only for the available original sample.

[^] Response rates were computed only for the available replenishment sample.

o=retained in classification tree analysis for adjustment of missing SP interview of the original sample.

r=retained in classification tree analysis for adjustment of missing SP interview of the replenishment sample.

  N=629 (450 respondents and 179 nonrespondents); Variable names used in classification trees shown parenthetically.

Appendix Table 4. Sampled Person Interview Response Rates Among Cases with Completed Facility Questionnaires, by Various Indicators: NHATS Round 5, 2011 Cohort

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
OVERALL	69.8%	COUNTY LEVEL INDICATORS	
BENEFICIARY INDICATORS		% Black 65+ (deciles)[2] (PCTBLK)	
Age[1] * (H_AGECAT)		0: 1st decile	63.9%
1: 65-69	64.5%	1: 2nd decile	79.9%
2: 70-74	78.8%	2: 3rd decile	62.0%
3: 75-79	80.1%	3: 4th decile	69.5%
4: 80-84	64.3%	4: 5th decile	74.0%
5: 85- 89	61.6%	5: 6th decile	54.8%
6: 90+	73.2%	6: 7th decile	76.0%
		7: 8th decile	77.9%
R1 Race Ethnicity[4] (RL1DRACEHISP_R)		8: 9th decile	75.8%
1: White, non-Hispanic	71.5%	9: 10th decile	74.5%
2: Black, non-Hispanic	68.3%		
3: Other, non-Hispanic	59.0%	% Hispanic 65+ (deciles)[2] (PCTHISP)	
4: Hispanic	39.1%	0: 1st decile	70.2%
5: DK/RF	49.7%	1: 2nd decile	75.5%
		2: 3rd decile	76.9%
Gender[1] (H_SEX)		3: 4th decile	70.7%
1: Male	73.5%	4: 5th decile	81.8%
2: Female	68.4%	5: 6th decile	72.4%
		6: 7th decile	47.7%
Census Region[1] (S_REGION)		7: 8th decile	59.3%
1: Northeast	52.6%	8: 9th decile	75.8%
2: Midwest	71.3%	9: 10th decile	65.7%
3: South	77.5%		
4: West	73.1%	% Poverty (deciles)[2] * (PCTPOV)	
Census Division[1] * (DIVISION)		0: 1st decile	56.4%
1: New England	52.1%	1: 2nd decile	66.3%
2: Middle Atlantic	52.8%	2: 3rd decile	77.6%
3: East North Central	73.3%	3: 4th decile	60.3%
4: West North Central	69.2%	4: 5th decile	69.1%
5: South Atlantic	76.3%	5: 6th decile	76.8%
6: East South Central	82.8%	6: 7th decile	59.1%
7: West South Central	77.0%	7: 8th decile	72.6%
8: Mountain	74.9%	8: 9th decile	90.9%
9: Pacific	72.5%	9: 10th decile	85.1%
Census Metro/Micro Area Designation (2013)[2]		OTHER INDICATORS	
(S_METMICRO)		Facility Type Indicator[3] * (FQ5DLOCSP)	
1: Metropolitan area	69.9%	1: Independent living/other	75.7%
2: Micropolitan area	67.7%	2: Assisted Living	71.6%
3: Non-metro	71.8%	3: Special care/memory care/Alzheimers unit	68.7%
		4: Nursing home	60.1%
Health Maintenance Organization Beneficiary[1]		8: Facility type not reported	100.0%
(HMOTYPE)			
0: Yes	72.1%	R1 RESIDENTIAL CARE STATUS[4] * (R1DRESID_R)	
9: No	69.1%	1: Community	79.9%
		2: Residential Care Resident not nursing home	57.0%
Age First Enrolled in Medicare[1] (MEDIC_BEG)			
1: Prior to age 65	71.3%		
2: At or after age 65	69.6%		

Variable & Values		Weighted Response Rate	Variable & Values		Weighted Response Rate
OTHER INDICATORS			**OTHER INDICATORS**		
R2 NURSING HOME STATUS[5]	(R2NH)		**R2 RESIDENTIAL CARE STATUS[5]**	(R2DRESID_R)	
1: Yes		70.6%	1: Community in R2		78.4%
2: No		69.8%	2: Residential care in R2		62.0%
R3 NURSING HOME STATUS[6]	(R3NH)		3: Nursing home in R2		70.1%
1: Yes		53.3%	**R3 RESIDENTIAL CARE STATUS[6]**	(R3DRESID_R)	
2: No		71.8%	1: Community in R3		80.5%
R4 NURSING HOME STATUS[7]	(R4NH)		2: Residential care in R3		66.7%
1: Yes		54.4%	3: Nursing home in R3		53.3%
2: No		73.5%	**R4 RESIDENTIAL CARE STATUS[7]**	(R4DRESID_R)	
R5 NURSING HOME STATUS[8] *	(R5NH)		1: Community in R4		83.5%
1: Yes		58.5%	2: Residential care in R4		71.0%
2: No		74.4%	3: Nursing home in R4		54.4%
			R5 RESIDENTIAL CARE STATUS[8]	(R5DRESID_R)	
			2: Residential care in R5		74.4%
			3: Nursing home in R5		58.5%

