NHATS Technical Paper #39

NATIONAL HEALTH AND AGING TRENDS STUDY (NHATS)
Development of Round 12 Accelerometry Weights

November 2023

Suggested Citation: Rui Jiao, Vicki A. Freedman, Jennifer Schrack. 2023. National Health and Aging Trends Study Development of Round 12 Accelerometry Weights. NHATS Technical Paper #39. Baltimore: Johns Hopkins Bloomberg School of Public Health. Available at www.NHATS.org. We thank Irene Molina Manrique, who played an instrumental role in the development of the Accelerometry weights and produced several tabulations that appear in this paper. This technical paper was prepared with funding from the National Institute on Aging (U01AG032947).

NHATS Accelerometry Sample and Need for Weights

The baseline Accelerometry Sample was selected in Round 11, and then the respondents were followed to collect physical activity data in Round 12. Data were collected using the Actigraph CentrePoint Insight Watch ("Activity Watch"). The survey weights included with the Round 12 public use file support weighted analysis of Medicare beneficiaries ages 65 and older as of October 1, 2014 who were alive in 2022 and resided in the contiguous United States.

Prior to Round 12 data collection, the 747 Round 11 Accelerometry Sample respondents were flagged to again receive Activity Watches during Round 12. Of those, 25 cases had either moved out of the country or died before the Round 12 interview, and were deemed as ineligible. Of the remaining eligible cases, 639 (89%) completed Part 2 of the SP interview and returned an Activity Watch with usable data.

The survey weights account for differential probabilities of selection and adjust for potential bias related to unit nonresponse to the Round 11 and Round 12 Activity Watch data.

This technical paper describes the development of the Round 12 NHATS Accelerometry SP weights and design variables for variance estimation.

Overview of Weight and Design Variables

The Accelerometry Summary File contains the weights (1 full sample weight, 56 replicate weights) and design variables (stratum, cluster) for making population estimates and proper variance estimation.

File	Full sample weight	Replicate weights	Stratum	Cluster
Accelerometry Summary File : NHATS_Round_12_Accel_Summ_File	w12agfinwgt0	w12agfinwgt1-w12agfinwgt56	w12agvarstrat	w12agvarunit

The weights are designed for generating Sample Person-level estimates. The design variables (stratum and cluster) should be specified when using software that uses Taylor series linearization to estimate the variances of estimates from complex sample surveys.

Replicate weights are also provided and may be used with software that uses replication methods to estimate the variances of estimates from complex sample surveys. The replication approach that was used is the modified balanced repeated replication (BRR) method suggested by Fay (Judkins 1990). Fay's method perturbs the weights by ±100 (1-K) percent where K is referred to as "Fay's factor" or a perturbation factor. The perturbation factor for standard balanced repeated replication (BRR) is K=0 or 100 percent. For NHATS and Accelerometry samples, K = 0.3 was used.

How to Use Sample Weights and Design Variables

Stata Example for Full Weights and Sample Design Variables. In Stata, users should specify the following svyset command to use full sample weights and design variables (Taylor series method) with the summary file.

```
/*Summary file*/
    svyset w12agvarunit [pweight=w12agfinwgt0], strata(w12agvarstrat)
    svy: [stata procedures]
```

Stata Example for Replication Weighting Methods. In Stata, users should specify the following svyset command to use replicate weights with the summary file.

```
/*Summary file*/
    svyset [pweight= w12agfinwgt0], brrweight(w12agfinwgt1-w12agfinwgt56) fay(.3) vce(brr) mse
    svy: [stata procedures]
```

SAS Example for Full Weights and Sample Design Variables. In SAS, users should specify the following command to use full sample weights and design variables (Taylor series method) with the summary file.

```
/*Summary file*/
    [sas survey procedure]
    weight w12agfinwgt0;
    cluster w12agvarunit;
    strata w12agvarstrat;
    [model or other statement]
    run;
```

SAS Example for Replication Weighting Methods In SAS, users should specify the following command when using replicate weights with the summary file.

```
/*Summary file*/
    [sas survey procedure] varmethod=brr (fay=.30);
    weight w12agfinwgt0;
    repweight w12agfinwgt1- w12agfinwgt56;
    [model or other statement]
    run;
```

