NATIONAL CENTER FOR HEALTH STATISTICS

White Paper

July, 2026

The Impact on Bias Reduction from adding a Weighting Calibration Variable derived from a Recursive Partitioning Model

Matthew D. Bramlett, Ph.D., Lindsey I. Black, M.P.H., James M. Dahlhamer, Ph.D. & Benjamin Zablotsky, Ph.D.

Abstract

Objective

This report describes the improvement in bias reduction from including a variable generated from recursive partitioning models (RPM) into the raking dimensions for survey weighting calibration.

Methods

Data are drawn from the National Health Interview Survey 2021-2023 and the National Health Interview Survey – Teen 30-month file. RPM was used to leverage the rich information available for respondents and nonrespondents to generate a predictor of reinterview status that captures complex interactions among variables related to nonresponse and outcome measures of interest. This was then used as a raking dimension for weighting calibration.

Results

Adding the RPM-generated variable to the raking dimensions reduced the sum of biases across a broad array of indicators by an additional 4.7% as compared with alternative weights that omitted the RPM variable. The impact on variance inflation was minimal when adding the RPM variable to the raking dimensions.

Conclusion

Using RPM methods to derive a raking variable allowed for the efficient integration of multiple sample frame variables and the effective targeting of weighting adjustments where correlations with response propensity and correlations with survey outcomes of interest coincide.

Introduction

Followback surveys, in which respondents from a previous survey are recontacted for additional data collection, offer unique opportunities for precise adjustments to their sampling weights to correct for followback nonresponse and to reduce nonresponse bias. Because all eligible sample cases in a followback survey have completed the previous survey, the data available for followback respondents and nonrespondents are atypically broad, specific to the individual sample case, and often related to the followback survey outcomes of interest. Furthermore, because the original survey estimates can serve as gold standards, followback nonresponse bias and the reduction in nonresponse bias attributable to weighting adjustments are quantifiable.

This report highlights one unique contribution that was developed for the National Health Interview Survey (NHIS) 2019 Sample Adult Followback (SAFB) and subsequently used for the NHIS-Teen, another followback survey of NHIS Sample Child participants: generating a variable from a recursive partitioning model of reinterview status, to be used for raking to control totals derived from the original survey. The analysis presented in this report quantifies the impact of adding this variable to the raking dimensions, by calculating the additional percent reduction in bias (across a broad array of demographic characteristics and health survey measures) directly attributable to the addition of the raking variable.

This procedure, while developed for followback surveys, may also be broadly applicable to standalone surveys, where the sample frame contains relevant information on all eligible sample cases, or when auxiliary data (e.g., contextual data appended by geography, interviewer neighborhood observations, contact attempt histories) are related to nonresponse and to the survey outcomes of interest.

Background

NHIS, conducted by the National Center for Health Statistics' Division of Health Interview Statistics (DHIS), is a nationally representative household survey of the civilian noninstitutionalized U.S. population. Within each sampled household, one adult (Sample Adult) and one child, if present, (Sample Child) are randomly selected for a detailed health interview. Historically, NHIS was primarily based on face-to-face interviews, supplemented by telephone interviewing. However, due to the COVID-19 pandemic interviewing procedures were disrupted and a temporary shift to all-telephone interviewing was made. To address possible bias resulting from noncoverage of nontelephone households or underrepresentation of subgroups due to telephone interviewing, and to enable comparisons between the year prior to the pandemic and the first year of the pandemic, approximately one-third of the 2020 NHIS sample was replaced with a followback component consisting of reinterviews of Sample Adult respondents from the 2019 NHIS. These reinterviews were conducted by telephone. The resulting SAFB component was integrated into the 2020 NHIS annual file (2020 NHIS Questionnaires, Datasets, and Documentation | National Health Interview Survey | CDC) via weighting procedures described previously (1).

The procedures developed for SAFB weighting included a new innovation regarding the use of recursive partitioning models (RPM) to leverage the broad array of variables available for both respondents and nonrespondents to the SAFB (including the full data from the original 2019 NHIS interview). RPM was used to generate a variable that differentiated sample cases by response propensity. This variable was then used as an additional raking variable for calibration to population control totals. This procedure was subsequently used for another followback survey, National Health Interview Survey-Teen (NHIS-Teen), and that analysis is the focus of this report.

The NHIS Sample Child interview is a detailed health interview of children age 17 years or younger. However, Sample Children themselves are not interviewed. Instead, a knowledgeable adult responsible for the child's health and health care (typically a parent or guardian) answers questions about the Sample Child's health. For older children, such as teens, there may be limitations to parental report of their child's health, in particular for sensitive topics such as mental health, screen time, weight concerns, receipt of social and emotional support, and bullying. To address this, as part of the CDC Public Health Data Modernization Initiative, the NHIS-Teen, a followback survey with teenagers (Sample Children ages 12-17) for whom a parent had completed the Sample Child interview was developed.

Methods

National Health Interview Survey - Teen

The NHIS-Teen was conducted as a followback to the NHIS Sample Child interview (2). Upon completion of the NHIS Sample Child interview, sampled children ages 12-17 were invited to complete a self-administered survey conducted over the web. Parental permission along with teen contact information (cell phone number and/or e-mail) was obtained during the NHIS interview. Teens that participated typically completed the survey within 2-3 weeks from recruitment. More information about NHIS-Teen is available online (NHIS-Teen | National Health Interview Survey | CDC).

NHIS Teen recruitment began in July 2021 and ended in December of 2023. Teen weighting procedures were developed as data collection progressed, resulting in the creation of a 6-month and a 12-month weight for internal evaluation and an 18-month weight for use in producing official estimates on key health indicators. Bias reduction was calculated at each stage and weighting methodology was evaluated to identify strategies for the creation of a final weight to be used with a 30-month data file. Analysis for this report is based on the 30-month file and weight, covering data collection from July 2021 through the end of 2023.

The sample dispositions of teen cases in the 30-month data collection are shown in Table A.

Table A: NHIS-Teen 30-Month sample dispositions

Sample disposition	Number of cases
Total (full sample frame) Sample Children ages 12-17	7,323
Ineligible, NHIS interview ended before permission request	296
Parental permission requested but not received	2,783
Permission received but teen did not complete	2,223
Teen completed but case was dropped due to poor quality*	63
Teen completed interview, present in final data file	1,958

*Dropped cases are subsequently treated as 'permission received but teen did not complete'
SOURCE: National Center for Health Statistics, National Health Interview Survey – Teen 30-month file.

The teen permission rate is 100*[(2,223 + 63 + 1,958)/(2,223 + 63 + 1,958 + 2,783)] = 60.4%.
The teen interview completion rate is 100*[1,958/(1,958 + 2,223 + 63)] = 46.1%.
The teen response rate is 100*[1,958/(1,958 + 2,783 + 2,223 + 63)] = 27.9%.
The teen reinterview rate is 100*[1,958/(1,958 + 2,783 + 2,223 + 63 + 296)] = 26.7%.

"Reinterview" may normally refer to a followback interview that has been completed with the original respondent, but in this case the respondents for the original NHIS interview and the NHIS-Teen interview were different people (parent and teen). However, "reinterview rate" is still preferable to "response rate" or "interview rate" (which could be confused with "interview completion rate"). At a household level, the teen interview is a reinterview of the NHIS housing unit, so "reinterview status" is used to refer to whether a teen interview was completed, or not (for any reason).

The reinterview rate of less than 27% for NHIS-Teen entails a high potential for nonresponse bias if the teens who responded differ substantively on NHIS-Teen survey estimates from the remainder who did not respond or were never given the opportunity to respond. Therefore, it is important to accurately adjust the weights so that the weighted pool of final respondents better mirrors the full sample frame in terms of the distributions of demographic and critical stratifying variables related to response propensity.

Recursive partitioning models

A recursive partitioning model was used to generate a predictor of reinterview status, to be added to the raking dimensions. The model was estimated using the function ctree() in the R package partykit (R version 4.4.0) (3).

The first step was to categorize the full sample frame of NHIS Sample Children ages 12-17 from July 2021-2023 by reinterview status as either 1) a completed NHIS-Teen interview, or 2) not a completed NHIS-Teen interview. This household reinterview categorization combines in category 2 all cases that did not complete an interview for any reason, whether due to teen nonresponse; parental permission refusal; ineligibility due to the NHIS respondent not being the teen's parent or guardian or the parent/guardian ending the interview prior to being asked permission; or interviews that were dropped (completed interviews that were removed from the file due to poor data quality).

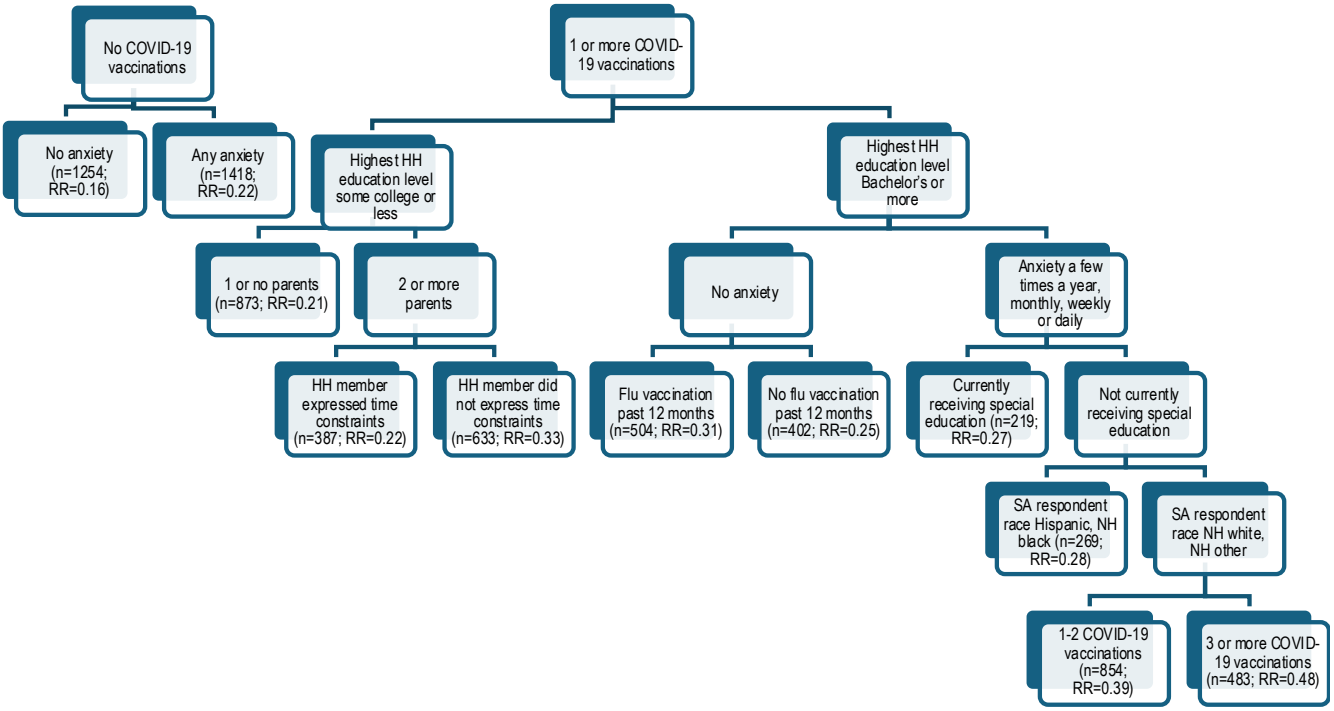
The next step was to model reinterview status, where the set of predictors included demographic, socioeconomic and geographic characteristics as well as a broad array of NHIS health measures. NHIS health measures that were selected had significant associations with teen response and – among the respondents – with key teen survey outcomes of interest. The set of NHIS health measures that were screened for inclusion in the model, available for both respondents and nonrespondents, included a broad array of variables from the NHIS Sample Child interview in addition to auxiliary (contextual or operational) data.