[1]Based on Information on the September 30, 2010 CMS 20% Health Insurance Skeleton Eligibility Write-Off (HISKEW) file.
[2]Based on county-level information from the September 30, 2014 CMS 5% EDB extract linked to the beneficiary's EDB address.
[3]Based on the responses to two items on the type of facility from the FQ, FQ6 (fq5facdescri; including answers from FQ6A) and FQ10 (fq5faaretype).
[4]Based on responses to items in the Round 1 interview or interview process.
[5]Based on responses to items in the Round 2 interview or interview process.
[6]Based on responses to items in the Round 3 interview or interview process.
[7]Based on responses to items in the Round 4 interview or interview process.
[8]Based on responses to items in the Round 5 interview or interview process.
*=retained in classification tree analysis for adjustment of missing SP interview.
 N=411 (285 respondents and 126 nonrespondents); Variable names used in classification trees shown parenthetically.

Figure 1. Round 5 2015 Cohort tracker weight nonresponse adjustment cells – non nursing home cases in original sample
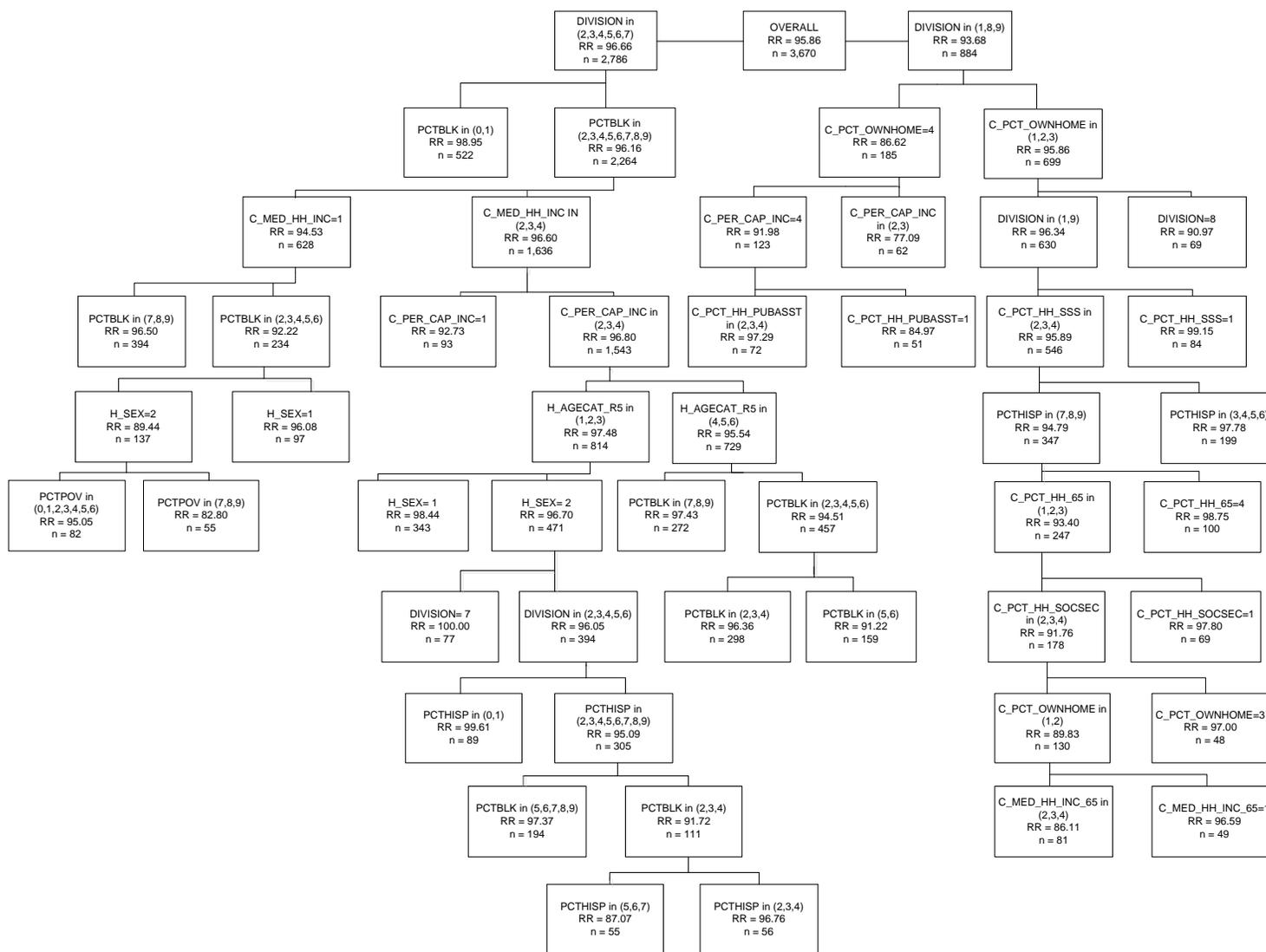
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 2. Round 5 2015 Cohort tracker weight nonresponse adjustment cells – nursing home cases in original sample

```
                    ┌─────────────────┐
                    │    OVERALL      │
                    │  RR = 97.13     │
                    │   n = 186       │
                    └────────┬────────┘
              ┌──────────────┴──────────────┐
    ┌─────────────────────┐      ┌─────────────────────┐
    │ DIVISION in (3,5,9)  │      │    DIVISION in       │
    │   RR = 93.30        │      │   (1,2,4,6,7,8)      │
    │    n = 84           │      │   RR = 100.00        │
    │                     │      │    n = 102           │
