



Weighting Procedures and Bias Assessment for the 2020 National Health Interview Survey

Matthew D. Bramlett, James M. Dahlhamer
and Jonaki Bose

Division of Health Interview Statistics
National Center for Health Statistics
Hyattsville, Maryland

Centers for Disease Control and Prevention
U.S. Department of Health and Human Services

September 2021

Abstract

Due to the COVID-19 pandemic, a portion of the 2020 sample for the National Health Interview Survey (NHIS) was replaced with a reinterview component consisting of sample adults from the 2019 NHIS. This report describes the weighting procedures that were used to produce three weights for sample adults: a) a weight for just the reinterviewed cases, representing the 2019 population, for longitudinal analysis; b) a weight for just the remaining 2020 sample cases (partial sample), representing the 2020 population, for specialized analysis; and c) a weight for the combined file (including both the reinterviewed cases and the remaining 2020 sample cases), representing the 2020 population, for analysis of 2020 sample adults. Since the reinterview component did not include sample children from the 2019 NHIS, only one child weight was created for 2020 using a method similar to the one used in 2019. A bias assessment of survey estimates based on these weights was also performed. Results reveal that:

- Implementation of weighting adjustments among reinterview cases led to an 80% reduction in observed bias across 78 estimates.
- The combined file retains some biases after weighting adjustments, notably an underrepresentation of adults living alone and those in the lowest income category, and an overrepresentation of adults living in households with both landline and cell telephones.
- Use of the partial sample should largely produce the same estimates as the combined sample (within sampling error): A comparison of estimates based on the partial sample weight to estimates based on the combined weight revealed only 2 (out of 78) significant differences.

Suggested citation: Bramlett MD, Dahlhamer JM and Bose J. Weighting procedures and bias assessment for the 2020 National Health Interview Survey. Published September 2021.

Weighting Procedures

I. Introduction

The National Health Interview Survey (NHIS) has been conducting interviews as part of ongoing data collection activities continuously since 1957 (1). Periodically, changes are made to data collection procedures to take advantage of advances in the field. The NHIS fielded a newly redesigned questionnaire in 2019, and the 2019 weighting procedures were updated as well. The new procedures use multilevel logistic regression models for predicting household, sample adult, and sample child response. The models utilize variables from multiple auxiliary data sources and the predicted probabilities or response propensities from these models are used for nonresponse adjustments to the household, sample adult, and sample child weights. In addition, raking procedures were introduced with the sample adult and sample child weights for calibration to population control totals. Raking is an iterative process that adjusts weights to match auxiliary data sources one variable at a time and is repeated multiple times until the distribution from the survey converges on (or closely matches) the distribution in the population. Raking allows for the inclusion of more calibration variables than was possible with post-stratification in the past. The new weighting procedures are described in more detail in the 2019 public-use file documentation (2).

When the new procedures were implemented for the 2019 sampling weights, a multiyear plan was designed to periodically update the set of covariates included in the response propensity models and evaluate whether any changes were needed in the calibration dimensions or adjustment capping levels. (Capping is a process that truncates weighting adjustments beyond a certain threshold to limit the production of extreme weights that will impact precision.) The plan was to keep the previous year's model and other parameters consistent while producing the current year's weights for the NHIS Early Release (ER) program, and then at the end of data collection for each calendar year, update the model covariate selection and other weighting adjustment parameters while producing the annual weights for that year. The COVID-19 pandemic

introduced several complications that quickly made this multiyear plan, newly enacted in 2019, inoperable for 2020.

Due to data collection difficulties posed by the COVID-19 pandemic, the NHIS shifted in quarter 2 (Q2) of 2020 from in-person to all-telephone interviewing. In Q3 and Q4, data collection included some in-person interviewing but was still predominantly telephone-based. Due to concerns about possible loss of coverage and lower response rates in telephone interviewing, and also to assess differences in health outcomes before and during the pandemic, approximately half of the usual sample for the last 5 months of 2020 was replaced with a longitudinal component where a subset of the 2019 sample adults were reinterviewed over the telephone using the 2020 NHIS questionnaire. The reinterview component did not include sample children from the 2019 NHIS.

The 2020 NHIS data collection effort thus effectively consisted of four separate designs: primarily in-person (Q1); telephone only (Q2); telephone attempts first with some in-person where feasible (Q3 & Q4, 2020 sample); and telephone reinterview (Q3 & Q4, 2019 sample). Combining these four components into one annual sample weighted to estimate the 2020 adult population presented unique benefits and challenges. This report describes the weighting procedures for the 2020 NHIS sample adult and sample child files and the analysis of bias reduction due to weighting adjustments.

II. Background

Typically, each year's NHIS public data release consists of an annual file that is created by concatenating four quarterly files, each of which is nationally representative on its own. Data collection in 2020 resulted in the following components, each of which can be weighted to represent the national population alone or in conjunction:

1. Q1: a normal quarter conducted primarily in-person as in previous years, using the 2020 sample;
2. Q2: a quarter conducted by telephone, using the 2020 sample;

3. Q3 truncated: 2020 sample from July, plus approximately half of the 2020 sample from August and September, conducted via telephone with some in-person interviewing in select parts of the country;
4. Q4 halved: 2020 sample from half of original Q4, conducted similarly to Q3 ; and
5. The reinterview sample: approximately half of the 2019 Q1-Q3 and all of the 2019 Q4 sample adult respondent cases, reinterviewed by telephone only in August-December 2020 using the 2020 questionnaire.

Note that reinterview cases only include sample adults. There was no reinterview for sample children, so the sample child weighting is restricted to regular 2020 sample cases only and will follow the typical procedures (2). The only differences between the 2020 and 2019 sample child weighting procedures are a) the 2020 sample child sample size is smaller than usual, because the 2020 sample was reduced for 5 of 12 months; and b) adjustments were made to the weighting procedures to account for the loss of non-telephone households and households for which a telephone number could not be obtained, and the different paradata (i.e., auxiliary data measuring aspects of the interviewing process) available for response propensity modeling for telephone versus in-person households. This weighting approach was also implemented for the Q2 Early Release estimates, as described in the documentation for that release (3). Therefore, the remainder of this document focuses on sample adult weighting.

Each NHIS quarterly sample is divided into four separate panels, each of which is nationally representative on its own. The Agency for Healthcare Quality and Research's Medical Expenditures Panel Survey (MEPS) uses two of four NHIS panels from the previous year as its sample frame. If NCHS conducted the reinterview for all 2019 cases, then the MEPS sample would also be receiving the 2020 reinterview. Thus, to avoid overburdening 2019 respondents with both an NHIS reinterview and a MEPS interview in 2020, the non-MEPS panels were selected for the NHIS reinterview component. To provide needed resources for the data collection effort for the reinterview cases, the regular 2020 sample was halved. Again, this was accomplished by including two of four panels (subsampling in) and dropping two (subsampling out). Thus, a more detailed breakdown of the 2020 sample follows:

Regular 2020 sample:

- 4 panels from Q1;
- 4 panels from Q2;
- 4 panels from July;
- 2 subsampled panels from August and September; and
- 2 subsampled panels from Q4.

Reinterview sample (originally interviewed in 2019):

- 2 subsampled panels of Q1/2019 reinterviewed in August 2020;
- 2 subsampled panels of Q2/2019 reinterviewed in September 2020;
- 2 subsampled panels of Q3/2019 reinterviewed in October 2020; and
- 4 panels of Q4/2019 reinterviewed in November-December 2020.

There are three sample adult weights that have been produced. Since the reinterview cases are nationally representative and have both a 2019 and 2020 interview, they may be used separately, as a stand-alone longitudinal file to analyze within-person changes over time. A separate weight for reinterview cases only, adjusted to represent the original 2019 population, was produced to enable such analyses. In addition, since the reinterview cases received the 2020 questionnaire for the 2020 reinterview, they can be combined with the regular 2020 sample to form the 2020 NHIS sample adult file, using one common set of sampling weights to represent the 2020 population of noninstitutionalized civilian adults in the United States. Finally, because some analyses may be negatively impacted by including the reinterview cases in the 2020 file, including the dependence introduced from the reinterview adults appearing in both the 2019 and 2020 sample adult files, a partial weight restricted to cases that were sampled in 2020 was also produced.

The next four sections cover adjustment steps used in production of the three sample adult weights and the final 2020 sample child weight. Development of the reinterview sample longitudinal weight is discussed first. While most analysts will use the final sample adult weight to produce 2020 estimates, development of the weighting procedures began with the production of the longitudinal weight. The new procedures were developed for the longitudinal weight and then adapted for the final sample adult weight, so it makes sense to describe them in that order.