R Example for Full Weights and Sample Design Variables. In R, users should specify the following command when using full sample weights and design variables (Taylor series method) with the summary file.

```
/*Summary file*/
    library(survey) #need this line only once per session
```

```
nhats.dsgn <- svydesign(id=~w12agvarunit, strata=~w12agvarstrat,
weights=~w12agfinwgt0, data = [data frame name], nest=TRUE)
[model or other statement]
```

R Example for Replication Weighting Methods. In R, users should specify the following command when using replication weights with the summary file.

```
library(survey) #need this line only once per session
nhatsrep<-svrepdesign(weights=~w12agfinwgt0, data=[data frame name], type="Fay",
rho = 0.3, repweights="w12agfinwgt[1-56]+")
[model or other statement]
```

For more information about how to account for sample design in NHATS, please refer to "Accounting for Sample Design in NHATS and NSOC Analyses: Frequently Asked Questions" (Freedman et al. 2020), available at www.NHATS.org.

Calculation of Weights

The Round 12 Accelerometry SP weight began with the Round 11 nonresponse adjusted Accelerometry SP weight, the weight prior to raking, which accounted for differential probabilities of selection of the Accelerometry sample and nonresponse happened in Round 11. See Jiao et al. 2022 for details of Round 11 Accelerometry SP weighting procedures. To produce the Round 12 weight additional adjustments were made: a three-stage nonresponse adjustment and a raking adjustment to the same control totals used in Round 11, estimated by Round 11 Analytic weights.

The 747 baseline Accelerometry sample respondents were classified into three response categories, ineligible due to death or moving out of the contiguous United States by Round 12 interview (n=25), respondent (n=639), and nonrespondent (n=83). The nonresponse happened at different stages during data collection. Of the 83 final nonrespondents, 24 didn't complete either the Sample Person interview (SP) or the Facility Questionnaire (FQ) and they were subject to stage 1 nonresponse adjustment; 2 cases had completed FQ but not SP who were adjusted at stage 2; 57 cases failed to provide valid activity data after completing SP and they were adjusted at stage 3. Table 1 shows the disposition codes map into ineligible, respondent and nonrespondent for all stages.

Table 1. Classification of Round 12 Accelerometry sample for Weight Development Purposes

		Continuing Accelerometry Sample		
Disposition code	**N**	**Classification for Stage 1**	**Classification for Stage 2**	**Classification for Stage 3**
60, 63 Complete SP				
Valid Activity Watch data returned among 60/63	639	Respondent	Respondent[#]	Respondent
Activity Watch data not collected/not returned/not valid among 60/63	57	Respondent	Respondent[#]	Nonrespondent[*]
64 Complete FQ, SP not complete	2	Respondent	Nonrespondent[*]	Nonrespondent
75 Physically/mentally unable to participate, no proxy	0	Nonrespondent[*]	Nonrespondent	Nonrespondent
76 Too ill to participate, no proxy	2	Nonrespondent[*]	Nonrespondent	Nonrespondent
77 Refusal, Sample Person	16	Nonrespondent[*]	Nonrespondent	Nonrespondent
78 Language barrier	0	Nonrespondent[*]	Nonrespondent	Nonrespondent
79 Unable to locate	1	Nonrespondent[*]	Nonrespondent	Nonrespondent
80 Unavailable during field period	2	Nonrespondent[*]	Nonrespondent	Nonrespondent
82 Outside of Primary Sampling Unit	0	Nonrespondent[*]	Nonrespondent	Nonrespondent
85 Refusal, facility	0	Nonrespondent[*]	Nonrespondent	Nonrespondent
87 Refusal, proxy	3	Nonrespondent[*]	Nonrespondent	Nonrespondent
88 Work stopped	0	Nonrespondent[*]	Nonrespondent	Nonrespondent
89 Final other/specify*	0	Nonrespondent[*]	Nonrespondent	Nonrespondent
62, 83, 86 SP deceased, or moved out of contiguous US	25	Ineligible	Ineligible	Ineligible
Total and number assigned weight	747			664