    └─────────────────────┘      └─────────────────────┘
```
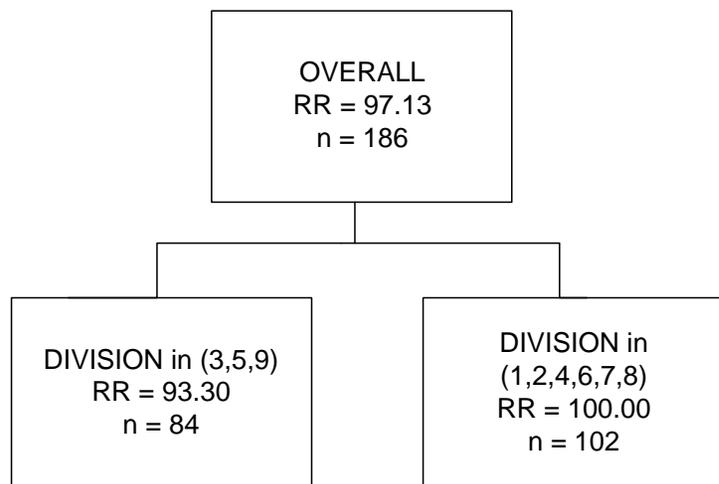
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 3. Round 5 2015 Cohort tracker weight nonresponse adjustment cells – deceased cases in original sample


```
                              ┌─────────────────┐
                              │     OVERALL     │
                              │   RR = 97.10    │
                              │    n = 296      │
                              └─────────────────┘
                    ┌──────────────────┴──────────────────┐
        ┌───────────────────────┐              ┌───────────────────────┐
        │  R4DRESID in (1,2,4)   │              │ R4DRESID in (3,5,7,8)  │
        │     RR = 98.51         │              │     RR = 89.88         │
        │      n = 234           │              │      n = 62            │
        └───────────────────────┘              └───────────────────────┘
        ┌───────────┴───────────┐
┌───────────────────────┐  ┌───────────────────────┐
│ DIVISION in (1,3,4,6,9) │  │ DIVISION in (2,5,7,8)  │
│     RR = 100.00         │  │     RR = 96.49         │
│      n = 134            │  │      n = 100           │
└───────────────────────┘  └───────────────────────┘
                       ┌───────────┴───────────┐
           ┌───────────────────────┐  ┌───────────────────────┐
           │ PCTBLK in (0,2,3,4,5)  │  │ PCTBLK in (6,7,8,9)    │
           │     RR = 94.98         │  │     RR = 97.93         │
           │      n = 48            │  │      n = 52            │
           └───────────────────────┘  └───────────────────────┘
```
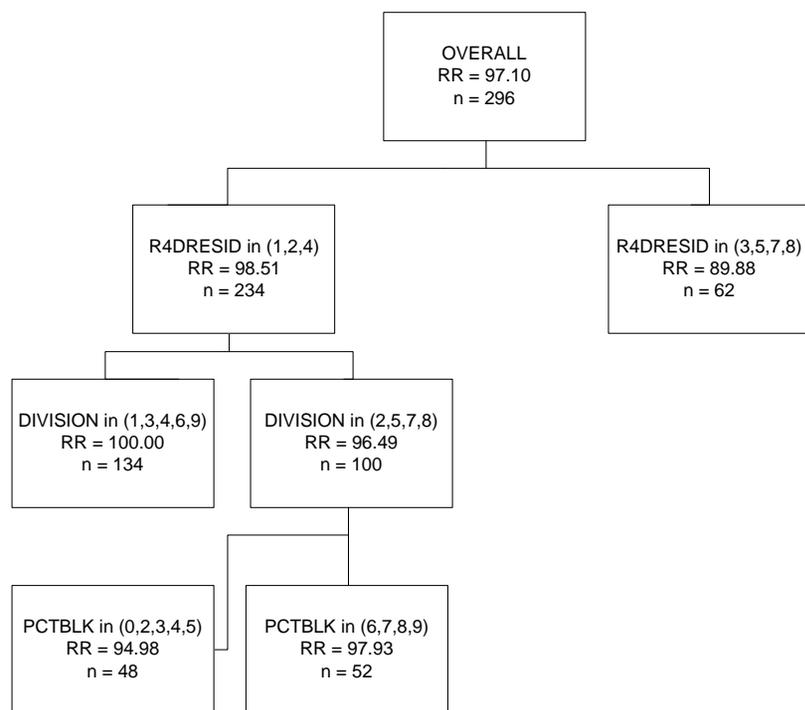
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 4. Round 5 2015 Cohort tracker weight nonresponse adjustment cells – non nursing home cases in replenishment sample