III. The Reinterview Sample Longitudinal Weight

The longitudinal weight is designed for intra-individual time-1 (2019)/time-2 (2020) analysis. The goal in weighting the reinterview cases for longitudinal analysis is to make the reinterview respondents (n=10,415) represent the original 2019 sample adult population, so that time-1 and time-2 estimates can be compared in a comparable manner and observed differences can be attributed to changes over time in the health and well-being of the sample cases rather than changes in the distribution of the population over time. Thus, the starting point, or base weight, is the final 2019 sample adult weight.

Step 1: Base weight

The base weight is equal to the 2019 sample adult weight if the household was subsampled in for 2020 telephone reinterview and equals zero otherwise.

The base weight has already been adjusted for 2019 nonresponse and calibrated to the 2019 population. However, since 2020 reinterview respondents comprise approximately 1/3 of the original 2019 sample adults due to subsampling, noncontact, and nonresponse, the 2019 sample adult weight can no longer be considered representative and must be adjusted for 2020 nonreinterview. A 2019 sample adult can be classified as a 2020 nonreinterview at each of the following three stages:

1. Subsampling out by panel
2. Nonrecontact (including adults who were not successfully recontacted and adults who were excluded from the reinterview study due to incomplete contact information)
3. Refusal to participate (nonresponse).

Adjustments were made to the base weights to account for each of these stages, and the resulting weights were calibrated to 2019 population control totals. The specific steps are described below, while the overall scheme is presented in Flowchart 1 (see appendix).

Due to subsampling, nonrecontact, and nonresponse, the size of the reinterview respondent group is approximately one-third that of the original 2019 sample adult file. A reinterview rate that low carries with it the potential for considerable bias in estimates, if estimates

are related to panel inclusion, recontact likelihood, or response propensity. However, a major advantage of the reinterview design is the availability of information about nonrespondents that typically is not available when adjusting weights to account for nonresponse bias. Specifically, the complete data from the 2019 sample adult interview is available for both 2020 reinterview respondents and nonrespondents. Thus, it is possible to identify the specific 2019 characteristics that are associated with subsequent nonreinterview, and to use those covariates to form adjustment cells for weighting adjustments. This technique has been demonstrated to reduce or eliminate biases in reinterview samples (4-6).

Because of the very large number of potential predictors of response propensity, recursive partitioning models (RPM) in the software package R were used to model nonreinterview propensity. RPM produces a set of nodes by first splitting the sample by the strongest predictor of response, and then within each of the resulting subgroups, splitting the subgroup by the strongest predictor, continuing this process until no predictor meets the significance criteria for further splitting, ending up with a tree of branches, each of which ends in a terminal node. The model selects the best categorizations for covariates; for example, if race/ethnicity is included in the set of covariates with 4 categories (Hispanic, non-Hispanic white, non-Hispanic black and non-Hispanic other race) the model may split that covariate into two categories (non-Hispanic black versus all others, e.g.) in creating the nodal splits, because that is where the differentiation on response is concentrated. The model also selects the relevant interaction effects among combinations of variables, without having to prespecify the interaction terms. This is particularly useful when dealing with a large set of potential covariates. The primary benefit of this approach, however, is that it results in a set of mutually exclusive and exhaustive categories (terminal nodes) which are defined by the predictors of nonresponse and can be used to form adjustment cells for weighting adjustments.

The set of covariates used as potential predictors of nonreinterview overall and its various stages was chosen to best leverage the full survey data that was available for both respondents and nonrespondents (since every case had completed the 2019 interview). The set of covariates included demographics (age, sex, race/ethnicity, marital status, nativity, household size, family

structure), socioeconomics (income, education, employment status, housing tenure), geographic indicators (region, metropolitan statistical area (MSA) status), telephone status (cell/landline/both/neither), and multiple health and healthcare measures from each of five main domains: health status, health insurance coverage, healthcare access, healthcare utilization, and health behaviors.

RPM was used to separately model overall nonreinterview (comparing reinterviewed cases to nonreinterviewed cases among the entire 2019 sample adult file) as well as the recontact and response stages: nonrecontact (comparing those contacted and those not contacted, among those subsampled for the reinterview from the 2019 sample adult file); and refusal (comparing respondents and nonrespondents among those contacted). The same set of potential covariates was used for models of nonreinterview, nonrecontact, and refusal and is shown in Table 1.

Theoretically, because each of the separate panels was drawn to be nationally representative, subsampling entire panels in or out of the sample should not introduce bias. Thus, only a simple ratio adjustment, applied equally to all subsampled-in cases to inflate the base weights to sum to the appropriate total (including the population counts previously allocated to the cases that were subsampled out) should be necessary at the subsampling stage. However, because the 2019 sample adult weights were not originally calibrated to population control totals within each panel, it is possible that 2019 nonresponse could have been associated with panel assignment in some way, resulting in subsampling bias when cases were selected for reinterviewing. The preliminary bivariate analysis of differences due to subsampling showed that the reinterview subsample of the 2019 sample adult file did not significantly differ from the rest of the 2019 sample adult file on any sociodemographic attribute, and only one significant difference at the 0.05 alpha level and two significant differences at the 0.10 alpha level were found among the health and healthcare indicators (results not shown).

Comparisons of 25 health and healthcare variables (29 indicators accounting for extra categories of insurance coverage and body mass index) should expect a false positive (at the 0.05 level) or two (at the (0.10 level), suggesting that no bias was introduced at the subsampling stage; therefore, a simple ratio adjustment was used.

Step 2: Subsampling adjustment

Cases that were subsampled in for reinterviewing received a ratio adjustment to the base weight: (the sum of the base weights for the full 2019 sample adult file) / (the sum of the base weights for the subsampled-in reinterview sample cases).

Among subsampled-in cases, some households could not be recontacted. The RPM recontact model resulted in 28 nodes (ranging from 290 to 2,895 cases and with recontact rates ranging from 66.7% to 83.9%). The primary split was by housing tenure with subsequent splits by marital status, years at current residence, used prescription medications, family income as a percentage of the federal poverty level, eye exam, education, age, region, metropolitan statistical area (MSA) status, and arthritis. The recontacted cases were adjusted to account for the nonrecontacts in step 3.

Step 3: Nonrecontact adjustment

Cases that were recontacted received a ratio adjustment to the subsample-adjusted weight: $1 /$ (the recontact propensity in the RPM recontact model terminal node).

Among recontacted cases, some of the adults did not participate, i.e., were nonrespondents. The RPM response model resulted in 22 nodes (ranging from 207 to 5,900 cases and with participation rates ranging from 41.5% to 70.9%). The primary split was by education with subsequent splits by age, years at current residence, used prescription medications, MSA status, household size, body mass index (BMI), telephone status, ever diagnosed with high cholesterol, eye exam, flu vaccination, dental cleaning, housing tenure, family income as a percentage of the federal

poverty level and overall excellent/very good health. The respondent cases were adjusted to account for the nonrespondents.

Step 4: Nonresponse adjustment

Cases that were reinterviewed received a ratio adjustment to the nonrecontact-adjusted weight: $1 /$ (the response propensity in the RPM response model terminal node).

Finally, the weights were calibrated to the original 2019 control totals. Rather than raking the weights to the external population control totals that the original 2019 sample adult weights were raked to, internal control totals (estimated from the 2019 sample adult file) were used to enable the inclusion of raking dimensions which are not typically available for the population.

The overall nonreinterview RPM had 27 nodes (ranging from 189 to 2,452 cases and with reinterview rates ranging from 26.0% to 65.6%). The primary split was by education, with subsequent splits by age, race, marital status, housing tenure, household size, telephone status, arthritis, wearing glasses, health insurance coverage type, prescription medication usage, eye exam, dental cleaning, and current cigarette smoking. The 27 nodes were used to create a single variable with 27 categories.

Step 5: Final Calibration

For cases that were reinterviewed, the nonresponse-adjusted weight was raked to internal control totals based on the 2019 sample adult raking dimensions (age, sex, race/ethnicity, education, MSA status and Census division) but replacing education with the 27-category node variable (because education was the primary split in forming the nodes).

The resulting weight is the final weight for longitudinal analysis of the reinterview cases: WTSA_L.

Note that every domain except access to health care contributed at least one variable to either the recontact or response adjustments or overall nonreinterview calibration steps (Table 1).

This is more specific adjustment than is typically possible and targeted directly at increasing (or decreasing) the estimated prevalence of the attributes that are under-(over-)represented in the respondent cases.

IV. The 2020 Sample Adult Weight

The 2020 sample adult file is composed of both the reinterview cases and the 2020-sampled cases (n=31,568). To continue to take advantage of the rich information available for nonreinterviewed cases, the weighting procedures for the combined annual weight consist of parallel processing streams for the reinterview cases and the regular 2020 sample cases, before the streams converge to treat the separate samples together. This is shown in Flowchart 2.