SP=Sample Person interview; FQ=Facility Questionnaire; NH=Nursing home

*: the nonrespondents subject to nonresponse adjustment in its stage

[#]: only a subset of the respondents subject to nonresponse adjustment in its stage

To adjust the stage 1 nonresponse, the potential variables used for creating weighting cells came from the same sources that were used for 2015 Cohort Round 12 Tracker weight nonresponse adjustment (Jiao et al. 2023):

- Beneficiary information from the sampling frame (the 20% HISKEW File for the Round 1 sample and the 20% extract of the Medicare Enrollment Database for the Round 5 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 from the EDB;
- County-level demographic information (e.g., percent of beneficiaries in the county who are Black and percent of beneficiaries in the county who are Hispanic, based on 5% extract of the EDB in 2021; percent of 2021 poverty of all ages in the county, estimated by the Census Bureau) for the county linked to the beneficiary's address from the EDB;
- Census tract-level information based on the 2017-2021 5-year American Community Survey (e.g. tract-level demographic information), based on linkages to the beneficiary's address from the EDB;
- Variables from the NHATS Rounds 1 to 11 interviews (race/ethnicity, highest education, and residential settings)

Appendix Table 1 provides weighted response rates (using the Accelerometry sample Round 11 nonresponse adjusted weights) by response 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 SAS HPSPLIT 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 containing at least 30 respondents. Final nonresponse cells included a total of 6 indicators (designated with "*" in Appendix Table 1). Combinations of these variables created 7 unique nonresponse cells for the nonresponse adjustment (see Appendix Figure 1).

There were only two cases completed FQ but not SP and needed nonresponse adjustment at stage 2. A single nonresponse cell was formed by including the two cases with those completed both FQ and SP.

After completing the SP interview, cases were subject to stage 3 nonresponse adjustment if they (1) did not complete Part 2 of the NHATS Round 12 interview, or (2) completed the Part 2 interview but refused to wear the Activity Watch, or (3) returned the Activity Watch but the data was invalid. To create stage 3 nonresponse cells, we added several Round 12 interview variables (gender, age, residential settings, and mobility level) besides a subset of stage 1 variables that were used for the classification tree analysis. Appendix Table 2 shows variables that were input to the classification tree analysis, along with weighted response rates for each level of each of these variables. Final nonresponse cells included a total of 6 indicators

(designated with "*" in Appendix Table 2). Combinations of these variables created 10 unique nonresponse cells for the nonresponse adjustment (see Appendix Figure 2).

Within each cell at each stage, the input weight for the respondents was inflated by the ratio of the weighted sum of the respondents and nonrespondents to that of the respondents.

Finally, a raking adjustment was applied for the respondents and the ineligible cases so that the weighted marginal totals match the Round 11 NHATS population totals estimated by the Round 11 NHATS analytic weights. The raking adjustment consisted of six dimensions: 5-year age groups, sex, race, region, micro/metropolitan status, and whether Medicare was received before age 65.

The estimated overall design effect due to variation in the Round 11 nonresponse adjusted Accelerometry SP weights was 1.03. After applying Round 12 nonresponse adjustments, the estimated overall design effect due to unequal weighting increased to 1.06. After the raking adjustment, the overall design effect for the final Round 12 Accelerometry SP weights remained at 1.06. No cases were identified as influential outliers, thus no trimming was needed.

REFERENCES

Jiao, Rui, Freedman, Vicki A., Schneider, Benjamin, and Schrack, Jennifer 2023. National Health and Aging Trends Study Development of Round 12 Survey Weights. NHATS Technical Paper #37. Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Rui Jiao, Vicki A. Freedman, Jennifer Schrack. 2022. National Health and Aging Trends Study Development of Round 11 Accelerometry Weights. NHATS Technical Paper #32. Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Freedman, V. A., Hu, M., DeMatteis, J., & Kasper, J. D. 2020. Accounting for sample design in NHATS and NSOC analyses: Frequently asked questions. NHATS Technical Paper #23. Baltimore: Johns Hopkins University School of Public Health. Available at www.NHATS.org.