OVERALL
RR = 63.10
n = 3,998

C_PER_CAP_INC in (3,4)
RR = 59.27
n = 1,930

C_PER_CAP_INC in (1,2)
RR = 67.71
n = 2,068

PCTBLK=0
RR = 73.37
n = 166

PCTBLK in (1,2,3,4,5,6,7,8,9)
RR = 58.12
n = 1,764

C_MED_HH_INC in (2,3,4,9)
RR = 64.60
n = 1,050

C_MED_HH_INC=1
RR = 72.04
n = 1,018

PCTBLK in (8,9)
RR = 49.68
n = 217

PCTBLK in (1,2,3,4,5,6,7)
RR = 59.26
n = 1,547

PCTBLK in (4,5,6,7,8,9)
RR = 61.49
n = 565

PCTBLK in (0,1,2,3)
RR = 67.80
n = 485

H_ENHRACEETH=3
RR = 68.67
n = 482

H_ENHRACEETH in (1,2)
RR = 78.65
n = 536

PCTHISP in (0,1,2,3,4,5)
RR = 61.64
n = 718

PCTHISP in (6,7,8,9)
RR = 57.30
n = 829

HMOTYPE=0
RR = 66.98
n = 222

HMOTYPE=9
RR = 58.50
n = 343

PCTPOV in (0,1,2,3,4,5,6,7)
RR = 66.19
n = 421

PCTPOV in (8,9)
RR = 79.66
n = 64

PCTBLK in (5,6,7,8,9)
RR = 62.77
n = 178

PCTBLK in (0,1,2,3,4)