Because of the very large number of potential predictors and the desire to incorporate interactions, RPM was used to model reinterview propensity. RPM produces a set of nodes by first splitting the sample by the strongest predictor of reinterview, 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 (4). The model selects the best categorizations for covariates; for example, if age in single years is included in the set of covariates with 6 categories 12-17, the model may split that covariate into noncontiguous categories (e.g., collapsing 13 & 17 year-olds versus all others) in creating the nodal splits, because that is where the differentiation on reinterview status is concentrated. The model also effectively selects the relevant interaction effects among combinations of variables (subsequent splits within subgroups are essentially interactions between the variables used to predict reinterview propensity), without the researcher having to prespecify the interaction terms. This is particularly useful when dealing with a large set of potential covariates and little *a priori* knowledge. The primary benefit of this approach is that it results in a set of mutually exclusive and exhaustive

categories (terminal nodes) which are differentiated by reinterview propensity and defined by interactions of variables associated with key survey outcomes of interest. Terminal nodes can be used to form adjustment cells for precise weighting adjustments.

Figure 1 shows the tree that resulted from the RPM of reinterview status. The tree has 11 terminal nodes. In Figure 1, the terminal nodes are represented by the boxes that include sample size and reinterview propensity for the node in the box text. The first terminal node, at far left, represents the interaction of number of COVID-19 vaccinations with frequency of anxiety and shows that 1,254 teens had had zero COVID-19 shots and never had anxiety, and among them, 16.3% completed the interview. The last node, at far right, shows that there were 483 teens who had had any COVID-19 shots, who lived in households where the highest education was Bachelor's or higher, who ever had anxiety and were not currently receiving special education services, whose household Sample Adult respondent was of non-Hispanic White or Other race, and who had 3 or more COVID-19 shots, and these teens had a reinterview propensity of 48.4%.

Figure 1: Recursive partitioning model splits and terminal nodes



NOTES: RR is reinterview rate. HH is household. SA is Sample Adult. NH is non-Hispanic.
SOURCES: National Center for Health Statistics, National Health Interview Survey 2021-2023 and National Health Interview Survey – Teen 30-month file.

Translating RPM results into a raking variable

The typical approach to modeling response propensity for weighting adjustment is to generate classes of predicted response propensity (e.g., quintiles) and then apply a nonresponse adjustment based on those classes (e.g., multiplying the weight by the inverse of the mean response propensity in the class) (5). The novel approach to use RPM results to create a variable for calibration was first used in the 2020 NHIS SAFB (1). The terminal nodes from the RPM (Figure 1) were each used as one category in

a variable derived from the interactions of the covariates used for the subgroup splits. Every terminal node equates to a category, and the overall calibration variable had 11 mutually exclusive and exhaustive categories, matching the number of terminal nodes.

Weighted estimates of this RPM-derived calibration variable, based on the full sample frame for Sample Children ages 12-17 (i.e., both respondents and nonrespondents to NHIS-Teen) and weighted by the original final Sample Child weight from NHIS, provided the estimated population control totals that served as the inputs for the raking program (i.e., the totals that the teen sample are raked to match to).

For more detail on the weighting adjustments for the 30-month NHIS-Teen file, including more information on the raking calibration procedures, see Bramlett *et al,* 2024 (6).

Quantification of bias and bias reduction

The objective in adjusting the NHIS-Teen weights is for the final set of teen respondents to represent the same population as was represented by the full sample frame (i.e., the NHIS final weighted sample of Sample Children ages 12-17 for July 2021-December 2023). Thus, weighted estimates from the NHIS Sample Child interview are “gold standard” estimates for NHIS-Teen bias evaluation.

The term “gold standard” is not meant to suggest judgment as to whether teen or parent responses are more or less accurate, nor to suggest a reference to the “best” estimates available. In this context, gold standard refers to the estimates for the full sample (all Sample Children ages 12-17) that would be obtained by the subsample (the 27% reinterviewed teens) if the weighting adjustments were successful in eliminating nonreinterview bias.

Estimates for selected health measures were calculated for followback respondents using both the original Sample Child weights and adjusted followback weights. Estimates based on the followback sample were considered unbiased if they were similar to the original gold standard estimates. The amount of bias was measured directly as the difference between gold standard estimates and followback estimates weighted by the original NHIS Sample Child weights (i.e., in the absence of followback weighting adjustments); remaining bias after weighting adjustments was similarly measured as the difference between gold standard estimates and followback estimates weighted by the adjusted final followback weights.

The change in bias directly attributable to weighting adjustments was calculated as the difference between measured bias and measured remaining bias after weighting adjustments. The change in bias is typically a reduction in bias if the weights are functioning as intended.

By selecting a broad array of health outcomes of interest and summing the absolute values of their biases before and after weighting adjustments, the overall bias reduction quantification is distilled down to a single number: the percent of total bias observed that was eliminated by weighting adjustments. This metric was used to assess which version of a final weight performed better at reducing bias. The bias assessment includes all the survey variables that were potential RPM predictors, as well as representative indicators from the major health and healthcare domains measured by NHIS, plus demographic, socioeconomic, and geographic measures.

The full list of indicators included in the bias assessment is shown in Table 1. Although the primary analysis measures total bias across all these indicators, a supplemental analysis examines the efficacy of the weighting adjustments in reducing bias in subsets of the full set of variables.

Quantifying the bias reduction attributable to any individual step or combination of steps in the weighting adjustment process was achieved by assessing and comparing the bias reduction from two versions of the final adjusted weight: the actual final weight and a version that omits the weighting step(s) under consideration. For example, by measuring the bias reduction with a version of the final weight that differs only in that it omits the RPM variable from the raking, the reduction that would have occurred in the absence of this variable's inclusion can be quantified. Comparing that result to the actual bias reduction from the final weight quantifies the bias reduction directly attributable to the inclusion of that variable in the raking.

Therefore, to assess the additional reduction in bias that results from using an RPM-derived variable for raking, the bias reduction was calculated using the fully-adjusted final NHIS-Teen weight, and calculated using an alternate version of the NHIS-Teen weight that omits the RPM-derived variable from the raking, and the results were compared. This comparison derives the bias reduction that results when the RPM variable is added to the rest of the final NHIS-Teen 30-month raking dimensions.

Analysis

Seven different weights were used to generate bias reduction measures under different scenarios:

Weight A: Unadjusted teen base weight: original NHIS Sample Child weight (30 months, annualized);
Weight B: Final 30-month NHIS Teen weight: the weight as described in Bramlett *et al*, 2024 (6);
Weight C: Weight B, except the RPM variable was excluded from the raking;
Weight D: Weight B, except the raking dimensions are the original NHIS Sample Child dimensions (i.e., excluding income and insurance coverage and with age collapsed) and the RPM variable;
Weight E: Weight D, except the RPM variable was excluded from the raking;
Weight F: Weight B, except the RPM variable was the only variable included in the raking;
Weight G: Weight A raked to RPM totals with no other adjustments.

Weights B, C, D, E and F all had the same nonresponse adjustments (the permission and participation adjustments as described in Bramlett *et al*, 2024) and differ only in the raking dimensions used to calibrate to control totals. Weight A is the unchanged original NHIS Sample Child weight for the 30 months, rescaled to annual population totals (essentially, the base weight in the NHIS-Teen sample frame); Weight G is weight A, raked to the RPM variable totals.

The impact of the RPM variable is shown in multiple ways:

As described above, the bias reduction was compared using the actual final NHIS-Teen weight (weight B), to the bias reduction using a version of the final weight that omitted the RPM variable from the raking (weight C). This is the primary analysis.

The impact of the RPM variable's addition to the original NHIS Sample Child raking dimensions was also assessed. The raking dimensions used for the original NHIS were expanded for NHIS-Teen by adding insurance coverage and income to the raking dimensions and breaking age groups into single years of age (based on evaluation of the 18-month teen weight) (7). In this secondary assessment, the bias reduction using an alternate teen weight that was raked to the original Sample Child totals, rather than the expanded teen totals (weight E), was compared to the bias reduction using an alternate teen weight that was raked to the original Sample Child totals plus the RPM variable (weight D). This is the secondary analysis.

Given that the variables included in the raking do not necessarily have independent effects on bias reduction and that some of the reduction that could have been achieved with the RPM variable was already being achieved with the other sociodemographic characteristics included in the raking, the bias reduction was also calculated using two versions of the weight that included only the RPM variable in the raking (weights F and G).

The bias reduction effects of individual variables included in the raking are not necessarily independent nor summative, if multiple raking variables are related to the outcomes of interest. Removing a variable that would have reduced bias in certain outcomes may still result in unbiased outcomes if other raking variables are also related to the outcomes. To disentangle overlapping raking effects, weight F was generated, which omitted all raking variables except the RPM variable so the total potential impact it might have had as a raking variable could be measured.

It is also possible that some of the bias reduction achieved with the RPM variable as a raking dimension would still have occurred via the earlier permission and/or participation adjustments. To disentangle overlapping effects, weight G was generated, which omitted all raking dimensions except the RPM variable and also omitted the permission and participation adjustments prior to raking, so the maximum possible impact of the RPM variable as the sole weighting adjustment could be measured.

Additionally, the impact of adding the RPM variable on reducing bias in subgroups of the full list of indicators was assessed. The full set of indicators includes some that are utilized as raking dimensions, some that are components of the RPM variable, and many others that are not entered into the raking either directly or indirectly. Bias reduction in these three subgroups was assessed separately using each weight, and compared by weight.

Finally, the impact of adding the RPM variable to the raking on the variance of estimates was assessed. Relative standard errors were calculated and averaged across the set of indicators using each weight, and compared by weight.