The 2020 sample cases (the right-hand stream in Flowchart 2) were processed according to the Q2/2020 procedures which were modified to account for pandemic-related changes to data collection procedures. With the shift to telephone interviewing in Q2, some modifications to the typical procedures were necessary to account for the loss of some contact history and neighborhood observation paradata in the nonresponse model, and to account for the loss of coverage of nontelephone households. The solution for the 2020 Q2 ER release was to drop the affected covariates from the nonresponse propensity model and to add housing tenure to the raking controls. This is detailed in the bias analysis report issued with the ER release (3).

Both weighting streams begin with the base weights. Unlike the longitudinal weight, which began with the final 2019 sample adult weight, the base weights for the combined reinterview and 2020 sample needed to be generated from population household counts, apportioned to the subsampled-in panels of the 2020 sample and the reinterview sample such that the base weights of the combined sample sum to the appropriate population counts within sampling strata.

Base weights are in effect an apportionment of population counts, within sampling strata, to sample cases according to their probabilities of selection; the new combined base weights simply apportion the 2020 population counts to the combined 2019 and 2020 sample cases rather

than just the 2020 sample cases. This effectively spreads the original 2020 base weights out to cover the 2019 sample as well, except for the subsampled-out panels.

Step 1: Household-level base weights

Base weights were generated for combined reinterview and 2020 sample cases by allocating population household counts according to the probability of selection within strata.

At this point the weighting streams diverge. The steps for the reinterview stream (left side of Flowchart 2) are described first.

The combined base weight for reinterview cases was adjusted for household-level nonresponse in 2019 by predicting response propensities, grouping response propensities into quintiles, and adjusting within propensity quintiles by the inverse of the median response propensity of the quintile. The 2019 multilevel logistic regression of household response was used to predict response propensities. Details of the 2019 response models were documented with the 2019 public data use release (2).

Step 2: Household-level 2019 nonresponse adjustment for reinterview sample cases

Reinterview sample cases received a ratio adjustment to the combined base weight: $1 / (\text{the median 2019 household response propensity in the response propensity quintile})$.

The next step was to transform the 2019-nonresponse-adjusted household-level combined base weights for the reinterview cases to person-level weights by the inverse of the probability of selection among all the adults within the household.

Step 3: Adjust reinterview sample cases for sample adult probability of selection in 2019

Reinterview sample cases received a ratio adjustment to the household-level nonresponse-adjusted weight: $1 /$ (the sample adult probability of selection within the household in 2019).

The person-level weights adjusted for within-household selection probabilities were then adjusted for nonresponse to the 2019 sample adult interview. As in 2019, adjustments were within quintiles formed from predicted response propensities generated from a multilevel logistic regression model.

Step 4: Sample adult 2019 nonresponse adjustment for reinterview sample cases

Reinterview sample adult cases received a ratio adjustment to the household-selection-adjusted weight: $1 /$ (the median 2019 sample adult interview response propensity in the response propensity quintile).

The next three steps for the reinterview cases are largely a repeat of the steps described above for the longitudinal weight, including raking to 2019 internal control totals to maximize the utility of the 2019 data available for both reinterviewed and nonreinterviewed cases. As described above (in step 3 of the longitudinal weight procedures), step 5 uses the RPM results from a model of recontact to adjust the weights of the reinterview cases for nonrecontact.

Step 5: Nonrecontact adjustment for reinterview cases

Reinterview cases that were recontacted received a ratio adjustment to the 2019-nonresponse-adjusted weight: $1 /$ (the recontact propensity in the RPM recontact model terminal node).

As described in step 4 of the longitudinal weight procedures, step 6 uses the RPM results from a model of response to adjust the weights of the reinterview cases for nonresponse.

Step 6: 2020 Nonresponse adjustment for reinterview cases

Reinterview cases that were reinterviewed received a ratio adjustment to the nonrecontact-adjusted weight: $1 / (\text{the response propensity in the RPM response model terminal node})$.

As described in step 5 of the longitudinal weight procedures, step 7 calibrates the weights for the reinterview cases to 2019 population control totals. The usual raking dimensions are used, except that education is replaced with a variable derived from the RPM results from a model of overall nonreinterview.

Step 7: Calibration to 2019 totals for reinterview cases

For reinterview cases for whom a 2020 reinterview was completed, the 2020-nonresponse-adjusted weight was raked to 2019 internal totals by age, sex, race/ethnicity, MSA status, Census division, and the 27-category variable derived from the RPM overall nonreinterview model terminal nodes.

The 2020 sample stream (the right side of Flowchart 2) uses the same procedures as were used for the 2019 sample weights (2). The first step in that stream (step 8) is to estimate the usual multilevel logistic regression model of response at the household level and adjust for nonresponse within response propensity classes.

Step 8: 2020 Household-level nonresponse adjustment for 2020 sample cases

2020 sample cases received a ratio adjustment to the base weight: $1 /$ (the median household response propensity in the response propensity quintile).

The household level weights were then transformed into person-level weights for adults and the nonresponse model was estimated at the person level.

Step 9: Adjust 2020 sample cases for sample adult probability of selection

2020 sample cases received a ratio adjustment to the household-level nonresponse-adjusted weight: $1 /$ (the sample adult probability of selection within the household).

The person-level weights adjusted for within household selection probabilities were then adjusted for nonresponse to the 2020 sample adult interview.

Step 10: Sample adult nonresponse adjustment for 2020 sample cases

2020 sample adult cases received a ratio adjustment to the household-selection-adjusted weight: $1 /$ (the median 2020 sample adult interview response propensity in the response propensity quintile).

In preparation for combining two sample components into one combined sample, the 2020 sample cases were raked to external population counts to represent the 2020 population.

Step 11: Calibration to 2020 totals for 2020 sample cases

2020 sample cases were raked to the sample adult raking dimensions (age, sex, race/ethnicity, education, MSA status, Census division, plus housing tenure).

At this point, there were two separate sample components, each representing the full population (with the reinterview cases representing the 2019 population and the 2020 sample cases representing the 2020 population). Combining the two into one without doubling the estimated population counts requires a proportional adjustment. Typically, when concatenating multiple annual samples together for pooled-years analysis, the procedure to reset the sampling weights such that they sum to the appropriate total would be to divide each weight by the number of years being concatenated. For example, if three years of data were being combined, the weights for each would be divided by three such that the sum of weights across the three years' worth of data would equal the average of the three population counts as estimated by each single year, and not three times the size of that population.

Because the combined 2020 sample adult file is concatenating two independent samples, each of which represents the population of a given year, a similar approach would be to divide the weights by two. However, $\frac{1}{2}$ and $\frac{1}{2}$ is not the only set of proportions that could be used. So long as the two proportions sum to 1, the resulting weights will sum to the appropriate population total. But an even allocation of $\frac{1}{2}$ and $\frac{1}{2}$ is not the best approach here. The sample sizes of the two components are not approximately equal: the reinterview component is about half the size (approximately 10,000) of the 2020 sample component (approximately 20,000), yet each represents essentially the same population. Effectively, the average weight for the reinterview cases is approximately double the average weight for the 2020 sample cases. A proportional adjustment of $\frac{1}{2}$ and $\frac{1}{2}$ would retain that heterogeneity in sampling weights between the two subgroups. An adjustment proportional to sample size, however, would smooth (i.e., homogenize) the weights. The reinterview cases comprise 33% and the 2020 sample cases comprise 67% of the combined sample. Therefore, the proportional adjustments to the weights were set to 0.33 and 0.67, respectively.

Step 12: Proportional adjustment to reinterview cases

The 2019-raked weights for the reinterview cases received a ratio adjustment of 0.33.

Step 13: Proportional adjustment to 2020 sample cases

The 2020-raked weights for the 2020 sample cases received a ratio adjustment of 0.67.

At this point the two streams converged, and the weights for the combined reinterview and 2020 sample were raked jointly to external population control totals for 2020.

Step 14: Raking to external population totals for combined sample

The combined sample was raked to the sample adult raking dimensions for 2020 (age, sex, race/ethnicity, education, MSA status, Census division, plus housing tenure).

The resulting weight is the final annual weight for the 2020 sample adult file, WTFA_A.

V. The Partial 2020 Weight

Because some analyses may be negatively impacted by including the reinterview cases in the 2020 file, a partial weight restricted to cases that were originally sampled in 2020 (i.e., the cases that were not reinterviews of 2019 sampled cases) was also produced. Specifically, pooled analysis of concatenated 2019 and 2020 sample adult files will include a subset of approximately 10,000 people who were interviewed both in 2019 and 2020. These individuals would appear twice in the pooled 2019-2020 data file – once as a 2019 respondent and once as a 2020 respondent. Although their 2020 data values could differ from their 2019 values for many variables, it is still expected that these cases will be correlated, which would have an impact on variance estimation, potentially increasing the possibility of type 1 error (rejecting a true null hypothesis). The simplest method to handle this potential complication is to drop the reinterview cases from the 2020 file when combining 2019 and 2020 data, so that the cases who were sampled in 2019 do not appear

twice in the concatenated datafile. The partial weight is designed to allow the 21,153 sample cases from the 2020 NHIS to represent the 2020 population by themselves.