RR = 72.80
n = 304

H_AGECAT_R5 in (2,3,4,5,6)
RR = 65.61
n = 506

H_AGECAT_R5=1
RR = 57.43
n = 212

H_ENHRACEETH in (2,3)
RR = 56.43
n = 222

H_ENHRACEETH=1
RR = 70.67
n = 121

PCTHISP in (1,2,3,4,5,6,7,8)
RR = 68.29
n = 364

PCTHISP=0
RR = 56.20
n = 57

DIVISION in (5,7,9)
RR = 59.17
n = 590

DIVISION in (1,2,8)
RR = 53.11
n = 239

C_PCT_POV_65=1
RR = 66.27
n = 57

C_PCT_POV_65 in (2,3,4)
RR = 53.47
n = 165

C_PER_CAP_INC=2
RR = 69.69
n = 330

C_PER_CAP_INC=1
RR = 56.89
n = 34

H_AGECAT_R5=2
RR = 71.68
n = 148

H_AGECAT_R5 in (3,4,5,6)
RR = 61.78
n = 358

C_MED_HH_INC_65 in (1,2)
RR = 67.24
n = 82

C_MED_HH_INC_65 in (3,4,9)
RR = 57.97
n = 508

DIVISION in (1,2,7)
RR = 61.08
n = 72

DIVISION in (3,4,5,8,9)
RR = 72.65
n = 258

C_PCT_POV_65=4
RR = 49.25
n = 74

C_PCT_POV_65 in (1,2,3)
RR = 59.57
n = 434