Appendix Table 1. Weighted Responses Rates for Variables used in Stage 1 Nonresponse Adjustment - No SP/FQ Interview

Variable & Values	Weighted Response Rate
OVERALL	96.5%
BENEFICIARY INDICATORS	
Age[1] * (H_AGECAT_R5)	
1: 65-69	97.0%
2: 70-74	97.2%
3: 75-79	95.2%
4: 80-84	94.7%
5: 85- 89	96.0%
6: 90+	100.0%
Gender[1] (H_SEX)	
1: Male	97.4%
2: Female	95.8%
Census Region[2] (S_REGION)	
1: Northeast	98.2%
2: Midwest	95.3%
3: South	97.2%
4: West	95.1%
Census Division[2]* (DIVISION)	
1: New England	100.0%
2: Middle Atlantic	97.5%
3: East North Central	96.4%
4: West North Central	94.0%
5: South Atlantic	96.3%
6: East South Central	100.0%
7: West South Central	97.5%
8: Mountain	100.0%
9: Pacific	94.2%
Census Metro/Micro Area Designation (2020) [2] (S_METMICRO)	
1: Metropolitan area	96.7%
2: Micropolitan area	93.5%
3: Non-metro	100.0%
Health Maintenance Organization Beneficiary[1] (HMOTYPE)	
0: Yes	97.4%
9: No	96.1%
Age First Enrolled in Medicare[1] (ENROLL_AGE)	
1: Prior to age 65	98.3%
2: At or after age 65	96.3%
R1/R5 RACE ETHNICITY[4] * (RL5DRACEHISP_R)	
1: White, non-Hispanic	97.7%
2: Black, non-Hispanic	98.4%
3: Other, non-Hispanic	85.8%
4: Hispanic	88.3%
5: DK/RF	94.8%
R1/R5 HIGHEST EDUCATION [4] (EL5HIGSTSCHL_R2)	
1: Below high school	96.1%
2: High school	96.7%
3: Above high school	96.6%

Variable & Values	Weighted Response Rate
TRACT-LEVEL INDICATORS (Quartiles)	
Household Income[3] (C_AGG_HH_INC)	
1: 1st quartile	95.3%
2: 2nd quartile	96.0%
3: 3rd quartile	97.1%
4: 4th quartile	97.8%
Median Household Income[3] (C_MED_HH_INC)	
1: 1st quartile	96.7%
2: 2nd quartile	96.0%
3: 3rd quartile	96.2%
4: 4th quartile	97.1%
Median Household Income 65+[3] (C_MED_HH_INC_65)	
1: 1st quartile	96.0%
2: 2nd quartile	97.7%
3: 3rd quartile	96.4%
4: 4th quartile	97.3%
9: Missing	92.2%
% Households with Adult 65+[3] (C_PCT_HH_65)	
1: 1st quartile	93.7%
2: 2nd quartile	98.3%
3: 3rd quartile	95.5%
4: 4th quartile	97.4%
% Households in Poverty[3] (C_PCT_HH_POV)	
1: 1st quartile	97.4%
2: 2nd quartile	96.8%
3: 3rd quartile	94.6%
4: 4th quartile	97.2%
% Households Reporting Public Assistance[3] (C_PCT_HH_PUBASST)	
1: 1st quartile	97.4%
2: 2nd quartile	97.7%
3: 3rd quartile	96.2%
4: 4th quartile	94.5%
% Households Reporting Retirement Income[3] (C_PCT_HH_RETIREINC)	
1: 1st quartile	94.8%
2: 2nd quartile	95.6%
3: 3rd quartile	97.2%
4: 4th quartile	97.5%
% Households Reporting Social Security[3] (C_PCT_HH_SOCSEC)	
1: 1st quartile	94.8%
2: 2nd quartile	95.8%
3: 3rd quartile	97.9%
4: 4th quartile	96.7%