Results

Table 1 presents the bias assessment for the primary and secondary analyses and shows all the indicators that were included in the bias assessment, as well as the domains from which they were drawn. The first column of data (1) shows the gold standard estimates based on the NHIS-Teen sample frame: the actual NHIS estimates for Sample Children ages 12-17 using the Sample Child weights, annualized for the 30-month period (weight A). The second column of data (2) shows the estimates for the NHIS-Teen respondents using the same weight A – had there been a 100% reinterview rate, these would be the same estimates as in the previous column. The difference between estimates in these two columns is the non-reinterview bias in teen estimates if there were no weighting adjustments at all (this bias is shown in the third column of data).

The fourth column of data (3) shows the actual NHIS-Teen 30-month estimates and standard errors, using the final 30-month weight as described in Bramlett *et al*, 2024 (weight B). The fifth data column shows the remaining bias in teen estimates when using the final 30-month weight, calculated as the absolute difference between (3) and (1). The sixth data column (4) shows the teen estimates that would have been obtained had the RPM variable been left out of the raking of the final 30-month weight (weight C), and the seventh data column shows the bias (|(4)-(1)|) that would have remained had the RPM variable been left out of the raking. Comparing the bias reduction in the official estimates with the bias reduction when the RPM variable was omitted reveals the impact of adding the RPM variable to the teen weight raking dimensions (the primary analysis).

The eighth and ninth data columns show the bias reduction using weights D and E. These columns show the biases that would remain in the presence or absence of the RPM variable had the original Sample Child raking dimensions not been expanded for NHIS-Teen to include single years of age or to add insurance coverage and income to the raking dimensions. Comparing the bias reduction in these two columns is the secondary analysis, revealing the impact of adding the RPM variable to the original Sample Child raking dimensions.

(The estimates and standard errors for weights D and E are omitted here for brevity but are presented in Table I in the appendix.)

Table 2 shows the sum of biases across all the indicators in Table 1 as generated by each of the seven weights A-G, and the percent bias reduction for each weight relative to weight A. (The full set of estimates, standard errors, and biases for weights F and G are presented in Table I in the appendix.)

Impact of adding the RPM variable to teen weight raking dimensions

As shown in Table 2, the total initial bias of 293.2 was reduced to 73.1 with the final weight B (a 75.1% reduction) and reduced to 82.9 when the final weight omitted the RPM variable (weight C; a 71.7% reduction). This resulted in a 4.7% reduction in the bias attributable to the addition of the RPM variable to the raking dimensions.

Impact of adding the RPM variable to original Sample Child dimensions

The total initial bias of 293.2 was reduced to 87.9 with weight D (a 70% reduction) and reduced to 96.4 with weight E (a 67.1% reduction) (Table 2). This resulted in a 4.3% improvement in the bias reduction. It appears that the bias reduction impacts of a) adding the RPM variable to the raking and b) expanding the original raking dimensions to include insurance, income, and single years of age were largely independent effects, as the bias reduction boost from adding the RPM variable was similar whether it was added to the original or the expanded dimensions. The impact was slightly greater when the RPM variable was added to the expanded dimensions rather than the original dimensions, which would not be expected if the impacts of adding the RPM variable and expanding the original dimensions were overlapping effects. This is consistent with the internal (i.e., unpublished) evaluation based on the 18-month NHIS-Teen weight that showed that persistent biases remained in insurance and income estimates after weighting adjustments, until those dimensions were explicitly controlled in the raking.

Given the alternate weights that were created for this analysis, it is also possible to quantify the bias impact of expanding the original Sample Child control totals to include single years of age and add insurance and income. Weights B and D are identical except that B was raked to the expanded totals (plus the RPM variable) and D was raked to the original totals (plus the RPM variable). The bias reduction with the original totals was 70.0% while the reduction with the expanded totals was 75.1%, a relative increase in bias reduction of 7.3%. This again is consistent with the internal evaluation based on 18-month data; insurance and income variables had some of the largest biases that remained after all other weighting adjustments, until they were explicitly controlled in the raking and their biases were eliminated to zero.

Impact of raking to the RPM variable alone

Because some of the bias reduction accomplished by the inclusion of the RPM variable in the raking could likely have occurred in its absence, weight F was generated to show the impact if the RPM

variable had been the only variable included in the raking. Table 2 shows that omitting the other raking variables and raking to the RPM variable alone would have resulted in an almost two-thirds (65.1%) reduction in bias.

Because some of that 65.1% bias reduction could have occurred as a result of the initial permission and participation adjustments (see Bramlett *et al*, 2024), weight G was generated to show the total sole impact of the RPM variable. If no adjustments were made to the base weights (i.e., the original NHIS Sample Child weights), other than raking to the RPM variable, the bias reduction would have totaled 58.2%.

Bias reduction in variables not explicitly controlled in the raking

The sum of biases across a set of indicators is a metric that simplifies analysis, but is a crude measure. When two weights produce a similar amount of overall bias reduction, it is not necessarily the case that the same indicators are becoming unbiased with either weight, nor that bias for specific indicators is reduced by the same amount regardless of which weight is used. If certain indicators are critical for broad-based analysis or measure the main survey outcomes of interest, it may be more important to ensure that bias in those specific indicators is minimized rather than focusing on overall bias reduction.

When a variable is included directly in the raking, its bias is eliminated to zero. For variables that were used to code categories of the RPM variable, the specific categories effectively controlled in the raking will not match to (will generally be collapsed versions of) the categories of those indicators in the bias assessment, so for components of the RPM variable, the expectation would be that the biases would be reduced, but not quite eliminated. All the other indicators in Table 1 are variables that do not enter the raking in any way; bias in these variables is reduced because they are correlated with the variables included in the raking.

Table 3 disentangles the sum of bias and percent bias reduction from Table 2 by subgrouping the indicators into raking variables; RPM components; and other variables, to enable a focus on the other variables. (These are the majority of the indicators in Table 1.) The impact of adding the RPM variable to the raking can be seen in the comparison of weights B and C and in the comparison of weights D and E. In both cases, the weight that included the RPM variable reduced the bias in non-raked variables further, from 70 or more to a little more than 66. In fact, the results for weight G show that as much bias reduction in non-raked variables is achieved by raking to the RPM variable alone as is achieved by the final teen weight (about 55% bias reduction in both cases).

The results for weight G in Table 3 indicate that if no adjustments were made to the base weights other than raking to the RPM variable, there still would have been almost 50% bias reduction in the non-raked indicators.

Variance

Weighting adjustments tend to involve a trade-off between bias reduction and variance inflation. That is, the more the weights are adjusted in attempts to reduce bias, the greater the variance in estimates that may result from the weights being more heterogenous (8). Table 4 presents relative standard errors (RSE; standard error as a proportion of the estimate), averaged across the Table 1 indicators, for each weight A-G. As expected, the unadjusted weight A shows the lowest average variance while the adjusted weights are all associated with higher variance. However, the differences among the adjusted weights are slight. Adding the RPM variable to the raking did not greatly increase the variance inflation. Weights B and C have the same average RSE, while weights D and E differ by only 0.001.

Discussion

These results support the broader use of any variable related to survey response and key survey outcomes as a calibration variable. The RPM variable included sociodemographic measures not already being used for raking (highest education among parents, number of parents); health and healthcare measures from the Sample Child interview (vaccinations, anxiety, special education services); and field operational data from the Contact History instrument (whether a household member had expressed time constraints). Researchers are not limited to external control totals provided by other agencies or ancillary data sources. These sources may not be aggregated in the best categories for a particular survey given its specific content and outcomes of interest. If data are available for both respondents and non-respondents, so they can be modelled to determine what matters for predicting nonresponse, any sample frame variable can become part of a tree variable for raking. Additionally, anything on the sample frame (defined broadly, including appendable contextual data or operational data) that continues to show unacceptable bias after utilizing RPM to improve the raking can be added to the raking as well, by deriving internal control totals from the frame (as was done for NHIS-Teen regarding income and insurance coverage in the expanded raking dimensions).

When a followback survey like NHIS-Teen attempts a reinterview with every sample case that completed the original survey, the set of original survey respondents is the full sample frame for the followback survey. Adapting these procedures for a standalone survey like NHIS requires a focus on the sample draw rather than the full sample frame. For example, the sample frame for NHIS is all housing units in the United States, but the set of housing units selected for contact within each geographic area that will be included in the sample is the sample draw. The set of respondents and nonrespondents among the sample draw households is the analysis pool for the recursive partitioning model that will produce the variable for raking. NCHS has begun exploring the utility of these procedures in the construction of sample weights for the NHIS annual data release.

Conclusion

When response rates were relatively high, it was sufficient to adjust for nonresponse geographically and post-stratify exclusively to age/sex/race/ethnicity (as NHIS did until the 2019 redesign) (9). As response rates have dropped (10), resulting final samples have required additional effort in adjusting the weights to adequately represent the original population. Traditionally, calibration has relied primarily on population control totals from external sources such as U.S. Census population estimates. However, the current analysis suggests that raking variables are not necessarily so limited, as long as they are available for both respondents and nonrespondents (i.e., present in or can be appended to the sample frame) and are related to both response propensity and key survey outcomes. Recursive partitioning models can harness many sample frame variables to create a single composite measure that captures complex interactions among variables related to nonresponse and outcome measures of interest. This may include sample frame variables not already controlled in the raking process, offering an efficient process for improving bias reduction.

In this study, adding the RPM-derived variable to the NHIS-Teen raking dimensions improved bias reduction by more than 4% without appreciably inflating the variance. Using RPM methods to derive this variable allowed for the efficient integration of multiple sample frame variables and the effective targeting of weight adjustments where correlations with response propensity and correlations with survey outcomes of interest coincide. While this approach has demonstrated clear utility for followback surveys, its promising performance suggests potential for broader application to standalone surveys such as the NHIS.

References

1. Bramlett MD, Dahlhamer JM and Bose J. Weighting procedures and bias assessment for the 2020 National Health Interview Survey. Hyattsville, MD: National Center for Health Statistics. September 2021. Available from: Weighting Procedures and Bias Assessment for the 2020 National Health Interview Survey (cdc.gov).

2. Zablotsky B, Black LI, Ng AE, Bose J, Jones J, Maitland A, and Blumberg SJ. Methodology Report for the 2021-2023 National Health Interview Survey – Teen 30-month File. Hyattsville, MD: National Center for Health Statistics. September 2024. Available from: Methodology Report for the 2021-2023 National Health Interview Survey - Teen 30-month File.

3. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2021. Available from https://www.R-project.org/.

4. Hothorn T, Zeileis A. partykit: A modular toolkit for recursive partytitioning in R. *Journal of Machine Learning Research* 16: 3905-9, 2015.

5. Valliant R, Dever JA, Kreuter F. Practical Tools for Designing and Weighting Survey Samples. New York: Springer. 2013.