C_PCT_HH_POV in (2,3,4)
RR = 63.75
n = 176

C_PCT_HH_POV=1
RR = 57.02
n = 258

H_AGECAT_R5=1
RR = 52.83
n = 74

H_AGECAT_R5 in (2,3,4,5,6)
RR = 61.06
n = 184

Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 5. Round 5 2015 Cohort tracker weight nonresponse adjustment cells – nursing home cases in replenishment sample
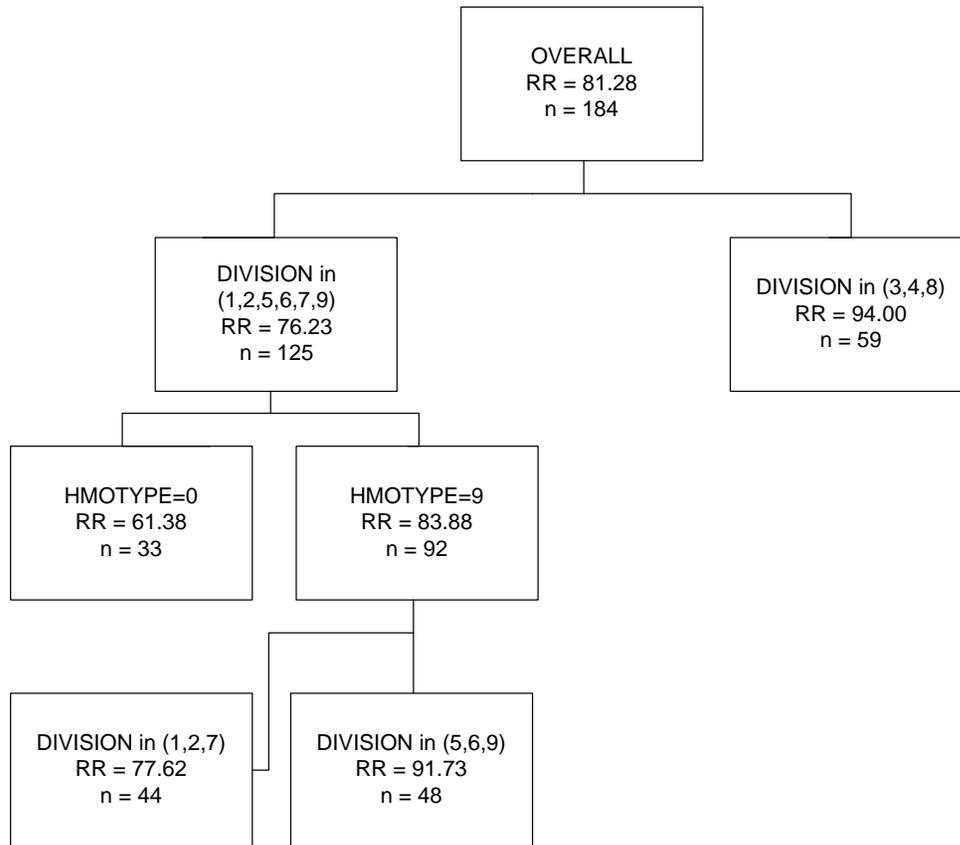
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 6. Round 5 2011 Cohort tracker weight nonresponse adjustment cells – non nursing home cases in original sample
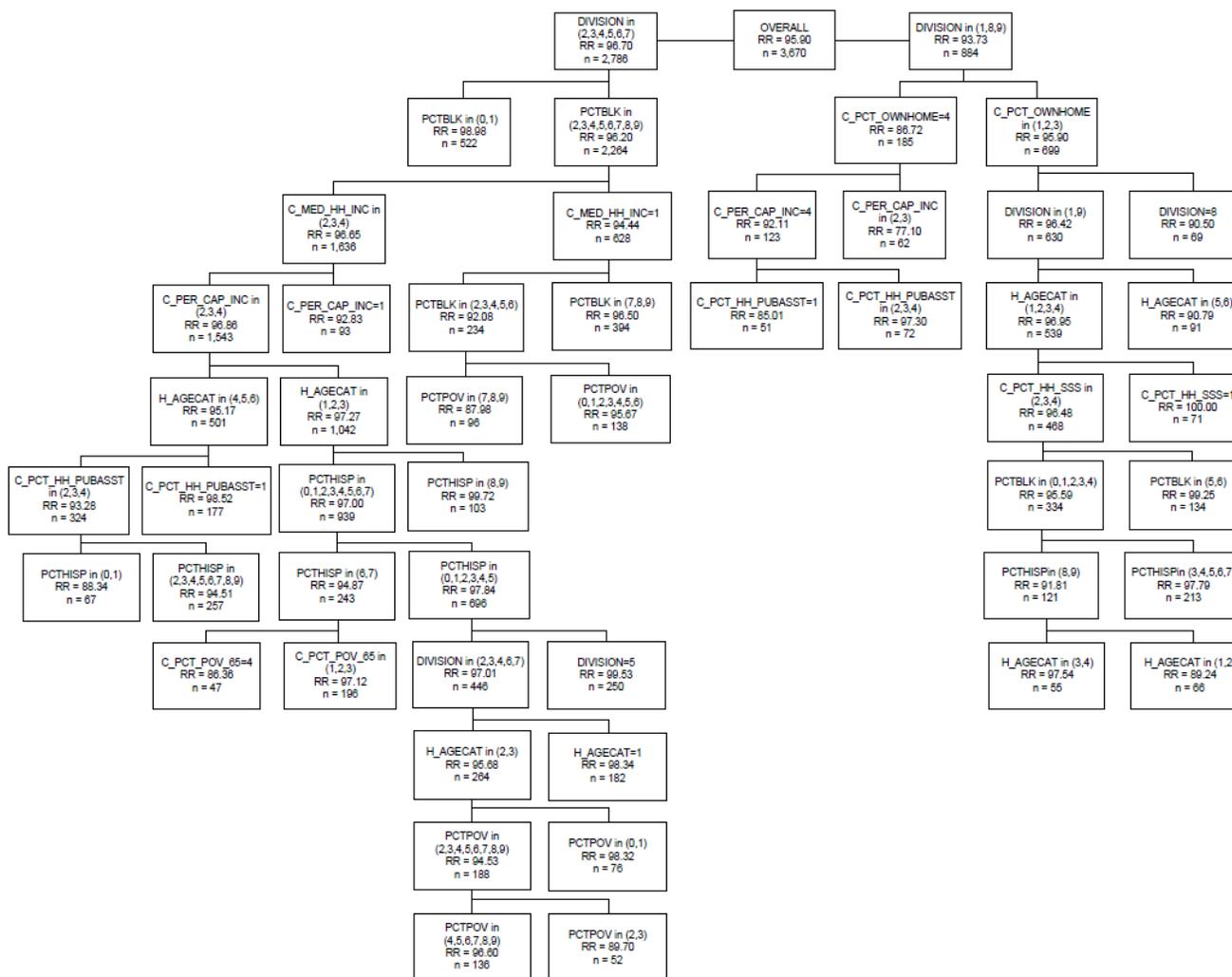
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 7. Round 5 2011 Cohort tracker weight nonresponse adjustment cells – nursing home cases in original sample

```
                    ┌─────────────────┐
                    │    OVERALL      │