Variable & Values		Weighted Response Rate
COUNTY LEVEL INDICATORS		
% Black 65+ (deciles)[2]	(PCTBLK_N)	
0: 1st decile		97.3%
1: 2nd decile		94.2%
2: 3rd decile		96.6%
3: 4th decile		99.1%
4: 5th decile		95.7%
5: 6th decile		95.6%
6: 7th decile		97.2%
7: 8th decile		97.2%
8: 9th decile		96.0%
9: 10th decile		94.6%
% Hispanic 65+ (deciles)[2]	(PCTHISP_N)	
0: 1st decile		100.0%
1: 2nd decile		94.7%
2: 3rd decile		99.0%
3: 4th decile		97.8%
4: 5th decile		96.3%
5: 6th decile		97.3%
6: 7th decile		98.7%
7: 8th decile		97.9%
8: 9th decile		91.2%
9: 10th decile		90.3%
% Poverty (deciles)[2] *	(PCTPOV_N)	
0:1st decile		97.2%
1: 2nd decile		97.1%
2: 3rd decile		94.8%
3: 4th decile		100.0%
4: 5th decile		97.0%
5: 6th decile		94.3%
6: 7th decile		93.6%
7: 8th decile		98.0%
8:9th decile		97.2%
9: 10th decile		93.2%
OTHER INDICATORS		
R11 RESIDENTIAL CARE STATUS[4]	(R11DRESID)	
1: Community		96.7%
2: Residential Care Resident not nursing home (SP interview complete)		92.8%
4: Nursing home (SP interview complete)		100.0%

Variable & Values		Weighted Response Rate
TRACT-LEVEL INDICATORS (Quartiles)		
% Households Reporting SSI[3]	(C_PCT_HH_SSS)	
1: 1st quartile		97.9%
2: 2nd quartile		94.9%
3: 3rd quartile		96.1%
4: 4th quartile		97.8%
% Households Owning Their Home[3]	(C_PCT_OWNHOME)	
1: 1st quartile		91.9%
2: 2nd quartile		95.3%
3: 3rd quartile		98.5%
4: 4th quartile		98.5%
% Households 65+ Owning Their Home[3]	(C_PCT_OWNHOME_65)	
1: 1st quartile		92.6%
2: 2nd quartile		96.9%
3: 3rd quartile		98.4%
4: 4th quartile		97.4%
% Households 65+ Below Poverty[3] *	(C_PCT_POV_65)	
1: 1st quartile		97.0%
2: 2nd quartile		97.7%
3: 3rd quartile		94.9%
4: 4th quartile		96.5%
Per Capita Income[3] *	(C_PER_CAP_INC)	
1: 1st quartile		94.1%
2: 2nd quartile		96.2%
3: 3rd quartile		98.4%
4: 4th quartile		96.6%

[1]Based on information from either the September 30, 2010 CMS 20% Health Insurance Skeleton Eligibility Write Off (HISKEW) file if the case is in the Round 1 sample, or the September 30, 2014 CMS 20% Enrollment Database (EDB) extract if the case is in the Round 5 replenishment sample.

[2]Based on county-level information from the September 30, 2021 CMS 5% EDB extract linked to the beneficiary's EDB address.

[3]Based on tract-level information from the 2017-2021 5-year American Community Survey file linked to the beneficiary's EDB address.

[4]Based on responses to items in the Rounds 1 to 11 interviews.

* Retained in classification tree analysis for nonresponse adjustment.