6. Bramlett MD, Black LI, Dahlhamer JM and Zablotsky B. Weighting procedures and bias assessment for the 2021-2023 National Health Interview Survey – Teen 30- month file. Hyattsville, MD: National Center for Health Statistics. September 2024. Available from: Weighting Procedures and Bias Assessment for the 2021-2023 National Health Interview Survey – Teen 30-month File.

7. Bramlett MD, Black L, Zablotsky B, and Dahlhamer JM. Weighting procedures and bias assessment for the 2021-2022 National Health Interview Survey – Teen 18- month file. Hyattsville, MD: National Center for Health Statistics. September 2023. Available from: Weighting Procedures and Bias Assessment for the 2021-2022 National Health Interview Survey – Teen 18-month File.

8. Biemer PP. Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly* 74(5): 817-848, 2010. https://doi.org/10.1093/poq/nfq058

9. 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. Available from: New Procedures for Nonresponse Adjustments to the 2019 National Health Interview Survey Sampling Weights (cdc.gov).

10. Williams D and Brick JM. Trends in U.S. face-to-face household survey nonresponse and level of effort. *Journal of Survey Statistics and Methodology* 6: 186-211, 2018. https://doi.org/10.1093/jssam/smx019

Detailed Tables

Detailed tables begin on next page.

Table 1. Non-reinterview bias in National Health Interview Survey – Teen estimates, by weight

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight A (2)	Bias using weight A (unadjusted weight)	Teen estimates using weight B (3)	Bias using weight B (final Teen weight)	Teen estimates using weight C (4)	Bias using weight C (omitting RPM)	Bias using weight D (with RPM)	Bias using weight E (omitting RPM)
Sociodemographic measures	Percent (standard error)		|(2)-(1)|	% (SE)	|(3)-(1)|	% (SE)	|(4)-(1)|	|DIFF|	|DIFF|
Age of teen: 12 years	16.1 (0.53)	13.9 (0.96)	2.1	16.1 (1.14)	0.0	16.1 (1.12)	0.0	0.1	0.3
13 years	15.6 (0.51)	15.6 (1.06)	0.1	15.6 (1.12)	0.0	15.6 (1.11)	0.0	0.3	0.3
14 years	17.8 (0.54)	18.6 (1.03)	0.7	17.8 (1.11)	0.0	17.8 (1.08)	0.0	0.5	0.6
15 years	17.1 (0.51)	18.6 (1.05)	1.5	17.1 (1.07)	0.0	17.1 (1.06)	0.0	1.6	1.6
16 years	17.5 (0.51)	18.0 (0.96)	0.4	17.5 (1.03)	0.0	17.5 (1.01)	0.0	0.3	0.4
17 years	15.9 (0.47)	15.5 (0.89)	0.4	15.9 (0.96)	0.0	15.9 (0.95)	0.0	1.2	1.2
Sex of teen: male	51.3 (0.72)	48.2 (1.38)	3.1	51.3 (1.48)	0.0	51.3 (1.46)	0.0	0.0	0.0
Race/ethnicity of teen: Hispanic	26.5 (0.71)	21.8 (1.15)	4.7	26.5 (1.36)	0.0	26.5 (1.35)	0.0	0.0	0.0
Black only, non-Hispanic	15.2 (0.58)	12.0 (0.95)	3.2	15.2 (1.20)	0.0	15.2 (1.19)	0.0	0.0	0.0
Asian only, non-Hispanic	5.9 (0.28)	7.1 (0.57)	1.3	5.9 (0.51)	0.0	5.9 (0.50)	0.0	0.0	0.0
Other races, non-Hispanic	52.5 (0.78)	59.0 (1.32)	6.6	52.5 (1.44)	0.0	52.5 (1.42)	0.0	0.0	0.0
Teen born in United States	93.8 (0.35)	93.7 (0.63)	0.0	93.2 (0.76)	0.6	93.1 (0.76)	0.6	0.6	0.6
Residence owned/mortgaged	68.6 (0.71)	75.7 (1.19)	7.1	68.6 (1.46)	0.0	68.6 (1.44)	0.0	0.0	0.0
In Metropolitan Statistical Area	87.0 (1.17)	87.5 (1.43)	0.6	87.0 (1.50)	0.0	87.0 (1.49)	0.0	0.0	0.0
Census region: Northeast	15.6 (0.56)	15.0 (1.00)	0.5	15.6 (0.90)	0.0	15.6 (0.88)	0.0	0.0	0.0
Midwest	20.7 (0.67)	23.3 (1.24)	2.6	20.7 (0.97)	0.0	20.7 (0.98)	0.0	0.0	0.0
South	39.5 (0.82)	36.3 (1.46)	3.2	39.5 (1.33)	0.0	39.5 (1.32)	0.0	0.0	0.0
West	24.3 (0.76)	25.4 (1.37)	1.1	24.3 (1.22)	0.0	24.3 (1.22)	0.0	0.0	0.0
Age, SCR: 18-34	8.4 (0.43)	5.7 (0.70)	2.6	7.0 (0.85)	1.4	7.0 (0.84)	1.4	1.3	1.3
35-44	42.2 (0.72)	41.2 (1.36)	1.1	42.7 (1.45)	0.5	42.8 (1.44)	0.5	0.8	0.9
45 and older	49.4 (0.71)	53.1 (1.29)	3.7	50.3 (1.36)	0.9	50.3 (1.35)	0.9	0.5	0.4
Race/ethnicity, SCR: Hispanic	26.2 (0.72)	21.7 (1.15)	4.5	26.4 (1.37)	0.2	26.4 (1.35)	0.2	0.2	0.2
White only, non-Hispanic	50.4 (0.78)	57.1 (1.34)	6.7	50.4 (1.43)	0.0	50.3 (1.42)	0.1	0.1	0.1
Black only, non-Hispanic	11.9 (0.55)	9.1 (0.87)	2.8	12.0 (1.16)	0.1	11.9 (1.13)	0.0	0.1	0.0
Other races, non-Hispanic	11.5 (0.47)	12.1 (0.86)	0.6	11.3 (0.93)	0.2	11.4 (0.92)	0.1	0.2	0.1

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight A (2)	Bias using weight A (unadjusted weight)	Teen estimates using weight B (3)	Bias using weight B (final Teen weight)	Teen estimates using weight C (4)	Bias using weight C (omitting RPM)	Bias using weight D (with RPM)	Bias using weight E (omitting RPM)
Education (Ed), SCR: <high school	11.1 (0.48)	9.5 (0.84)	1.6	12.0 (1.07)	0.9	12.1 (1.06)	1.0	1.2	1.4
High school graduate/GED	20.0 (0.58)	15.9 (0.95)	4.1	19.1 (1.19)	0.9	18.9 (1.17)	1.1	1.0	1.0
Some college/Associate's	28.0 (0.61)	24.6 (1.10)	3.4	27.7 (1.24)	0.3	27.5 (1.23)	0.4	0.1	0.1
Bachelor's or more	41.0 (0.73)	50.0 (1.30)	9.1	41.3 (1.35)	0.3	41.5 (1.33)	0.5	0.3	0.5
Highest Ed in family: <high school	6.0 (0.37)	4.5 (0.60)	1.5	5.8 (0.81)	0.2	5.9 (0.80)	0.1	0.0	0.1
High school graduate/GED	15.8 (0.53)	12.2 (0.91)	3.6	16.0 (1.20)	0.2	15.7 (1.17)	0.0	0.3	0.2
Some college/Associate's	28.6 (0.62)	24.4 (1.11)	4.2	29.3 (1.30)	0.6	29.3 (1.30)	0.7	0.9	1.4
Bachelor's or more	49.6 (0.74)	58.9 (1.27)	9.3	49.0 (1.40)	0.6	49.1 (1.38)	0.6	1.2	1.7
One parent in household (hh)	35.8 (0.72)	29.3 (1.33)	6.5	34.9 (1.53)	0.9	33.1 (1.49)	2.7	0.7	2.5
Income from: wages/salary/tips/self	93.4 (0.36)	94.6 (0.69)	1.2	93.2 (0.91)	0.2	93.3 (0.90)	0.1	0.3	0.2
Interest/investments	23.1 (0.63)	28.6 (1.20)	5.5	24.0 (1.13)	0.9	24.1 (1.13)	1.0	0.5	0.5
Social Security/Railroad	9.9 (0.40)	8.9 (0.72)	1.0	9.6 (0.81)	0.3	9.3 (0.78)	0.6	0.0	0.3
Social Security Disability	7.8 (0.39)	6.6 (0.69)	1.2	7.7 (0.81)	0.1	7.7 (0.80)	0.1	0.0	0.0
Retirement/survivor/disability pension	6.3 (0.33)	6.4 (0.61)	0.1	5.9 (0.59)	0.4	5.7 (0.58)	0.5	0.1	0.2
Other sources of income	15.6 (0.55)	14.2 (0.99)	1.4	14.9 (1.09)	0.7	14.6 (1.07)	1.0	0.6	0.8
Poverty status: in poverty	14.3 (0.55)	10.0 (0.93)	4.3	14.3 (1.32)	0.0	14.3 (1.30)	0.0	0.9	0.9
1.00-1.99*FPL	21.5 (0.62)	20.1 (1.10)	1.3	21.5 (1.21)	0.0	21.5 (1.21)	0.0	3.1	2.9
2.00-3.99*FPL	27.9 (0.62)	27.5 (1.22)	0.4	27.9 (1.25)	0.0	27.9 (1.23)	0.0	0.3	0.2
4.00* or higher FPL	36.3 (0.73)	42.4 (1.35)	6.1	36.3 (1.35)	0.0	36.3 (1.32)	0.0	1.9	1.7
Free/reduced meals past 12 months	52.9 (0.76)	50.3 (1.34)	2.7	53.3 (1.46)	0.4	53.3 (1.43)	0.4	1.2	1.2
Food stamps received past 12 months	20.8 (0.64)	17.1 (1.14)	3.6	21.6 (1.39)	0.8	21.4 (1.36)	0.6	1.1	0.9
WIC benefits received past 12 months	5.5 (0.42)	4.6 (0.73)	0.9	5.8 (0.92)	0.3	5.8 (0.92)	0.3	0.5	0.5
Could not afford balanced meals, past 30 (p30) days	11.8 (0.49)	10.2 (0.92)	1.7	11.9 (1.10)	0.1	12.0 (1.10)	0.1	0.5	0.5
Food did not last, p30 days	13.4 (0.51)	11.4 (0.96)	1.9	13.3 (1.15)	0.0	13.5 (1.16)	0.1	0.2	0.4
Worried food would run out, p30 days	16.0 (0.56)	13.5 (1.01)	2.5	16.0 (1.21)	0.0	16.1 (1.22)	0.1	0.4	0.5
Food security: secure	89.3 (0.46)	90.4 (0.88)	1.1	88.7 (1.06)	0.6	88.6 (1.06)	0.7	0.8	0.9
Insecure	6.5 (0.37)	6.1 (0.71)	0.4	7.4 (0.90)	0.9	7.4 (0.90)	0.9	0.9	1.0
Very insecure	4.2 (0.30)	3.5 (0.53)	0.7	3.9 (0.63)	0.3	3.9 (0.61)	0.2	0.1	0.1