                    │  RR = 97.38     │
                    │   n = 186       │
                    └────────┬────────┘
              ┌──────────────┴──────────────┐
    ┌─────────────────┐          ┌─────────────────┐
    │ DIVISION in (3,5,9) │      │   DIVISION in    │
    │   RR = 93.84    │          │  (1,2,4,6,7,8)  │
    │    n = 84       │          │  RR = 100.00    │
    │                 │          │   n = 102       │
    └─────────────────┘          └─────────────────┘
```
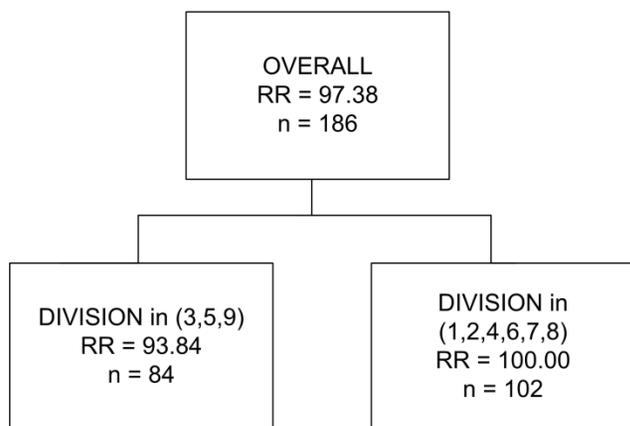
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 8. Round 5 2011 Cohort tracker weight nonresponse adjustment cells – deceased cases in original sample
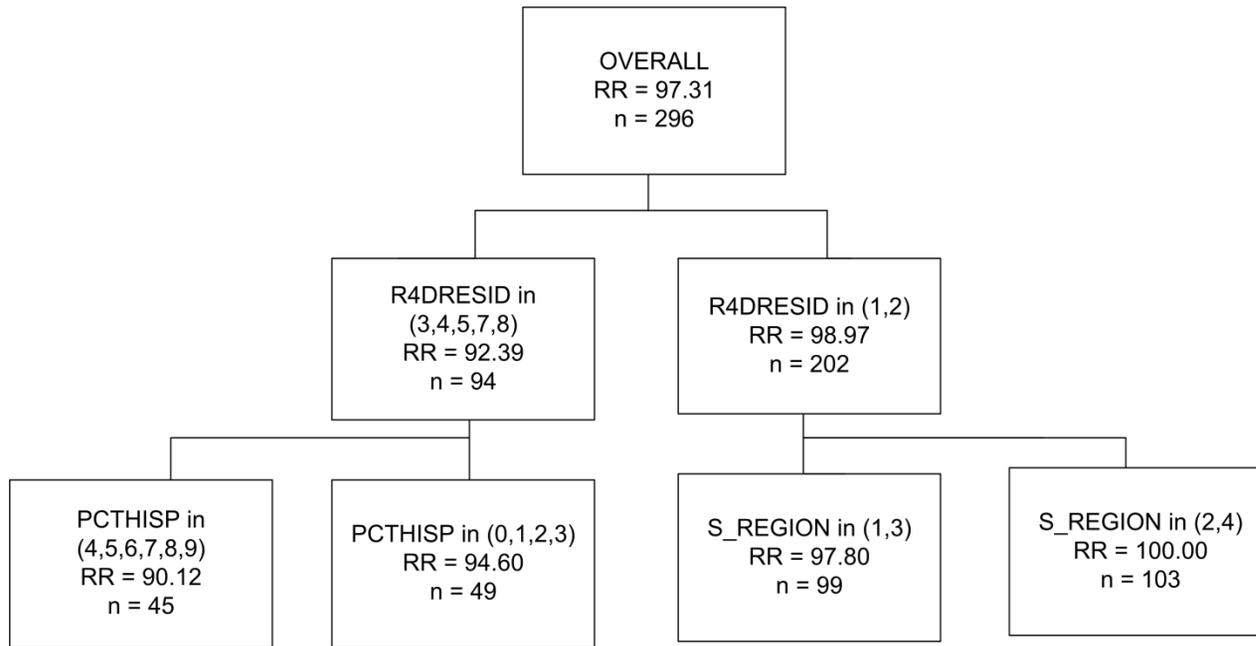