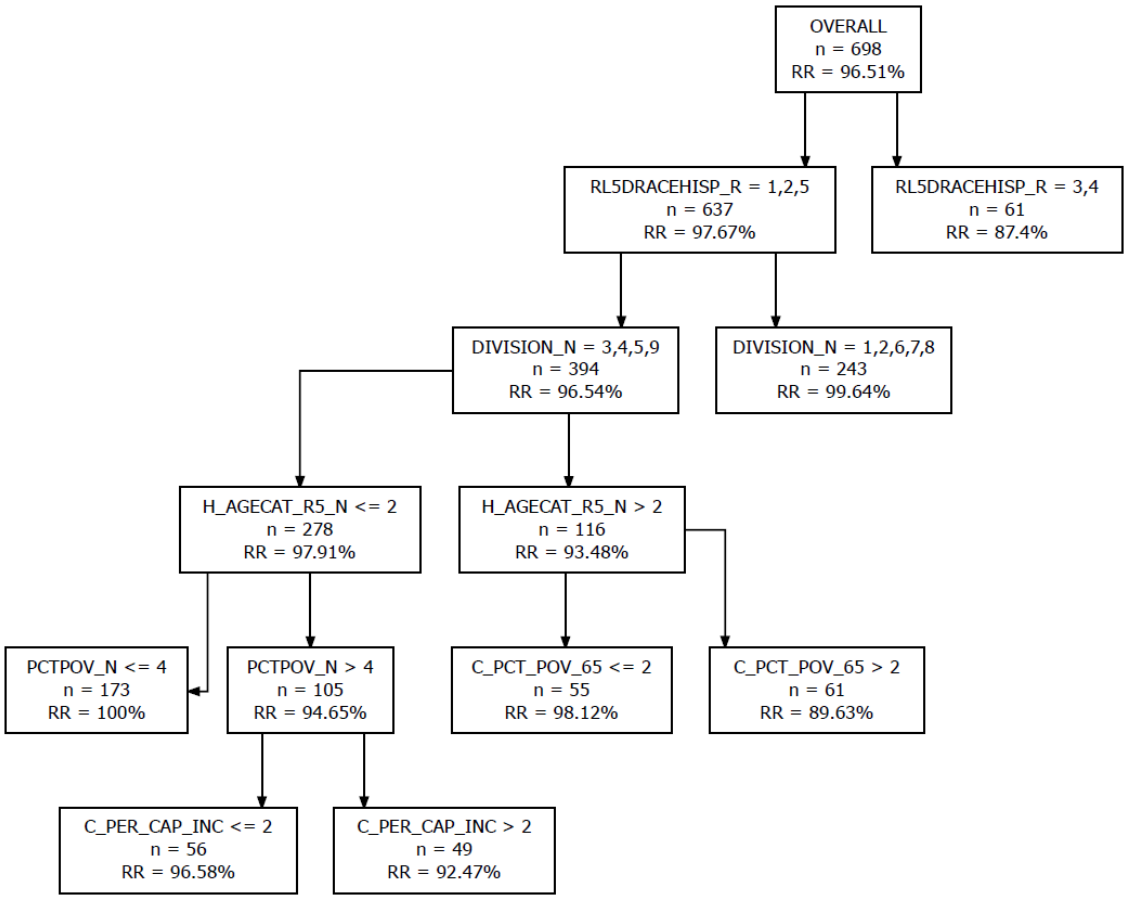
Variable names used in classification trees shown in parentheses

Appendix Table 2. Weighted Responses Rates for Variables used in Stage 3 Nonresponse Adjustment - No Actigraph Data after SP Interview

Variable & Values	Weighted Response Rate	Variable & Values	Weighted Response Rate
OVERALL	91.7%		
BENEFICIARY INDICATORS			
Census Division[1] * (DIVISION)		**R12 SP Age at interview**[2] * (R12D2INTVRAGE)	
1: New England	100.0%	2: 70-74	93.3%
2: Middle Atlantic	92.9%	3: 75-79	96.1%
3: East North Central	92.3%	4: 80-85	85.5%
4: West North Central	86.4%	5: 86-89	86.1%
5: South Atlantic	92.1%	6: 90+	93.7%
6: East South Central	95.5%		
7: West South Central	91.6%	**R12 SP gender**[2] * (R12DGENDER)	
8: Mountain	88.3%	1: Male	90.7%
9: Pacific	90.6%	2: Female	92.6%
Census Metro/Micro Area Designation (2020) [1] (S_METMICRO)		**R12 SP mobility**[2] * (MO12OUTOFT)	
1: Metropolitan area	91.4%	1: Every day (7 days a week)	94.9%
2: Micropolitan area	92.9%	2: Most days (5-6 days a week)	92.9%
3: Non-metro	94.9%	3: Some days (2-4 days a week)	84.1%
R1/R5 RACE ETHNICITY[2] * (RL5DRACEHISP_R)		4: Rarely (once a week)	87.1%
1: White, non-Hispanic	93.2%	5: Never	67.6%
2: Black, non-Hispanic	88.7%		
3: Other, non-Hispanic	86.4%	**R12 SP Residence**[2] (R12DRESID)	
4: Hispanic	82.2%	1: Community	92.4%
5: DK/RF	93.1%	2: Residential Care Resident not nursing home (SP interview complete)	93.7%
R1/R5 HIGHEST EDUCATION[2] * (EL5HIGSTSCHL_R2)		4: Nursing home (SP interview complete)	37.9%
1: Below high school	84.6%		
2: High school	91.1%		
3: Above high school	93.5%		