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight A (2)	Bias using weight A (unadjusted weight)	Teen estimates using weight B (3)	Bias using weight B (final Teen weight)	Teen estimates using weight C (4)	Bias using weight C (omitting RPM)	Bias using weight D (with RPM)	Bias using weight E (omitting RPM)
Number of adults in hh work fulltime: 0	11.7 (0.47)	9.7 (0.79)	2.0	11.7 (0.99)	0.0	11.5 (0.97)	0.2	0.2	0.0
1	44.0 (0.72)	43.9 (1.29)	0.1	44.9 (1.36)	0.9	44.8 (1.35)	0.8	1.3	1.1
2 or more	44.3 (0.70)	46.4 (1.30)	2.1	43.4 (1.37)	0.9	43.7 (1.35)	0.6	1.5	1.1
Number of adults in hh working: 0	6.1 (0.33)	4.5 (0.56)	1.7	5.7 (0.75)	0.5	5.4 (0.72)	0.7	0.5	0.7
1	33.8 (0.67)	30.7 (1.20)	3.1	33.4 (1.34)	0.5	33.2 (1.31)	0.6	0.0	0.2
2 or more	60.1 (0.71)	64.8 (1.28)	4.8	61.0 (1.44)	0.9	61.4 (1.41)	1.3	0.5	0.9
Duration in residence: 1-3 years	31.9 (0.69)	27.2 (1.23)	4.7	30.6 (1.40)	1.3	30.5 (1.38)	1.5	1.2	1.5
4-10 years	35.0 (0.67)	37.0 (1.27)	2.0	36.7 (1.37)	1.7	36.7 (1.35)	1.7	1.8	1.9
11 or more years	33.1 (0.65)	35.8 (1.23)	2.7	32.6 (1.28)	0.5	32.8 (1.26)	0.3	0.6	0.4
1 or more working landline phones	22.3 (0.59)	24.9 (1.18)	2.6	23.4 (1.20)	1.1	23.4 (1.18)	1.1	1.1	1.0
Health Status indicators									
Excellent or very good health	83.4 (0.50)	84.7 (0.95)	1.4	83.4 (1.08)	0.0	83.3 (1.07)	0.0	0.3	0.3
Anxiety frequency: never	41.3 (0.70)	34.1 (1.27)	7.2	41.7 (1.50)	0.4	41.2 (1.48)	0.1	0.4	0.0
A few times per year or monthly	38.5 (0.68)	43.6 (1.32)	5.1	37.7 (1.36)	0.8	38.3 (1.36)	0.2	0.8	0.4
Weekly or daily	20.3 (0.57)	22.4 (1.14)	2.1	20.6 (1.13)	0.4	20.5 (1.13)	0.3	0.4	0.3
Depression frequency: never	59.6 (0.74)	53.7 (1.33)	5.8	57.9 (1.42)	1.7	57.7 (1.40)	1.8	1.6	1.7
A few times per year or monthly	24.1 (0.63)	27.6 (1.18)	3.5	25.3 (1.22)	1.2	25.6 (1.21)	1.5	1.0	1.3
Weekly or daily	16.4 (0.53)	18.7 (1.05)	2.3	16.8 (1.05)	0.5	16.7 (1.03)	0.3	0.6	0.4
Ever had asthma	14.8 (0.50)	12.7 (0.86)	2.2	13.0 (0.94)	1.8	13.1 (0.94)	1.8	1.7	1.7
Current asthma	8.4 (0.38)	6.6 (0.63)	1.8	6.4 (0.64)	2.0	6.4 (0.65)	1.9	1.9	1.9
Any difficulty seeing	6.7 (0.38)	6.7 (0.67)	0.0	7.1 (0.74)	0.4	7.0 (0.73)	0.3	0.5	0.4
Wears glasses or contact lenses	38.2 (0.66)	41.6 (1.31)	3.4	41.0 (1.38)	2.8	41.0 (1.37)	2.8	2.9	2.9
Any learning difficulty	10.9 (0.45)	8.4 (0.76)	2.6	9.0 (0.83)	1.9	8.7 (0.80)	2.2	1.7	2.0
Any difficulty remembering	9.5 (0.43)	8.2 (0.74)	1.3	8.8 (0.80)	0.7	8.4 (0.78)	1.0	0.6	0.9
Disability status: has any disabilities	14.5 (0.50)	13.4 (0.97)	1.1	13.2 (0.98)	1.3	13.0 (0.97)	1.5	1.1	1.2
Ever diagnosed with ADD/ADHD	15.5 (0.50)	13.4 (0.90)	2.1	12.9 (0.91)	2.6	12.6 (0.89)	2.8	2.5	2.7
Currently (Curr) has ADD/ADHD	13.8 (0.47)	12.3 (0.86)	1.6	11.6 (0.85)	2.2	11.4 (0.83)	2.4	2.1	2.3
Ever had a learning disability	10.3 (0.42)	7.4 (0.70)	2.9	8.4 (0.82)	1.8	8.1 (0.79)	2.2	1.8	2.1

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight A (2)	Bias using weight A (unadjusted weight)	Teen estimates using weight B (3)	Bias using weight B (final Teen weight)	Teen estimates using weight C (4)	Bias using weight C (omitting RPM)	Bias using weight D (with RPM)	Bias using weight E (omitting RPM)
Curr has learning disability	8.6 (0.39)	6.0 (0.65)	2.6	7.0 (0.76)	1.6	6.6 (0.73)	2.1	1.5	2.0
Curr has autism spectrum disorder	4.0 (0.27)	2.8 (0.46)	1.1	3.0 (0.49)	0.9	2.9 (0.48)	1.1	0.9	1.0
Curr any other developmental disability	3.3 (0.26)	2.3 (0.44)	1.0	2.9 (0.55)	0.4	2.8 (0.54)	0.5	0.4	0.4
Curr any developmental disability	19.7 (0.56)	17.5 (1.05)	2.2	17.7 (1.11)	2.0	17.2 (1.08)	2.5	1.8	2.2
Injury/illness school days missed: 0	39.3 (0.69)	37.7 (1.36)	1.6	40.0 (1.48)	0.7	40.1 (1.47)	0.8	0.8	0.7
1-5	40.3 (0.67)	43.5 (1.33)	3.2	41.9 (1.43)	1.6	41.7 (1.41)	1.4	1.4	1.3
6-10	11.8 (0.46)	11.5 (0.86)	0.4	10.6 (0.85)	1.2	10.8 (0.86)	1.0	1.2	1.1
11 or more	8.6 (0.41)	7.4 (0.70)	1.2	7.4 (0.77)	1.1	7.4 (0.77)	1.1	0.9	0.9
Ever receive special education services	18.7 (0.53)	16.0 (1.00)	2.7	17.3 (1.12)	1.4	16.7 (1.08)	2.0	1.4	1.9
Curr receive special education services	12.3 (0.45)	9.1 (0.75)	3.1	11.11 (0.91)	1.2	10.3 (0.87)	2.0	1.1	1.9
SLE/Social Support/Behavioral Problems									
Lived with anyone with alcohol/drug problem	11.8 (0.48)	11.0 (0.85)	0.9	11.9 (0.96)	0.1	11.6 (0.95)	0.2	0.3	0.0
Lived with parent who was incarcerated	8.0 (0.40)	6.7 (0.74)	1.3	7.6 (0.84)	0.4	7.4 (0.81)	0.7	0.1	0.4
Lived with mentally ill/depressed person	11.3 (0.49)	12.4 (0.91)	1.1	12.4 (0.97)	1.1	12.4 (0.96)	1.1	1.2	1.2
Lifetime of being put down by adults in home	6.8 (0.37)	5.4 (0.59)	1.4	5.6 (0.66)	1.2	5.5 (0.64)	1.3	1.1	1.1
Treated/judged based on race/ethnicity	8.2 (0.39)	7.7 (0.68)	0.5	8.4 (0.80)	0.2	8.5 (0.79)	0.3	0.1	0.2
Victim of/witnessed violence	8.3 (0.42)	7.9 (0.77)	0.4	8.7 (0.87)	0.4	8.6 (0.86)	0.3	0.5	0.5
Experienced two or more SLEs	13.7 (0.52)	12.3 (0.91)	1.4	13.2 (1.02)	0.5	13.0 (1.01)	0.7	0.3	0.4
Community support	88.3 (0.49)	91.3 (0.73)	3.1	89.7 (0.89)	1.5	89.7 (0.89)	1.5	1.3	1.4
Basic needs not met	4.4 (0.31)	4.0 (0.55)	0.4	4.3 (0.63)	0.1	4.2 (0.61)	0.2	0.0	0.1
Any difficulty with changes in routine	22.2 (0.58)	23.2 (1.11)	1.0	22.2 (1.13)	0.1	22.1 (1.12)	0.2	0.2	0.3
Any difficulty controlling behavior	13.3 (0.46)	12.5 (0.92)	0.8	12.8 (0.98)	0.5	12.6 (0.97)	0.7	0.5	0.6
Any difficulty concentrating	7.4 (0.37)	7.5 (0.70)	0.1	7.7 (0.76)	0.3	7.6 (0.76)	0.2	0.4	0.3
Any difficulty making friends	13.2 (0.46)	14.9 (0.90)	1.7	14.7 (0.98)	1.4	14.5 (0.96)	1.3	1.6	1.5
Health Care Service Use indicators									
Saw doctor, past 12 months	92.3 (0.39)	92.2 (0.73)	0.1	91.5 (0.87)	0.9	91.6 (0.85)	0.7	0.7	0.5
Usual place: doctor office/health center	91.1 (0.45)	92.2 (0.78)	1.1	90.5 (0.99)	0.6	90.7 (0.97)	0.4	0.5	0.2