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 9. Round 5 2015 Cohort analytic weight nonresponse adjustment cells – original sample residential care (not nursing home) and nursing home cases with both an SP and FQ interview
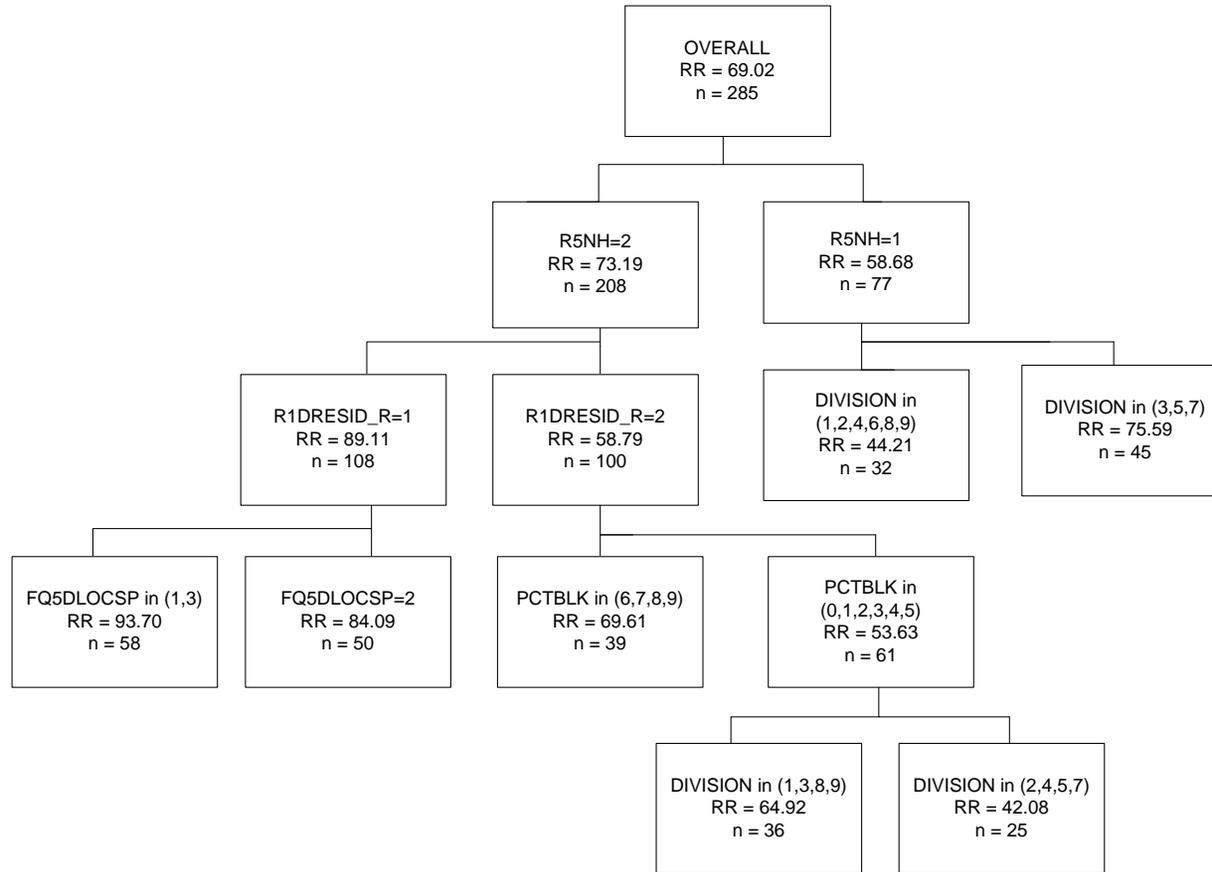
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 10. Round 5 2015 Cohort analytic weight nonresponse adjustment cells – replenishment sample residential care (not nursing home) and nursing home cases with both an SP and FQ interview
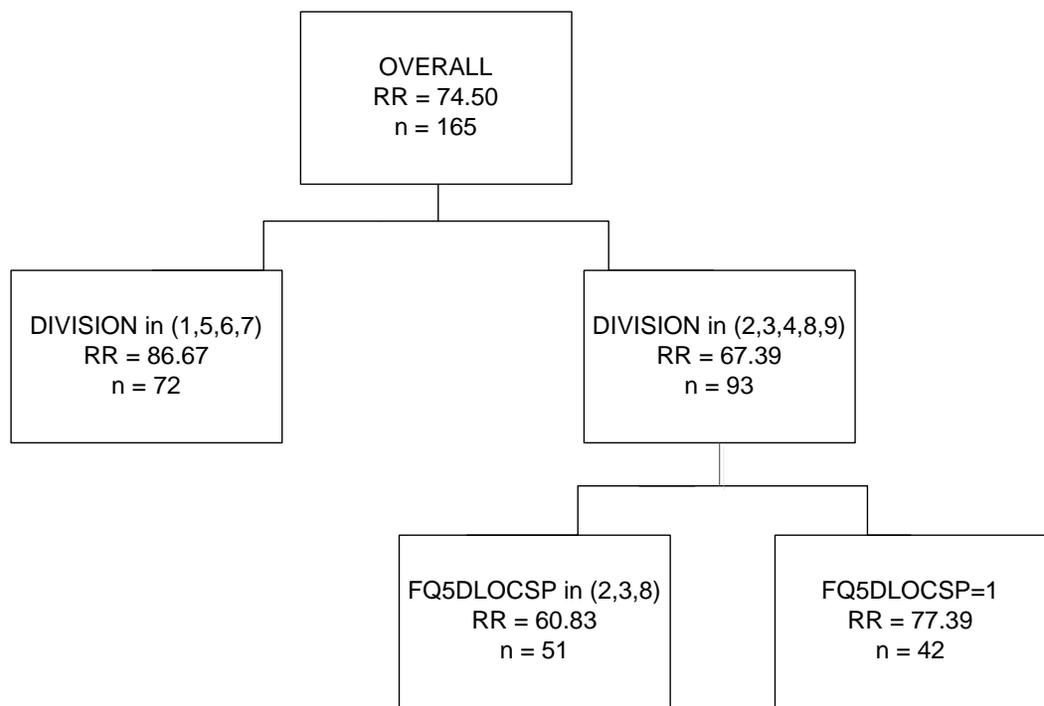
Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell

Figure 11. Round 5 2011 Cohort analytic weight nonresponse adjustment cells –original sample residential care (not nursing home) and nursing home cases with both an SP and FQ interview

Note: "RR" is the weighted response rate for the particular cell, and "n" is the number of respondents in the cell