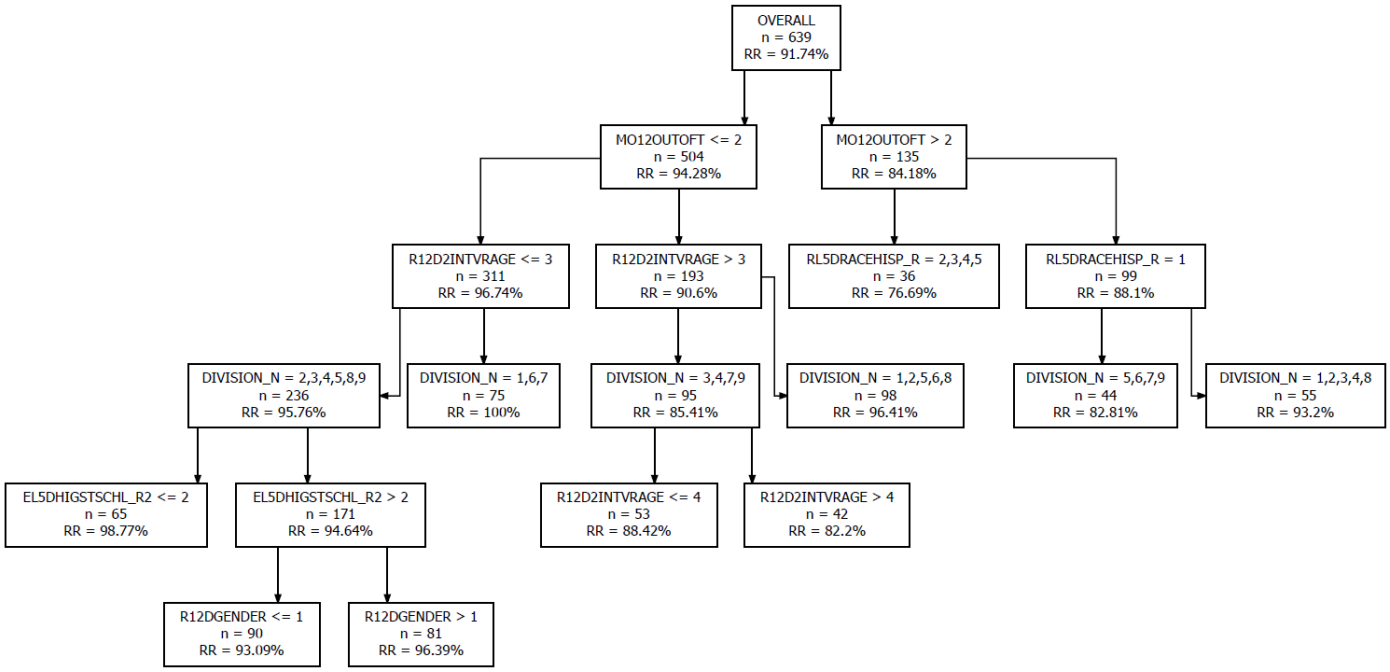
[1]Based on county-level information from the September 30, 2021 CMS 5% EDB extract linked to the beneficiary's EDB address.
[2]Based on responses to items in the Rounds 1 to 12 interviews.
* Retained in classification tree analysis for nonresponse adjustment.
Variable names used in classification trees shown in parentheses

Appendix Figure 1. Accelerometry SP Weight stage 1 nonresponse adjustment cells



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

Appendix Figure 2. Accelerometry SP Weight stage 3 nonresponse adjustment cells



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