The weighting adjustment steps for the partial weight are shown in Flowchart 3. As with the other weight constructions, the adjustment process begins with the base weights. These base weights were created for the original 2020 sample, unrelated to the combined base weights created due to the inclusion of the reinterview cases as part of the 2020 annual file.

Step 1: Base weights

Generate base weights for the 2020 sample cases by allocating population counts according to the probability of selection within strata, to the subsampled-in panels of 2020 only.

With no need for a subsampling adjustment, the next steps are to account for 2020 household-level nonresponse, transform the weights to person-level, and adjust for nonresponse at the sample adult level. These are the same steps as Steps 8, 9 and 10 above for the 2020 Sample Adult weight, but they need to be repeated because this weight begins with a different set of base weights, apportioned to the 2020 sample cases only. The household-level nonresponse adjustment followed the usual procedure of predicting response propensities and adjusting within propensity quintiles by the inverse of the median response propensity of the quintile.

Step 2: 2020 Household-level nonresponse adjustment for 2020 sample cases

2020 sample cases received a ratio adjustment to the 2020 base weight equal to $1 /$ the median response propensity in the household-level response propensity quintile.

The household level weights were then transformed into person-level weights for adults and the nonresponse model was estimated at the person level.

Step 3: Adjust 2020 sample cases for sample adult probability of selection

2020 sample cases received a ratio adjustment to the household-level nonresponse-adjusted weight equal to $1 /$ the sample adult probability of selection.

The person-level weights adjusted for within household selection probabilities were then adjusted for nonresponse to the 2020 sample adult interview.

Step 4: Sample adult nonresponse adjustment for 2020 sample cases

2020 sample adult cases received a ratio adjustment to the household-selection-adjusted weight equal to $1 /$ the median response propensity in the sample adult-level response propensity quintile.

The nonresponse-adjusted sample adult weight was then calibrated to external control totals for the 2020 population.

Step 5: Calibration to 2020 totals

For 2020 sample cases, the nonresponse-adjusted weights were raked to the sample adult raking dimensions (age, sex, race/ethnicity, education, MSA status, Census division, plus housing tenure).

The resulting weight is the partial weight for analysis of 2020 sample adult cases only: WTSA_P.

Note, the process above for the partial weight is the process that would have been employed if there were no reinterview component, that is, if 2020 were a normal year.

VI. The 2020 Sample Child Weight

Development of the 2020 sample child weight uses the same procedures as used for the 2019 sample child weight (2) and 2020 partial weight (flowchart 3). The adjustment process begins with the base weights created for the original 2020 sample. The sample child weight is designed to allow the responding sample children ($n=5,790$) to represent the 2020 child population.

Step 1: Base weights

Generate base weights for the 2020 sample cases by allocating population counts according to the probability of selection within strata, to the subsampled-in panels of 2020.

The next steps account for 2020 household-level nonresponse, transform the weights to person-level, and adjust for nonresponse at the sample child level. The household-level nonresponse adjustment followed the usual procedure of predicting response propensities and adjusting within propensity quintiles by the inverse of the median response propensity of the quintile.

Step 2: 2020 Household-level nonresponse adjustment for 2020 sample cases

2020 sample cases received a ratio adjustment to the 2020 base weight equal to $1 /$ the median response propensity in the household-level response propensity quintile.

The household level weights were then transformed into person-level weights for children.

Step 3: Adjust 2020 sample cases for sample child probability of selection

2020 sample cases received a ratio adjustment to the household-level nonresponse-adjusted weight equal to $1 /$ the sample child probability of selection.

The person-level weights adjusted for within household selection probabilities were then adjusted for nonresponse to the 2020 sample child interview.

Step 4: Sample child nonresponse adjustment for 2020 sample cases

2020 sample child cases received a ratio adjustment to the household-selection-adjusted weight equal to $1 /$ the median response propensity in the sample child-level response propensity quintile.

The nonresponse-adjusted sample child weight was then calibrated to external control totals for the 2020 population.

Step 5: Calibration to 2020 totals

The nonresponse-adjusted weights were raked to the sample child raking dimensions (age, sex, race/ethnicity, MSA status, Census division, plus housing tenure).

The resulting weight is the final annual weight for analysis of 2020 sample child cases: WTFA_C.

Bias Assessment

I. Introduction

After final adult and child weights were developed, bias in sociodemographic and health estimates due to nonresponse and coverage error was assessed for the longitudinal file, the 2020 sample adult file, and the 2020 partial sample file. The results of those assessments are presented in the next three sections.

II. Bias analysis in the reinterview sample

One major advantage of the reinterview design is the ability to directly quantify nonreinterview bias and the reduction in that bias when the weights are adjusted. Under the assumption that estimates based on the 2019 sample adult file are unbiased, or at least a reasonable benchmark, estimates based on the reinterview cases (before and after weighting adjustments) can be compared with 2019 estimates and the bias resulting from completing 2020 reinterviews with a subset of 2019 original respondents can be assessed. The analysis of nonreinterview bias in the reinterview sample is presented in Table 2 and includes all the demographic, socioeconomic, geographic and health variables that were included as potential RPM predictors (a total of 40 variables, or 78 indicators when accounting for variables with more than two categories).

Bias was measured directly, not estimated, as the difference in the 2019 estimates as estimated using data from the reinterview respondents and the 2019 estimates using data from the full set of original 2019 sample adult respondents. Bias reduction was assessed by summing the absolute values of the biases, across all the indicators, for the reinterview respondents weighted by the final 2019 sample adult weight (i.e., before the adjustments described above) and for the reinterview respondents weighted by the final adjusted weight `WTSA_L`. The ratio of summed bias after adjustment to summed bias before adjustment is the percentage of total bias (across 78

indicators) remaining after adjustment; its complement is the reduction in bias directly attributable to the weighting adjustments.

Of the 78 indicators shown in Table 2, 74 improved (i.e., showed less absolute bias) while only 4 worsened after weighting adjustment; that is, 74 indicators show smaller differences between columns (1) and (3) than between columns (1) and (2). Bias in the estimated percentage of adults with household incomes two to four times the federal poverty level, landline-only telephone status, disability status and public health insurance coverage worsened after adjustment, by half a percentage point or less.

For 63 indicators in Table 2, remaining bias after weighting adjustments is zero or statistically nonsignificant; 15 show remaining biases that are significantly different than zero. (A conservative significance test was used, that adjusted the error term for the overlap of reinterview respondents to the 2019 sample adults, to account for the covariance due to the nonindependence of the samples; this technique is conservative in the context of testing for bias nonsignificance because it makes smaller differences appear to be statistically significant.) Of the 15 significant indicators, five had biases of less than one percentage point and 9 had biases between 1 and 1.5 percentage points; only excellent/very good health status had a larger bias, overestimating excellent or very good health in the reinterview sample by 1.57 percentage points.

It is not surprising that the reinterview sample slightly overestimates excellent or very good health. Especially given the COVID-19 pandemic, there were almost certainly some deaths among the original respondents to the 2019 survey, and the reinterview sample is a sample of survivors, weighted back to represent the 2019 population. Survivors, in the aggregate, are likely healthier than the deceased respondents would have been had they also survived.

Several of the demographic and access to care indicators with statistically significant bias remaining are associated with affluence (slightly overestimating home ownership, regular source of care and private health insurance coverage while slightly underestimating family incomes below 100% of the federal poverty level and landline-only telephone status). Statistically significant bias remaining for health status, health care utilization and health behaviors are less easily interpreted

(overestimating ever diagnosed with high cholesterol and underestimating disability; overestimating prescription medications and dental cleaning and underestimating emergency room visit and cholesterol check; overestimating obesity and underestimating cigarette smoking).

By summing the absolute biases across the 78 indicators, before and after adjustments to the weights, and comparing the two sums, the total bias reduction that can be directly attributed to the reinterview weighting adjustments was quantified at 80%. The remaining biases across the 78 indicators are:

- with four exceptions, reduced after weighting;
- mostly nonsignificant;
- with one exception (1.57), smaller than 1.5 percentage points; and
- not consistent in direction regarding the association with health.

This is strong evidence that the weighting adjustment process has functioned as intended and reduced nonreinterview bias considerably. Researchers using the longitudinal weight should be aware that there may be a slight bias in favor of healthier and more affluent adults in the 2020 portion of a longitudinal 2019-2020 data set. .

III. Bias analysis in the 2020 sample adult file

The longitudinal weight for reinterview cases, WTSA_L, was shown above to have little bias (and those biases, at least among the variables examined, are known) when representing the 2019 population. This weight was re-raked to the external 2020 population control totals used in Steps 11 and 14 of the weight constructions to provide a weight that would produce 2020 estimates with known biases. The reinterview cases weighted by the longitudinal weight re-raked to the 2020 population can thus serve as a benchmark for analysis of bias in estimates produced from the combined annual file.