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight A (2)	Bias using weight A (unadjusted weight)	Teen estimates using weight B (3)	Bias using weight B (final Teen weight)	Teen estimates using weight C (4)	Bias using weight C (omitting RPM)	Bias using weight D (with RPM)	Bias using weight E (omitting RPM)
Telehealth visits, past 12 months	18.1 (0.52)	19.1 (1.03)	1.0	18.3 (1.07)	0.2	18.2 (1.06)	0.2	0.0	0.1
Urgent care visits, past 12 months: 0	71.8 (0.64)	70.4 (1.29)	1.4	71.3 (1.35)	0.5	71.3 (1.33)	0.5	0.5	0.4
1 visit	16.0 (0.51)	16.5 (0.99)	0.5	16.1 (1.06)	0.1	16.1 (1.04)	0.1	0.1	0.1
2 or more visits	12.2 (0.47)	13.1 (0.94)	0.9	12.6 (0.96)	0.4	12.6 (0.96)	0.4	0.5	0.5
Wellness visits, past 12 months	89.5 (0.43)	89.2 (0.83)	0.3	89.6 (0.82)	0.1	89.6 (0.82)	0.1	0.2	0.2
ER visits, past 12 months: 0	86.8 (0.46)	87.8 (0.84)	1.1	87.0 (0.97)	0.8	87.1 (0.95)	0.7	0.6	0.6
1 visit	9.1 (0.38)	8.0 (0.70)	1.1	8.3 (0.80)	0.6	8.4 (0.79)	0.4	0.7	0.5
2 or more visits	4.1 (0.27)	4.1 (0.52)	0.0	4.7 (0.60)	0.6	4.5 (0.58)	0.4	0.7	0.5
Received flu shot, past 12 months (mos)	43.0 (0.73)	48.0 (1.39)	5.0	43.1 (1.46)	0.1	44.1 (1.45)	1.1	0.1	0.8
Medications for emotions, past 12 mos	13.3 (0.47)	13.6 (0.94)	0.3	13.1 (0.94)	0.2	12.7 (0.92)	0.6	0.3	0.6
Received mental health care, past 12 months	17.3 (0.52)	19.3 (1.03)	2.0	17.8 (1.02)	0.5	17.7 (1.01)	0.4	0.6	0.4
Ever diagnosed with COVID	34.2 (0.68)	33.9 (1.32)	0.3	33.4 (1.38)	0.8	33.2 (1.37)	0.9	0.8	1.0
Received COVID vaccination	59.2 (0.77)	69.5 (1.33)	10.3	59.7 (1.51)	0.5	61.5 (1.49)	2.3	0.5	1.7
Prescribed medication, past 12 months	38.4 (0.69)	40.0 (1.34)	1.5	38.7 (1.38)	0.3	38.5 (1.37)	0.1	0.2	0.0
Health Care Access/Insurance Coverage indicators									
Unable to pay medical bills, past 12 months	7.9 (0.40)	7.0 (0.76)	0.9	8.0 (0.86)	0.1	8.1 (0.86)	0.1	0.3	0.3
Worried about paying medical bills	32.4 (0.66)	32.2 (1.20)	0.2	33.0 (1.33)	0.5	33.1 (1.32)	0.7	0.4	0.5
Delay/defer medical care due to cost	2.7 (0.22)	2.3 (0.39)	0.2	2.6 (0.45)	0.1	2.6 (0.45)	0.1	0.1	0.1
Health insurance coverage: private	57.5 (0.74)	64.1 (1.31)	6.7	57.4 (1.48)	0.1	57.4 (1.46)	0.1	2.1	2.0
Public	38.2 (0.72)	32.7 (1.27)	5.6	38.3 (1.44)	0.1	38.3 (1.43)	0.0	2.3	2.3
None	4.3 (0.28)	3.2 (0.44)	1.1	4.3 (0.63)	0.0	4.3 (0.62)	0.0	0.2	0.3

NOTES: RPM is the raking variable derived from the recursive partitioning model. SE is standard error. DIFF is difference between estimates not shown and (1). SCR is Sample Child respondent. GED is General Equivalence Degree. FPL is Federal Poverty Level. WIC is Women, Infants & Children; ADD/ADHD is attention-deficit disorder/attention-deficit/hyperactivity disorder; SLE is Stressful Life Events. ER is emergency room .
SOURCE: National Center for Health Statistics, National Health Interview Survey, 2021-2023; National Health Interview Survey-Teen 30-Month File.

Table 2. Sum of non-reinterview biases and reduction in bias attributable to weighting adjustments, by weight

Sum of biases	293.2	73.1	82.9	87.9	96.4	102.3	122.7
Percent bias reduction relative to weight A	0	75.1	71.7	70.0	67.1	65.1	58.2

NOTES: RPM is the raking variable derived from the recursive partitioning model. SC is Sample Child.
SOURCE: National Center for Health Statistics, National Health Interview Survey – Teen 30-month file.

Table 3. Sum of non-reinterview biases and percent bias reduction, by weight and indicator subgroup

	Sum of biases (percent reduction relative to weight A)						
Total	293.2 (0)	73.1 (75.1)	82.9 (71.7)	87.9 (70.0)	96.4 (67.1)	102.3 (65.1)	122.7 (58.2)
Raking variables	69.2 (0)	0.2 (99.7)	0.1 (99.9)	14.8 (78.6)	14.7 (78.8)	23.9 (65.5)	30.1 (56.5)
RPM components	75.2 (0)	6.4 (91.5)	10.5 (86.0)	7.0 (90.7)	11.4 (84.8)	11.7 (84.4)	16.7 (77.8)
Other indicators	148.8 (0)	66.5 (55.3)	72.3 (51.4)	66.1 (55.6)	70.3 (52.8)	66.7 (55.2)	75.9 (49.0)

NOTES:RPM is the raking variable derived from the recursive partitioning model. SC is Sample Child. Raking variables are age, sex, race/ethnicity, region, metropolitan statistical area status, housing tenure, income as poverty status, and insurance; RPM components are receipt of COVID-19 vaccination, frequency of anxiety, highest education in household, number of parents, any flu vaccination, current receipt of special education services and SC respondent race/ethnicity; the operational auxiliary variable recording whether a householder expressed time constraints was an additional RPM component but was not considered a survey outcome or demographic or stratifying variable and was not included in the list of indicators assessed for bias reduction. Sample Adult (SA) respondent race/ethnicity was an RPM component while SC race/ethnicity was included in the bias assessment; in most cases the SA respondent is the same person as the SC respondent (most often, a parent of the teen), and even when not, their races/ethnicities are highly correlated, so for the purposes of this table, bias in SC respondent race/ethnicity stands in for bias in SA respondent race/ethnicity. Similarly, while receipt of COVID-19 vaccination was included in the bias assessment, the specific number of COVID-19 vaccinations was an RPM component; for this table, the bias in receipt stands in for the bias in number of COVID-19 vaccinations. Other indicators include all remaining indicators in Table 1.
SOURCE: National Center for Health Statistics, National Health Interview Survey – Teen 30-month file.

Table 4. Relative standard error averaged across bias assessment indicators, by weight

Variance metric	Unadjusted Sample Child weight (A)	Final teen weight (B)	Final teen weight omitting RPM (C)	Teen weight raked to original SC totals and RPM (D)	Teen weight raked to original SC totals (E)	Teen weight raked to RPM (F)	Weight A raked to RPM (G)
Average relative standard error	0.056	0.058	0.058	0.059	0.058	0.058	0.057

NOTES: RPM is the raking variable derived from the recursive partitioning model. SC is Sample Child.
SOURCE: National Center for Health Statistics, National Health Interview Survey – Teen 30-month file.

Appendix

This appendix contains Table I, which shows the estimates, standard errors, and biases for weights D, E, F, and G in addition to the gold standard estimates. The estimates and standard errors for weights D and E and the estimates, standard errors, and biases for weights F and G were omitted from Table 1 for brevity.

Table I begins on next page.

Table I. Non-reinterview bias in National Health Interview Survey – Teen estimates, by weight D through G

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight D (2)	Bias using weight D	Teen estimates using weight E (3)	Bias using weight E	Teen estimates using weight F (4)	Bias using weight F	Teen estimates using weight G (5)	Bias using weight G
Sociodemographic measures	Percent (standard error)		|(2)-(1)|	% (SE)	|(3)-(1)|	% (SE)	|(4)-(1)|	% (SE)	|(5)-(1)|
Age of teen: 12 years	16.1 (0.53)	15.9 (1.15)	0.1	15.8 (1.13)	0.3	16.0 (1.13)	0.1	14.5 (1.04)	1.6
13 years	15.6 (0.51)	15.3 (1.12)	0.3	15.3 (1.12)	0.3	15.6 (1.12)	0.1	15.6 (1.10)	0.1
14 years	17.8 (0.54)	18.3 (1.13)	0.5	18.4 (1.12)	0.6	18.5 (1.12)	0.6	18.7 (1.12)	0.9
15 years	17.1 (0.51)	18.6 (1.15)	1.6	18.7 (1.14)	1.6	18.4 (1.13)	1.3	18.6 (1.10)	1.5
16 years	17.5 (0.51)	17.2 (1.00)	0.3	17.2 (0.98)	0.4	17.1 (1.00)	0.4	17.6 (1.00)	0.1
17 years	15.9 (0.47)	14.7 (0.89)	1.2	14.7 (0.88)	1.2	14.5 (0.88)	1.4	15.0 (0.90)	0.9
Sex of teen: male	51.3 (0.72)	51.3 (1.49)	0.0	51.3 (1.48)	0.0	49.4 (1.48)	2.0	49.3 (1.46)	2.0
Race/ethnicity of teen: Hispanic	26.5 (0.71)	26.5 (1.38)	0.0	26.5 (1.36)	0.0	25.6 (1.33)	0.9	24.0 (1.24)	2.5
Black only, non-Hispanic	15.2 (0.58)	15.2 (1.21)	0.0	15.2 (1.19)	0.0	13.9 (1.12)	1.3	13.6 (1.08)	1.6
Asian only, non-Hispanic	5.9 (0.28)	5.9 (0.50)	0.0	5.9 (0.50)	0.0	5.8 (0.50)	0.0	6.1 (0.53)	0.2
Other races, non-Hispanic	52.5 (0.78)	52.5 (1.45)	0.0	52.5 (1.43)	0.0	54.7 (1.43)	2.2	56.3 (1.37)	3.9
Teen born in United States	93.8 (0.35)	93.2 (0.76)	0.6	93.2 (0.76)	0.6	93.5 (0.71)	0.2	93.7 (0.68)	0.1
Residence owned/mortgaged	68.6 (0.71)	68.6 (1.46)	0.0	68.6 (1.45)	0.0	71.0 (1.40)	2.4	71.9 (1.36)	3.3
In Metropolitan Statistical Area	87.0 (1.17)	87.0 (1.50)	0.0	87.0 (1.49)	0.0	86.2 (1.56)	0.7	86.2 (1.56)	0.8
Census region: Northeast	15.6 (0.56)	15.6 (0.90)	0.0	15.6 (0.89)	0.0	14.3 (0.84)	1.2	14.3 (0.82)	1.3
Midwest	20.7 (0.67)	20.7 (0.99)	0.0	20.7 (0.99)	0.0	21.3 (1.00)	0.7	22.4 (1.02)	1.7
South	39.5 (0.82)	39.5 (1.34)	0.0	39.5 (1.33)	0.0	40.1 (1.33)	0.6	39.2 (1.30)	0.3
West	24.3 (0.76)	24.3 (1.24)	0.0	24.3 (1.23)	0.0	24.2 (1.22)	0.0	24.2 (1.21)	0.1
Age, SCR: 18-34	8.4 (0.43)	7.1 (0.86)	1.3	7.1 (0.86)	1.3	7.0 (0.85)	1.3	6.7 (0.80)	1.7
35-44	42.2 (0.72)	43.0 (1.46)	0.8	43.2 (1.46)	0.9	43.1 (1.45)	0.9	43.1 (1.42)	0.9
45 and older	49.4 (0.71)	49.9 (1.37)	0.5	49.8 (1.36)	0.4	49.9 (1.35)	0.5	50.2 (1.33)	0.8
Race/ethnicity, SCR: Hispanic	26.2 (0.72)	26.4 (1.38)	0.2	26.4 (1.36)	0.2	25.5 (1.33)	0.7	23.9 (1.24)	2.3
White only, non-Hispanic	50.4 (0.78)	50.3 (1.44)	0.1	50.3 (1.43)	0.1	52.5 (1.43)	2.1	54.2 (1.37)	3.8
Black only, non-Hispanic	11.9 (0.55)	12.0 (1.16)	0.1	11.9 (1.14)	0.0	11.0 (1.08)	0.9	10.7 (1.03)	1.2
Other races, non-Hispanic	11.5 (0.47)	11.3 (0.92)	0.2	11.4 (0.93)	0.1	11.1 (0.90)	0.4	11.2 (0.88)	0.3