Whereas the biases presented in Table 2 are quantifiable and objectively measured, allowing for a direct quantification of the reduction in bias attributable to the weighting adjustments, the Table 3 analysis is less straightforward and relies on the assumption that the

reinterview sample, re-raked to 2020 totals, is the best standard to measure against. This is reasonable given the greater ability to adjust for biases using the wealth of sociodemographic and health information collected on these cases in 2019. While this is logical, it is still just an assumption, and Table 3 estimates should be interpreted with caution. While Table 2 biases are actual measurements of observed bias, Table 3 biases are merely estimates of potential bias.

Table 3 shows the comparison of the reinterview sample (weighted by WTSA_L re-raked to 2020) with the partial (WTSA_P) and combined (WTFA_A) samples. Comparisons of re-raked reinterview cases to the partial file used a t-test for independent samples to test differences between these separate sample components. Comparisons of re-raked reinterview cases to the combined file used the nested t-test described above, accounting for the nonindependence of the samples (and providing a conservative test). The indicator for “had cholesterol checked in previous 12 months” is not available from the 2020 questionnaire, and at the time of this writing the imputed income variables for 2020 were not yet available for analysis. Imputed poverty level was replaced with unimputed household income for Table 3. Because unimputed household income has one category more than imputed poverty level, there remain 78 indicators overall.

The comparison of estimates based on reinterview cases re-raked to 2020 and the combined annual file shows that for 60 of the 78 indicators, the differences were zero or nonsignificant (Table 3). Of the 18 significant differences:

- Two are consistent with the known biases in the reinterview sample: high cholesterol and emergency room visit. High cholesterol was overestimated by 0.93 percentage points by the reinterview cases representing the 2019 population (Table 2), a bias that likely carried through to the 2020 re-rake. The combined file is underestimating high cholesterol by 1.3 percentage points when compared against this biased standard (Table 3), suggesting that the combined file estimate is the true estimate (or at least closer to the truth). The emergency room visit biases are even closer to a complete offset (0.94 underestimate in reinterview (Table 2), 1.2 overestimate in combined sample (Table 3)). Other than high cholesterol, there were no significant health status biases found.

- Eleven indicators are multiple categories of four demographic variables with offsetting biases (for example, one category being overrepresented while another category of the same variable is underrepresented): household size, family structure, years at current residence and telephone status. Other than these, the only bias found among demographic variables was for one category of household income: adults with less than \$35,000 income were underestimated.
- The remaining biases were found in the health care access and utilizations domains, and while the biases are consistent within each domain, they are inconsistent across domains. The combined sample overestimates utilization (doctor visits and eye exams) but also overestimates difficulty with access to care (difficulty paying bills, foregoing medications due to cost). The access biases were also inconsistent with the income bias (the combined sample underestimates the percent in the lowest income category).

Overall, two indicators are reflecting bias in the reinterview sample and the remaining significant biases included six larger than 2 percentage points: one-person household size (underestimated in the 2020 sample adult file); four or more person households (overestimated); one adult, childless family structure (underestimated); landline and cell telephone status (overestimated); cell-only telephone status (underestimated); and eye exam (overestimated).

IV. Bias analysis in the partial sample

All the biases described above for the combined file are present and slightly larger when estimates are generated with the partial weight WTSA_P (Table 3). This is a natural consequence of comparing a combination with one portion of that combination: because the reinterview portion cannot differ from itself, differences between the reinterview portion and the combined file must be driven by the remainder of the combined file. Estimates from the combined file are essentially weighted averages of estimates from the two sample components (averages weighted by the proportional adjustments from Steps 12 and 13 above: 0.33 for reinterview and 0.67 for 2020 sample cases). In a sense, the biases in the partial file are being diluted by the reinterview cases in

the combined file, which are relatively unbiased due to the weighting adjustments. This suggests that the reinterview design is functioning as intended, in this case mitigating (but not eliminating) a coverage gap.

No additional significant biases were found, and of the 18 noted above for the combined file, although all are naturally slightly larger when estimated from the partial weight, two are no longer statistically significant. (This is likely due to the smaller sample size of the partial file and/or the independent t-test no longer providing a conservative test.)

Some estimates based on the partial weight appear to be more biased than estimates from the combined file, when compared against a standard of the reinterview cases re-raked to 2020 population counts. But analysts using the partial weight may be more concerned with whether partial estimates of health and health care will differ from the official estimates based on the full combined sample adult file. Table 3 also shows these comparisons, again using the conservative test for nested samples.

The comparison of estimates based on the combined file and on the partial weight shows that across 78 indicators, only 2 differed significantly (Table 3). Adults with 4-10 years at the current residence were underrepresented (0.8 percentage points) and adults from landline-only households were overrepresented (0.1 percentage points). There were no significant differences between the combined sample and the partial sample among all the health and health care domains, even with the conservative significance test.

When two samples produce an array of equivalent estimates, there are generally only two possibilities: 1) the weighting adjustments functioned as expected and both samples are accurately representing the population; or 2) the two samples with different weighting techniques are somehow representing the same inaccurate population. In this case, there are three samples (reinterview alone, partial alone, reinterview plus partial) that are largely estimating the same population.

V. Conclusion

The richness of the data available for nonrespondents to the 2020 reinterview of 2019 sample adults allowed for more precise adjustment to the reinterview sampling weights than is usually possible in most surveys-- adjustments tailored to reduce nonreinterview bias in health estimates. The observed bias was reduced by 80% after implementing weighting adjustments among reinterview cases. Thus, estimates based on the reinterview cases were shown to be relatively low in bias when representing the 2019 population. There is some evidence of a slight affluence bias remaining in the reinterview sample.

When the reinterview cases are re-raked to the 2020 population control totals, the similarity of 2020 estimates based on the reinterview cases and the 2020 sample cases, separately, as well as the similarity of reinterview estimates and the overall 2020 combined estimates, suggest that the 2020 combined file (combining reinterview and 2020 sample cases) and the 2020 partial file (regular 2020 sample cases only) are also largely unbiased when weighted by the corresponding adjusted weights. However, the combined file does retain a few biases after weighting adjustments. The combined sample appears to have underrepresented adults living alone, those in the lowest income category, and those who only have cell phones, while overrepresenting adults living in households with four or more persons or in households with landline and cell telephones. There were no biases detected for health status, health insurance coverage, or health behaviors. The lack of differences in estimates of health and health care between the overall 2020 combined estimates and those based on the partial weight may reassure researchers using the partial weight that their analysis should produce the same estimates as the combined file would have (within sampling error).

References

1. Parsons VL, Moriarity C, Jonas K, et al. Design and estimation for the National Health Interview Survey, 2006–2015. *Vital and Health Statistics* 2(165). Hyattsville, MD: National Center for Health Statistics, 2014.
2. Bramlett MD, Dahlhamer JM, Bose J and Blumberg SJ. New procedures for nonresponse adjustments to the 2019 National Health Interview Survey sampling weights. Hyattsville, MD: National Center for Health Statistics, September 2020.
3. Dahlhamer JM, Bramlett MD, Maitland A, Blumberg SJ. Preliminary evaluation of nonresponse bias due to the COVID-19 pandemic on National Health Interview Survey estimates, April-June 2020. Hyattsville, MD: National Center for Health Statistics, February 2021.
4. Bramlett MD, Foster EB, Frasier AM, et al. Design and operation of the National Survey of Adoptive Parents, 2007. *Vital and Health Statistics* 1(50). Hyattsville, MD: National Center for Health Statistics, 2010.
5. Bramlett MD, Brooks KS, Foster EB, et al. Design and operation of the National Survey of Adoptive Parents of Children with Special Health Care Needs, 2008. *Vital and Health Statistics* 1(51). Hyattsville, MD: National Center for Health Statistics, 2010.
6. Zablotsky B, Bramlett MD, George JM, et al. Design and operation: 2013 National Survey of Children in Nonparental Care and 2014 National Survey of the Diagnosis and Treatment of ADHD and Tourette Syndrome. *Vital and Health Statistics* 1(63). Hyattsville, MD: National Center for Health Statistics, 2019.

Table 1: Recursive Partitioning Model Covariates for Nonrecontact, Refusal and Nonreinterview Models

<i>Covariate</i>	<i>RC</i>	<i>R</i>	<i>RI</i>	<i>Covariate (continued)</i>	<i>RC</i>	<i>R</i>	<i>RI</i>
Person-level Demographic Characteristics				Health Status			
Age	√	√	√	Excellent/very good health	.	√	.
Gender	.	.	.	Disability (current)	.	.	.
Race and ethnicity	.	.	√	Wear glasses (current)	.	.	√
Marital status	√	.	√	High cholesterol (current)	.	.	.
Nativity	.	.	.	Asthma (ever diagnosed)	.	.	.
Household Demographic Characteristics				Arthritis (ever diagnosed)			
Household size	.	√	√	Pain at least 1 day in previous 3 months	.	.	.
Family Structure	.	.	.	Major depression	.	.	.
Years at current residence	√	√	.	Major anxiety	.	.	.
Telephone status	.	√	√	Access to Health Care			
Housing tenure	√	√	√	Regular source of care	.	.	.
Socioeconomic Characteristics				Delayed care due to cost			
Education	√	√	√	Difficulty paying medical bills	.	.	.
Employment Status	.	.	.	Did not get needed medications due to cost	.	.	.
Poverty level (imputed)	√	√	.	Did not get mental health care due to cost	.	.	.
Geographic Characteristics				Healthcare Utilization in previous 12 months			
Region	√	.	.	Doctor visit	.	.	.
Metropolitan Statistical Area status	√	√	.	Emergency room visit	.	.	.
Health Insurance Coverage				Prescription medications used			
Insurance coverage type	.	.	√	Flu vaccination	.	√	.
Health Behaviors				Cholesterol checked			
Body-mass Index	.	√	.	Eye exam	√	√	√
Current smoker	.	.	√	Mental health care	.	.	.
				Dental cleaning			

RC = used to form nonrecontact model nodes; R = used to form nonresponse nodes; RI = used to form overall nonreinterview nodes; √ = covariate included in final model.

Table 2: Bias analysis among reinterview (RI) cases: 2019 estimates

<i>Indicator</i>	<i>2019 Sample adult file (1)</i>	<i>RI cases with 2019 weight (2)</i>	<i>RI cases with WTSA_L (3)</i>	<i>Remaining bias (3) – (1)</i>
Age	<i>Percent (standard error)</i>			
18-29	20.9 (0.35)	16.8 (0.53)	20.9 (0.67)	0.01
30-44	25.1 (0.32)	24.2 (0.55)	25.1 (0.60)	0.01
45-64	33.0 (0.33)	34.7 (0.58)	33.0 (0.63)	0.00
65+	21.1 (0.28)	24.4 (0.51)	21.1 (0.51)	0.01
Sex				
Male	48.3 (0.35)	48.4 (0.60)	48.3 (0.68)	0.00
Race/ethnicity				
Hispanic	16.5 (0.43)	14.3 (0.57)	16.5 (0.71)	0.00
Non-Hispanic white	63.2 (0.52)	67.2 (0.81)	63.2 (0.95)	0.00
Non-Hispanic black	11.8 (0.34)	10.2 (0.50)	11.9 (0.61)	0.10
Non-Hispanic other	8.5 (0.31)	8.3 (0.50)	8.4 (0.57)	0.10
Education				
Less than high school	12.4 (0.32)	10.0 (0.44)	11.5 (0.54)	0.90
High school/General Equivalence Degree	27.5 (0.34)	24.1 (0.58)	28.3 (0.70)	0.85
Some college	31.1 (0.36)	31.5 (0.60)	31.2 (0.65)	0.05
Bachelor's or more	29.0 (0.38)	34.5 (0.66)	29.0 (0.65)	0.00
Legal Marital Status				
Never married	28.5 (0.36)	26.1 (0.61)	28.0 (0.45)	0.45
Currently married	52.5 (0.39)	59.6 (0.68)	52.8 (0.27)	0.27
Divorced/separated/widowed	19.0 (0.27)	14.3 (0.42)	19.2 (0.18)	0.18
Nativity				
U.S. born	81.4 (0.36)	82.4 (0.57)	80.7 (0.75)	0.75
Employment Status				
Employed full-time	52.4 (0.39)	51.9 (0.65)	52.7 (0.28)	0.28
Employed part-time	11.5 (0.25)	11.2 (0.38)	11.6 (0.07)	0.07
Not employed, retired	18.0 (0.27)	20.6 (0.47)	17.9 (0.11)	0.11
Not employed, other	18.1 (0.32)	16.3 (0.52)	17.8 (0.26)	0.26
Own or Rent Residence				
Own/buying	66.1 (0.44)	69.7 (0.65)	67.6 (1.48)	1.48*
Poverty Level (imputed income)				
In poverty	11.1 (0.28)	9.3 (0.42)	10.3 (0.87)	0.87*
1.00 - < 2.00	18.8 (0.30)	16.5 (0.47)	18.5 (0.29)	0.29
2.00 - < 4.00	30.8 (0.33)	30.3 (0.57)	31.5 (0.69)	0.69
4.00+	39.3 (0.42)	43.9 (0.69)	39.7 (0.47)	0.47
Household Size				
1 person	15.9 (0.22)	17.1 (0.35)	16.1 (0.38)	0.23
2 persons	34.5 (0.32)	37.0 (0.57)	34.9 (0.62)	0.35
3 persons	18.1 (0.27)	16.7 (0.45)	17.6 (0.53)	0.54
4+ persons	31.5 (0.39)	29.2 (0.65)	31.5 (0.73)	0.04

<i>Indicator</i>	<i>2019 Sample adult file (1)</i>	<i>RI cases with 2019 weight (2)</i>	<i>RI cases with WTSA_L (3)</i>	<i>Remaining bias</i>
Family Structure				
1 adult, 0 children	18.5 (0.26)	19.4 (0.41)	18.5 (0.44)	0.03
2+ adults, 0 children	47.8 (0.37)	48.7 (0.60)	47.0 (0.67)	0.79
1 adult, any children	3.0 (0.09)	2.6 (0.14)	3.1 (0.19)	0.17
2+ adults, any children	30.7 (0.38)	29.3 (0.60)	31.3 (0.68)	0.64
Years at Current Residence				
3 years or less	36.5 (0.41)	33.1 (0.63)	35.9 (0.73)	0.57
4 to 10 years	24.2 (0.31)	24.7 (0.54)	24.3 (0.60)	0.08
11 to 20 years	20.4 (0.31)	21.3 (0.51)	20.4 (0.55)	0.03
More than 20 years	19.0 (0.32)	21.0 (0.51)	19.4 (0.54)	0.46
Region				
Northeast	17.8 (0.43)	17.8 (0.86)	17.8 (1.00)	0.00
Midwest	21.0 (0.47)	22.6 (0.98)	21.0 (1.07)	0.00
South	37.7 (0.59)	35.1 (1.18)	37.7 (1.37)	0.00
West	23.5 (0.49)	24.6 (1.07)	23.5 (1.21)	0.00
Metropolitan Statistical Area Status				
Metropolitan	85.7 (1.20)	85.1 (1.54)	85.7 (1.59)	0.00
Micropolitan	9.0 (1.09)	9.4 (1.32)	9.0 (1.33)	0.00
Neither	5.4 (0.76)	5.5 (0.94)	5.4 (0.99)	0.00
Telephone Status				
Both landline and cell phone	36.2 (0.42)	38.6 (0.64)	36.3 (0.68)	0.11
Cell phone only	60.8 (0.44)	58.7 (0.66)	61.1 (0.70)	0.26
Landline only	2.4 (0.10)	2.2 (0.15)	2.1 (0.15)	0.35*
No phone	0.5 (0.06)	0.4 (0.10)	0.5 (0.13)	0.01
Health Status				
Excellent/very good health	57.4 (0.40)	59.2 (0.70)	59.0 (0.78)	1.57*
Disability	9.0 (0.21)	8.1 (0.32)	7.8 (0.32)	1.27*
Ever diagnosed with asthma	13.5 (0.25)	14.1 (0.45)	13.9 (0.48)	0.44
High Cholesterol	24.9 (0.31)	28.2 (0.53)	25.8 (0.56)	0.93*
Ever Diagnosed with Arthritis	21.4 (0.28)	23.5 (0.49)	21.4 (0.52)	0.04
Pain, 1+ Days in past 3 months	58.9 (0.40)	61.5 (0.63)	59.6 (0.72)	0.63
Major depression	6.9 (0.19)	6.4 (0.32)	6.7 (0.37)	0.23
Major anxiety	6.1 (0.18)	5.7 (0.29)	6.0 (0.34)	0.09
Wear glasses	61.7 (0.37)	65.7 (0.66)	61.9 (0.75)	0.17
Health Insurance Coverage				
Public coverage	26.5 (0.36)	26.1 (0.58)	25.6 (0.65)	0.88
Private coverage	61.9 (0.42)	64.8 (0.65)	63.1 (0.75)	1.26*
Uninsured	11.7 (0.27)	9.0 (0.39)	11.3 (0.51)	0.39
Health Care Service Access				
Regular source of care	87.8 (0.26)	90.3 (0.40)	88.6 (0.53)	0.81*
Delayed care due to cost	9.05 (0.22)	8.24 (0.35)	8.8 (0.40)	0.24
Difficulty paying medical bills	13.8 (0.27)	13.1 (0.46)	13.6 (0.50)	0.11