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight D (2)	Bias using weight D	Teen estimates using weight E (3)	Bias using weight E	Teen estimates using weight F (4)	Bias using weight F	Teen estimates using weight G (5)	Bias using weight G
Education (Ed), SCR: <high school	11.1 (0.48)	12.3 (1.11)	1.2	12.5 (1.09)	1.4	11.9 (1.06)	0.8	11.0 (0.99)	0.1
High school graduate/GED	20.0 (0.58)	19.0 (1.19)	1.0	19.0 (1.18)	1.0	19.0 (1.17)	1.0	18.5 (1.13)	1.5
Some college/Associate's	28.0 (0.61)	28.0 (1.25)	0.1	28.1 (1.25)	0.1	28.2 (1.24)	0.3	27.5 (1.20)	0.4
Bachelor's or more	41.0 (0.73)	40.7 (1.34)	0.3	40.5 (1.33)	0.5	40.9 (1.33)	0.0	43.0 (1.34)	2.0
Highest Ed in family: <high school	6.0 (0.37)	6.0 (0.84)	0.0	6.1 (0.84)	0.1	5.7 (0.78)	0.3	5.3 (0.73)	0.7
High school graduate/GED	15.8 (0.53)	16.0 (1.21)	0.3	15.9 (1.18)	0.2	15.7 (1.18)	0.0	14.8 (1.11)	0.9
Some college/Associate's	28.6 (0.62)	29.6 (1.32)	0.9	30.0 (1.32)	1.4	29.9 (1.29)	1.2	28.6 (1.24)	0.1
Bachelor's or more	49.6 (0.74)	48.4 (1.40)	1.2	48.0 (1.38)	1.7	48.8 (1.38)	0.8	51.3 (1.34)	1.7
One parent in household (hh)	35.8 (0.72)	35.1 (1.55)	0.7	33.3 (1.51)	2.5	34.5 (1.53)	1.3	34.3 (1.51)	1.5
Income from: wages/salary/tips/self	93.4 (0.36)	93.1 (0.92)	0.3	93.2 (0.91)	0.2	93.4 (0.87)	0.1	93.6 (0.82)	0.3
Interest/investments	23.1 (0.63)	23.6 (1.11)	0.5	23.6 (1.11)	0.5	24.1 (1.13)	1.0	24.6 (1.12)	1.5
Social Security/Railroad	9.9 (0.40)	9.9 (0.84)	0.0	9.6 (0.82)	0.3	9.8 (0.83)	0.0	9.7 (0.80)	0.1
Social Security Disability	7.8 (0.39)	7.8 (0.81)	0.0	7.7 (0.81)	0.0	7.6 (0.79)	0.2	7.4 (0.77)	0.4
Retirement/survivor/disability pension	6.3 (0.33)	6.2 (0.64)	0.1	6.1 (0.63)	0.2	6.3 (0.63)	0.0	6.4 (0.64)	0.2
Other sources of income	15.6 (0.55)	15.0 (1.09)	0.6	14.7 (1.08)	0.8	15.0 (1.07)	0.6	15.1 (1.05)	0.4
Poverty status: in poverty	14.3 (0.55)	13.3 (1.28)	0.9	13.4 (1.25)	0.9	12.9 (1.23)	1.4	11.9 (1.13)	2.3
1.00-1.99*FPL	21.5 (0.62)	24.5 (1.32)	3.1	24.3 (1.31)	2.9	24.0 (1.28)	2.5	23.2 (1.23)	1.7
2.00-3.99*FPL	27.9 (0.62)	27.7 (1.26)	0.3	27.7 (1.24)	0.2	28.0 (1.25)	0.1	28.2 (1.25)	0.3
4.00* or higher FPL	36.3 (0.73)	34.5 (1.31)	1.9	34.6 (1.29)	1.7	35.2 (1.31)	1.1	36.7 (1.33)	0.3
Free/reduced meals past 12 months	52.9 (0.76)	54.1 (1.47)	1.2	54.1 (1.44)	1.2	53.6 (1.46)	0.6	53.3 (1.42)	0.4
Food stamps received past 12 months	20.8 (0.64)	21.8 (1.40)	1.1	21.6 (1.37)	0.9	21.1 (1.36)	0.4	20.1 (1.30)	0.7
WIC benefits received past 12 months	5.5 (0.42)	5.9 (0.95)	0.5	6.0 (0.95)	0.5	5.7 (0.92)	0.3	5.3 (0.85)	0.2
Could not afford balanced meals, past 30 (p30) days	11.8 (0.49)	12.3 (1.14)	0.5	12.3 (1.14)	0.5	11.7 (1.08)	0.1	11.5 (1.05)	0.3
Food did not last, p30 days	13.4 (0.51)	13.6 (1.18)	0.2	13.7 (1.18)	0.4	13.0 (1.12)	0.4	12.9 (1.08)	0.5
Worried food would run out, p30 days	16.0 (0.56)	16.4 (1.24)	0.4	16.5 (1.25)	0.5	15.7 (1.18)	0.3	15.3 (1.14)	0.7
Food security: secure	89.3 (0.46)	88.5 (1.08)	0.8	88.4 (1.08)	0.9	89.0 (1.02)	0.3	89.3 (1.00)	0.1
Insecure	6.5 (0.37)	7.4 (0.91)	0.9	7.5 (0.92)	1.0	7.1 (0.85)	0.6	6.8 (0.81)	0.3
Very insecure	4.2 (0.30)	4.1 (0.64)	0.1	4.1 (0.63)	0.1	3.9 (0.61)	0.3	4.0 (0.61)	0.2

Indicator	Gold standard estimates using weight A (1)	Teen estimates using weight D (2)	Bias using weight D	Teen estimates using weight E (3)	Bias using weight E	Teen estimates using weight F (4)	Bias using weight F	Teen estimates using weight G (5)	Bias using weight G
Number of adults in hh work fulltime: 0	11.7 (0.47)	11.9 (1.01)	0.2	11.7 (0.99)	0.0	11.4 (0.96)	0.3	11.1 (0.92)	0.5
1	44.0 (0.72)	45.3 (1.39)	1.3	45.1 (1.38)	1.1	45.2 (1.37)	1.2	44.8 (1.35)	0.8
2 or more	44.3 (0.70)	42.8 (1.39)	1.5	43.2 (1.37)	1.1	43.4 (1.38)	0.9	44.1 (1.36)	0.3
Number of adults in hh working: 0	6.1 (0.33)	5.6 (0.74)	0.5	5.4 (0.71)	0.7	5.4 (0.69)	0.8	5.4 (0.68)	0.8
1	33.8 (0.67)	33.8 (1.35)	0.0	33.6 (1.33)	0.2	33.2 (1.34)	0.6	32.6 (1.31)	1.2
2 or more	60.1 (0.71)	60.6 (1.46)	0.5	61.0 (1.42)	0.9	61.4 (1.44)	1.3	62.0 (1.41)	2.0
Duration in residence: 1-3 years	31.9 (0.69)	30.7 (1.41)	1.2	30.5 (1.40)	1.5	30.0 (1.38)	2.0	29.5 (1.35)	2.4
4-10 years	35.0 (0.67)	36.8 (1.40)	1.8	36.9 (1.38)	1.9	36.9 (1.38)	1.9	37.3 (1.35)	2.3
11 or more years	33.1 (0.65)	32.5 (1.27)	0.6	32.7 (1.25)	0.4	33.2 (1.28)	0.1	33.2 (1.26)	0.1
1 or more working landline phones	22.3 (0.59)	23.4 (1.21)	1.1	23.3 (1.19)	1.0	23.3 (1.18)	1.0	23.4 (1.19)	1.2
Health Status indicators									
Excellent or very good health	83.4 (0.50)	83.1 (1.10)	0.3	83.1 (1.09)	0.3	83.4 (1.06)	0.1	83.9 (1.02)	0.6
Anxiety frequency: never	41.3 (0.70)	41.7 (1.51)	0.4	41.3 (1.49)	0.0	41.5 (1.49)	0.2	40.4 (1.44)	0.8
A few times per year or monthly	38.5 (0.68)	37.7 (1.36)	0.8	38.1 (1.37)	0.4	37.6 (1.34)	0.9	38.6 (1.32)	0.1
Weekly or daily	20.3 (0.57)	20.7 (1.15)	0.4	20.6 (1.14)	0.3	21.0 (1.15)	0.8	21.0 (1.12)	0.7
Depression frequency: never	59.6 (0.74)	58.0 (1.42)	1.6	57.8 (1.41)	1.7	57.7 (1.40)	1.9	56.8 (1.39)	2.7
A few times per year or monthly	24.1 (0.63)	25.1 (1.21)	1.0	25.4 (1.21)	1.3	25.1 (1.20)	1.0	25.6 (1.20)	1.5
Weekly or daily	16.4 (0.53)	17.0 (1.06)	0.6	16.8 (1.04)	0.4	17.2 (1.06)	0.9	17.5 (1.05)	1.2
Ever had asthma	14.8 (0.50)	13.1 (0.96)	1.7	13.2 (0.96)	1.7	12.8 (0.93)	2.0	12.7 (0.91)	2.1
Current asthma	8.4 (0.38)	6.4 (0.66)	1.9	6.5 (0.67)	1.9	6.2 (0.63)	2.1	6.3 (0.62)	2.1
Any difficulty seeing	6.7 (0.38)	7.2 (0.75)	0.5	7.1 (0.75)	0.4	7.2 (0.74)	0.5	7.1 (0.72)	0.4
Wears glasses or contact lenses	38.2 (0.66)	41.0 (1.41)	2.9	41.0 (1.40)	2.9	41.1 (1.39)	3.0	41.3 (1.36)	3.2
Any learning difficulty	10.9 (0.45)	9.3 (0.86)	1.7	8.9 (0.83)	2.0	9.0 (0.83)	1.9	8.7 (0.79)	2.2
Any difficulty remembering	9.5 (0.43)	8.9 (0.82)	0.6	8.6 (0.80)	0.9	8.8 (0.81)	0.6	8.6 (0.77)	0.9
Disability status: has any disabilities	14.5 (0.50)	13.4 (1.01)	1.1	13.3 (1.00)	1.2	13.6 (1.00)	0.9	13.4 (0.98)	1.1
Ever diagnosed with ADD/ADHD	15.5 (0.50)	13.0 (0.93)	2.5	12.8 (0.91)	2.7	13.0 (0.91)	2.5	13.3 (0.91)	2.2
Currently (Curr) has ADD/ADHD	13.8 (0.47)	11.8 (0.87)	2.1	11.5 (0.85)	2.3	11.8 (0.86)	2.1	12.0 (0.86)	1.8
Ever had a learning disability	10.3 (0.42)	8.5 (0.83)	1.8	8.1 (0.80)	2.1	8.2 (0.79)	2.1	7.9 (0.75)	2.4

Curr has learning disability	8.6 (0.39)	7.1 (0.78)	1.5	6.7 (0.74)	2.0	6.9 (0.75)	1.8	6.6 (0.70)	2.0
Curr has autism spectrum disorder	4.0 (0.27)	3.1 (0.51)	0.9	2.9 (0.49)	1.0	3.0 (0.48)	1.0	2.9 (0.46)	1.1
Curr any other developmental disability	3.3 (0.26)	2.9 (0.56)	0.4	2.8 (0.56)	0.4	2.8 (0.53)	0.5	2.6 (0.50)	0.7
Curr any developmental disability	19.7 (0.56)	17.9 (1.13)	1.8	17.4 (1.09)	2.2	17.6 (1.10)	2.1	17.6 (1.08)	2.1
Injury/illness school days missed: 0	39.3 (0.69)	40.1 (1.48)	0.8	40.0 (1.47)	0.7	39.7 (1.47)	0.4	39.6 (1.44)	0.3
1-5	40.3 (0.67)	41.7 (1.44)	1.4	41.6 (1.43)	1.3	42.0 (1.43)	1.7	42.2 (1.41)	1.9
6-10	11.8 (0.46)	10.6 (0.86)	1.2	10.8 (0.87)	1.1	10.7 (0.85)	1.1	10.7 (0.83)	1.1
11 or more	8.6 (0.41)	7.6 (0.79)	0.9	7.7 (0.78)	0.9	7.7 (0.78)	0.9	7.6 (0.75)	1.0
Ever receive special education services	18.7 (0.53)	17.3 (1.12)	1.4	16.8 (1.09)	1.9	17.1 (1.10)	1.6	16.6 (1.06)	2.1
Curr receive special education services	12.3 (0.45)	11.2 (0.93)	1.1	10.4 (0.88)	1.9	11.0 (0.89)	1.3	10.4 (0.84)	1.8
SLE/Social Support/Behavioral Problems									
Lived with anyone with alcohol/drug problem	11.8 (0.48)	12.1 (0.99)	0.3	11.9 (0.98)	0.0	12.2 (0.98)	0.4	12.2 (0.95)	0.4
Lived with parent who was incarcerated	8.0 (0.40)	7.9 (0.88)	0.1	7.7 (0.85)	0.4	8.0 (0.87)	0.0	8.0 (0.86)	0.1
Lived with mentally ill/depressed person	11.3 (0.49)	12.5 (0.98)	1.2	12.5 (0.97)	1.2	12.6 (0.96)	1.3	12.6 (0.95)	1.3
Lifetime of being put down by adults in home	6.8 (0.37)	5.7 (0.67)	1.1	5.7 (0.66)	1.1	5.8 (0.67)	1.0	5.9 (0.66)	0.9
Treated/judged based on race/ethnicity	8.2 (0.39)	8.3 (0.79)	0.1	8.4 (0.78)	0.2	8.0 (0.74)	0.3	7.8 (0.72)	0.5
Victim of/witnessed violence	8.3 (0.42)	8.9 (0.88)	0.5	8.8 (0.89)	0.5	8.7 (0.85)	0.4	8.7 (0.85)	0.3
Experienced two or more SLEs	13.7 (0.52)	13.4 (1.04)	0.3	13.3 (1.04)	0.4	13.4 (1.02)	0.3	13.3 (1.00)	0.4
Community support	88.3 (0.49)	89.6 (0.89)	1.3	89.6 (0.89)	1.4	89.8 (0.86)	1.6	90.4 (0.82)	2.1
Basic needs not met	4.4 (0.31)	4.3 (0.63)	0.0	4.2 (0.61)	0.1	4.4 (0.63)	0.0	4.4 (0.63)	0.1
Any difficulty with changes in routine	22.2 (0.58)	22.1 (1.14)	0.2	21.9 (1.14)	0.3	22.2 (1.14)	0.0	22.4 (1.13)	0.1
Any difficulty controlling behavior	13.3 (0.46)	12.8 (0.99)	0.5	12.7 (0.99)	0.6	12.7 (0.97)	0.6	12.6 (0.96)	0.7
Any difficulty concentrating	7.4 (0.37)	7.8 (0.78)	0.4	7.7 (0.78)	0.3	7.8 (0.77)	0.4	7.6 (0.73)	0.2
Any difficulty making friends	13.2 (0.46)	14.8 (1.00)	1.6	14.7 (0.99)	1.5	15.0 (0.99)	1.7	14.6 (0.94)	1.3
Health Care Service Use indicators									
Saw doctor, past 12 months	92.3 (0.39)	91.6 (0.84)	0.7	91.8 (0.82)	0.5	91.6 (0.84)	0.7	91.6 (0.83)	0.7
Usual place: doctor office/health center	91.1 (0.45)	90.6 (0.99)	0.5	90.8 (0.97)	0.2	90.8 (0.97)	0.3	91.2 (0.91)	0.1

Telehealth visits, past 12 months	18.1 (0.52)	18.1 (1.07)	0.0	18.0 (1.06)	0.1	18.2 (1.06)	0.1	18.1 (1.03)	0.0
Urgent care visits, past 12 months: 0	71.8 (0.64)	71.4 (1.36)	0.5	71.4 (1.34)	0.4	71.4 (1.34)	0.4	71.5 (1.31)	0.4
1 visit	16.0 (0.51)	15.9 (1.06)	0.1	15.8 (1.04)	0.1	15.9 (1.04)	0.1	16.1 (1.03)	0.1
2 or more visits	12.2 (0.47)	12.7 (0.98)	0.5	12.7 (0.97)	0.5	12.6 (0.96)	0.4	12.5 (0.92)	0.2
Wellness visits, past 12 months	89.5 (0.43)	89.8 (0.81)	0.2	89.7 (0.81)	0.2	89.7 (0.81)	0.1	89.7 (0.79)	0.1
ER visits, past 12 months: 0	86.8 (0.46)	86.7 (1.00)	0.6	86.9 (0.98)	0.6	86.9 (0.98)	0.8	87.2 (0.94)	0.9
1 visit	9.1 (0.38)	8.5 (0.82)	0.7	8.5 (0.81)	0.5	8.3 (0.79)	0.7	8.2 (0.77)	0.4
2 or more visits	4.1 (0.27)	4.8 (0.63)	0.7	4.6 (0.60)	0.5	4.8 (0.61)	0.7	4.5 (0.58)	0.4
Received flu shot, past 12 months (mos)	43.0 (0.73)	43.1 (1.47)	0.1	43.8 (1.46)	0.8	42.9 (1.46)	0.1	43.2 (1.44)	0.2
Medications for emotions, past 12 mos	13.3 (0.47)	13.1 (0.95)	0.3	12.7 (0.93)	0.6	13.2 (0.94)	0.1	13.6 (0.96)	0.2
Received mental health care, past 12 months	17.3 (0.52)	17.9 (1.03)	0.6	17.7 (1.02)	0.4	18.0 (1.02)	0.7	18.2 (1.02)	0.9
Ever diagnosed with COVID	34.2 (0.68)	33.4 (1.40)	0.8	33.2 (1.39)	1.0	33.5 (1.38)	0.7	33.6 (1.37)	0.5
Received COVID vaccination	59.2 (0.77)	59.7 (1.52)	0.5	60.8 (1.51)	1.7	59.7 (1.51)	0.5	59.7 (1.49)	0.5
Prescribed medication, past 12 months	38.4 (0.69)	38.6 (1.37)	0.2	38.5 (1.37)	0.0	38.8 (1.36)	0.3	38.9 (1.35)	0.4
Health Care Access/Insurance Coverage indicators									
Unable to pay medical bills, past 12 months	7.9 (0.40)	8.2 (0.89)	0.3	8.3 (0.89)	0.3	8.0 (0.86)	0.1	7.9 (0.84)	0.0
Worried about paying medical bills	32.4 (0.66)	32.8 (1.32)	0.4	32.9 (1.32)	0.5	32.7 (1.31)	0.2	32.2 (1.28)	0.2
Delay/defer medical care due to cost	2.7 (0.22)	2.5 (0.45)	0.1	2.5 (0.45)	0.1	2.6 (0.46)	0.1	2.5 (0.44)	0.2
Health insurance coverage: private	57.5 (0.74)	55.3 (1.49)	2.1	55.4 (1.47)	2.0	56.3 (1.47)	1.2	58.8 (1.44)	1.3
Public	38.2 (0.72)	40.6 (1.46)	2.3	40.6 (1.45)	2.3	39.6 (1.43)	1.4	37.4 (1.40)	0.8
None	4.3 (0.28)	4.1 (0.61)	0.2	4.0 (0.58)	0.3	4.1 (0.60)	0.2	3.8 (0.55)	0.5

NOTES: RPM is the raking variable derived from the recursive partitioning model. SE is standard error. DIFF is difference between estimates not shown and (1). SCR is Sample Child respondent. GED is General Equivalence Degree. FPL is Federal Poverty Level. WIC is Women, Infants & Children; ADD/ADHD is attention-deficit disorder/attention-deficit/hyperactivity disorder; SLE is Stressful Life Events. ER is emergency room.
SOURCE: National Center for Health Statistics, National Health Interview Survey, 2021-2023; National Health Interview Survey-Teen 30-Month File.

Suggested citation

Bramlett MD, Black LI, Dahlhamer JM, Zablotsky B. The impact on bias reduction from adding a weighting calibration variable derived from a recursive partitioning model. National Center for Health Statistics White Paper. 2026 July: 1-27.

Copyright information

National Center for Health Statistics

Keywords: nonresponse bias • classification tree • iterative proportional fitting • National Health Interview Survey – Teen