<i>Indicator</i>	<i>2019 Sample adult file (1)</i>	<i>RI cases with 2019 weight (2)</i>	<i>RI cases with WTSA_L (3)</i>	<i>Remaining bias</i>
Did not get needed meds due to cost	7.0 (0.19)	6.5 (0.32)	6.7 (0.35)	0.29
Did not get mental health care due to cost	4.4 (0.15)	4.3 (0.27)	4.3 (0.28)	0.01
Health Care Service Utilization				
Doctor visit in past 12 months	84.9 (0.27)	86.8 (0.44)	85.0 (0.55)	0.08
Emergency room visit, past 12 mos.	21.8 (0.31)	20.8 (0.50)	20.9 (0.55)	0.94*
Prescription medications, past 12 mos.	66.3 (0.36)	71.1 (0.59)	67.4 (0.70)	1.06*
Flu vaccination, past 12 mos.	46.8 (0.39)	50.3 (0.66)	45.9 (0.72)	0.91
Cholesterol checked, past 12 mos.	73.2 (0.35)	74.7 (0.60)	72.1 (0.70)	1.13*
Eye exam, past 12 mos.	54.1 (0.37)	57.1 (0.59)	53.9 (0.66)	0.17
Mental health care, past 12 mos.	9.5 (0.20)	9.8 (0.35)	9.8 (0.40)	0.24
Dental cleaning, past 12 mos.	65.3 (0.36)	68.2 (0.60)	66.4 (0.68)	1.13*
Health Behaviors				
Body-mass index Underweight/healthy weight	34.0 (0.35)	32.3 (0.61)	32.7 (0.69)	1.27*
Body-mass index Overweight	33.9 (0.34)	34.2 (0.59)	33.9 (0.64)	0.07
Body-mass index Obese	32.1 (0.36)	33.5 (0.63)	33.5 (0.67)	1.33*
Current smoker	14.0 (0.26)	12.0 (0.41)	12.9 (0.46)	1.09*

*Remaining bias is significantly >0 at the 0.05 level.

Source: National Health Interview Survey, 2019 & 2020

Table 3: Bias analysis using reinterview (RI) cases as standard for partial & combined files: 2020 estimates

<i>Indicator</i>	<i>RI cases re-raked to 2020</i>	<i>Combined file with WTFA_A</i>	<i>Partial file with WTSA_P</i>
Age	<i>Percent (standard error)</i>		
18-29	20.5 (0.74)	20.6 (0.38)	20.5 (0.48)
30-44	25.3 (0.62)	25.2 (0.33)	25.2 (0.40)
45-64	32.5 (0.64)	32.6 (0.35)	32.6 (0.45)
65+	21.8 (0.52)	21.7 (0.30)	21.7 (0.38)
Sex			
Male	48.3 (0.70)	48.3 (0.37)	48.3 (0.45)
Race/ethnicity			
Hispanic	16.7 (0.76)	16.6 (0.48)	16.7 (0.67)
Non-Hispanic white	64.1 (0.97)	63.9 (0.57)	63.8 (0.77)
Non-Hispanic black	11.7 (0.62)	11.6 (0.37)	11.6 (0.46)
Non-Hispanic other	7.5 (0.57)	7.9 (0.32)	7.9 (0.35)
Education			
Less than high school	11.4 (0.54)	11.9 (0.33)	11.9 (0.42)
High school/General Equivalence Degree	27.5 (0.72)	27.6 (0.38)	27.7 (0.47)
Some college	30.4 (0.66)	30.7 (0.36)	30.7 (0.44)
Bachelor's or more	30.7 (0.71)	29.8 (0.39)	29.8 (0.48)
Legal Marital Status			
Never married	28.7 (0.75)	29.1 (0.42)	29.3 (0.53)
Currently married	51.9 (0.76)	52.2 (0.42)	52.3 (0.53)
Divorced/separated/widowed	19.4 (0.53)	18.7 (0.27)	18.4 (0.32)
Nativity			
U.S. born	81.0 (0.69)	81.7 (0.39)	82.1 (0.52)
Employment Status			
Employed full-time	51.0 (0.75)	50.8 (0.39)	50.4 (0.48)
Employed part-time	10.0 (0.45)	10.0 (0.25)	10.1 (0.32)
Not employed, retired	19.3 (0.50)	19.1 (0.28)	19.3 (0.38)
Not employed, other	19.7 (0.63)	20.1 (0.36)	20.2 (0.45)
Own or Rent Residence			
Own/buying	68.0 (0.72)	67.9 (0.44)	67.9 (0.57)
Total Family Income			
< \$35,000	22.5 (0.69)	20.8 (0.36)*	20.8 (0.48)*
\$35,000 - < \$75,000	27.8 (0.62)	28.1 (0.36)	28.2 (0.46)
\$75,000 - < \$100,000	11.8 (0.46)	12.5 (0.26)	12.5 (0.33)
\$100,000 +	28.5 (0.72)	29.4 (0.41)	29.3 (0.52)
Unknown	9.5 (0.50)	9.2 (0.27)	9.2 (0.35)
Household Size			
1 person	18.0 (0.47)	15.4 (0.21)*	15.6 (0.27)*
2 persons	34.7 (0.63)	34.0 (0.36)	34.1 (0.45)

<i>Indicator</i>	<i>RI cases re-raked to 2020</i>	<i>Combined file with WTFA_A</i>	<i>Partial file with WTSA_P</i>
3 persons	17.2 (0.54)	18.0 (0.30)	17.7 (0.36)
4+ persons	30.2 (0.75)	32.7 (0.43)*	32.6 (0.53)*
Family Structure			
1 adult, 0 children	20.3 (0.51)	18.0 (0.25)*	18.2 (0.32)*
2+ adults, 0 children	46.9 (0.68)	48.8 (0.37)*	48.9 (0.47)*
1 adult, any children	3.2 (0.22)	2.6 (0.09)*	2.6 (0.12)*
2+ adults, any children	29.6 (0.69)	30.6 (0.39)	30.3 (0.49)
Years at Current Residence			
3 years or less	33.1 (0.76)	33.6 (0.43)	34.0 (0.53)
4 to 10 years	27.5 (0.68)	25.9 (0.34)*	25.1 (0.42)*†
11 to 20 years	19.6 (0.57)	20.7 (0.33)*	20.9 (0.42)
More than 20 years	19.8 (0.57)	19.8 (0.34)	20.1 (0.42)
Region			
Northeast	17.4 (0.99)	17.6 (0.46)	17.6 (0.82)
Midwest	20.8 (1.07)	20.9 (0.45)	20.9 (0.74)
South	38.1 (1.40)	37.9 (0.64)	37.9 (0.99)
West	23.8 (1.24)	23.7 (0.55)	23.7 (0.87)
Metropolitan Statistical Area Status			
Metropolitan	86.1 (1.57)	85.9 (1.20)	86.0 (1.39)
Micropolitan	8.8 (1.32)	8.9 (1.10)	8.9 (1.25)
Neither	5.2 (0.96)	5.2 (0.76)	5.1 (0.85)
Telephone Status			
Both landline and cell phone	31.0 (0.69)	33.4 (0.42)*	33.5 (0.52)*
Cell phone only	67.2 (0.71)	64.1 (0.43)*	63.8 (0.54)*
Landline only	1.7 (0.15)	2.2 (0.11)*	2.3 (0.14)*†
No phone	0.1 (0.03)	0.3 (0.06)*	0.4 (0.07)*
Health Status Measures			
Excellent/very good health	57.8 (0.69)	58.6 (0.39)	58.7 (0.48)
Disability	8.2 (0.36)	8.8 (0.22)	8.8 (0.26)
Ever diagnosed with asthma	14.5 (0.54)	14.2 (0.28)	14.1 (0.34)
High Cholesterol	27.4 (0.59)	26.1 (0.32)*	25.8 (0.39)*
Ever Diagnosed with Arthritis	21.3 (0.57)	20.9 (0.29)	20.8 (0.37)
Pain on at Least 1 Day in Past 3 Months	60.0 (0.79)	60.7 (0.43)	60.5 (0.54)
Chronic depression	4.5 (0.28)	4.5 (0.16)	4.6 (0.21)
Chronic anxiety	11.4 (0.45)	11.2 (0.24)	11.5 (0.32)
Wear glasses	61.5 (0.73)	62.6 (0.40)	63.0 (0.51)
Health Insurance Coverage			
Public coverage	25.6 (0.71)	25.6 (0.39)	25.3 (0.47)
Private coverage	64.0 (0.78)	64.0 (0.45)	64.1 (0.56)
Uninsured	10.4 (0.54)	10.4 (0.30)	10.6 (0.40)
Health Care Service Access			
Regular source of care	89.0 (0.50)	88.6 (0.28)	88.6 (0.37)

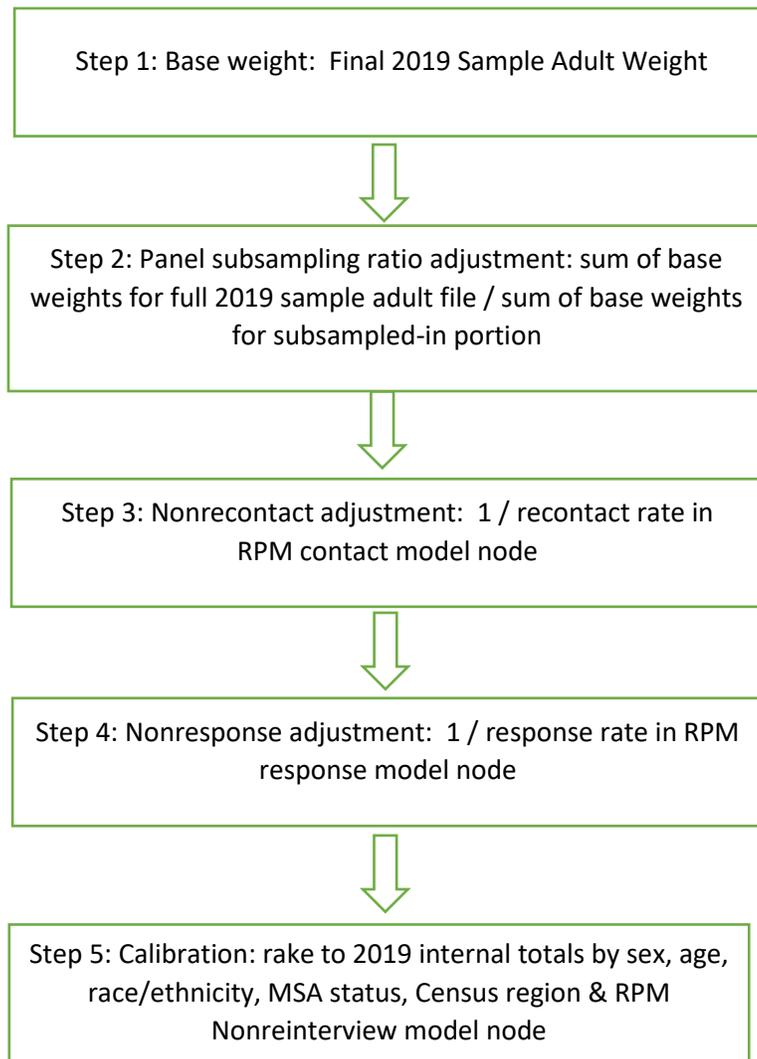
<i>Indicator</i>	<i>RI cases re-raked to 2020</i>	<i>Combined file with WTFA_A</i>	<i>Partial file with WTSA_P</i>
Delayed care due to cost	7.3 (0.40)	7.5 (0.22)	7.7 (0.28)
Difficulty paying medical bills	10.8 (0.45)	12.1 (0.28)*	12.2 (0.34)*
Did not get needed meds due to cost	4.9 (0.32)	5.7 (0.18)*	5.8 (0.23)*
Did not get mental health care due to cost	4.2 (0.31)	4.3 (0.16)	4.3 (0.19)
Health Care Service Utilization			
Doctor visit, past 12 mos.	81.9 (0.57)	83.4 (0.31)*	83.5 (0.37)*
Emergency room visit, past 12 mos.	17.8 (0.57)	19.0 (0.30)*	18.9 (0.36)
Prescription medications, past 12 mos.	65.1 (0.73)	65.3 (0.40)	65.3 (0.50)
Flu vaccination, past 12 mos.	46.7 (0.76)	47.9 (0.40)	48.2 (0.50)
Eye exam, past 12 mos.	44.1 (0.71)	46.4 (0.39)*	46.6 (0.49)*
Mental health care, past 12 mos.	10.3 (0.44)	10.1 (0.24)	10.0 (0.30)
Dental cleaning, past 12 mos.	62.0 (0.74)	63.0 (0.39)	62.8 (0.49)
Health Behaviors			
Body-mass index Underweight/healthy weight	32.5 (0.68)	33.1 (0.36)	33.2 (0.45)
Body-mass index Overweight	33.7 (0.65)	33.9 (0.35)	34.1 (0.43)
Body-mass index Obese	33.9 (0.70)	33.0 (0.37)	32.7 (0.48)
Current smoker	12.5 (0.49)	12.5 (0.27)	12.3 (0.35)

*Estimate is significantly different from that of RI cases re-raked to 2020, at the 0.05 level.

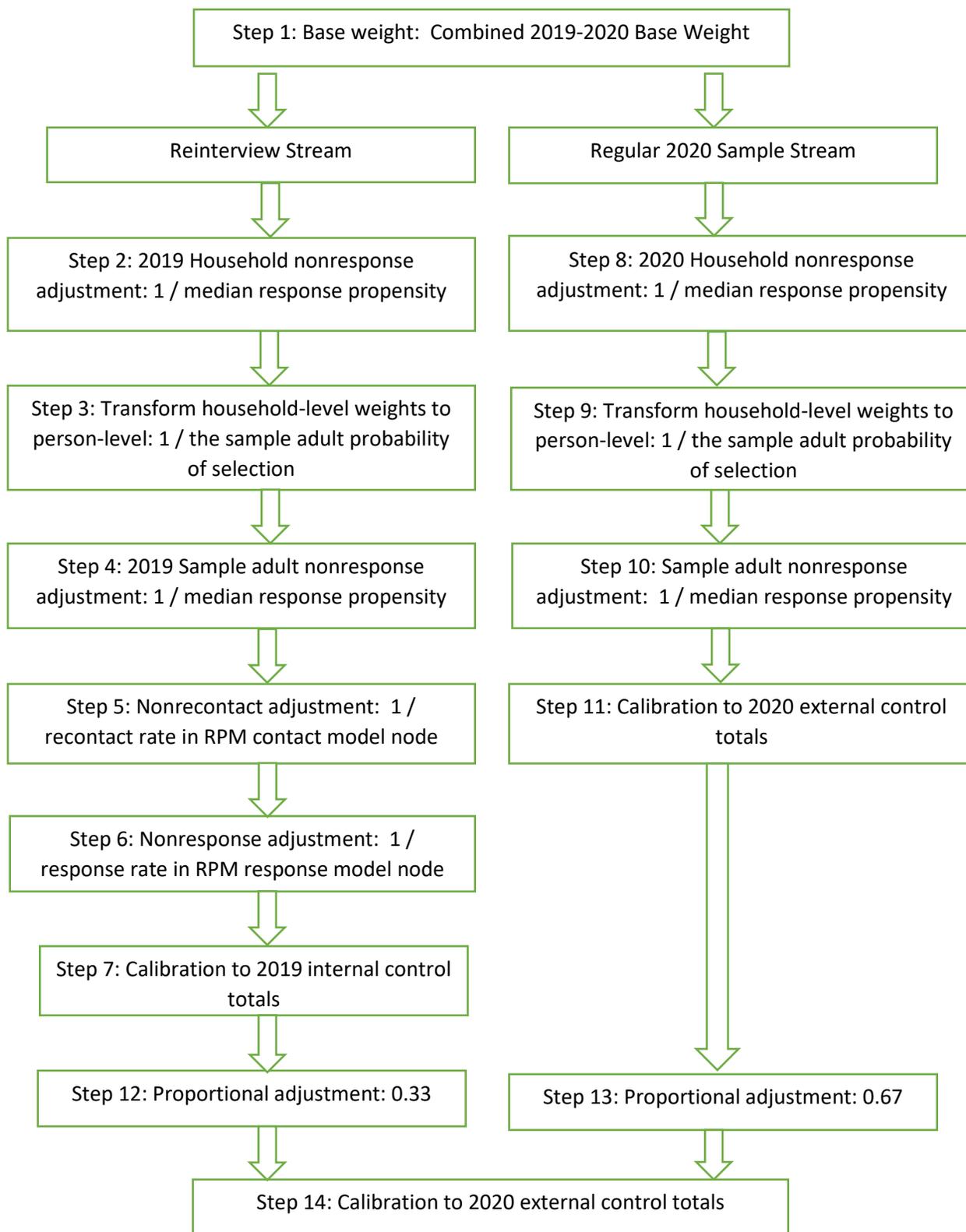
†Estimate is significantly different from that of the combined file with WTFA_A, at the 0.05 level.

Source: National Health Interview Survey, 2020

Flowchart 1: Longitudinal Weight (Reinterview cases only)



Flowchart 2: 2020 Combined Annual Weight



Flowchart 3: Partial 2020 Weight (2020 Sample cases only)

