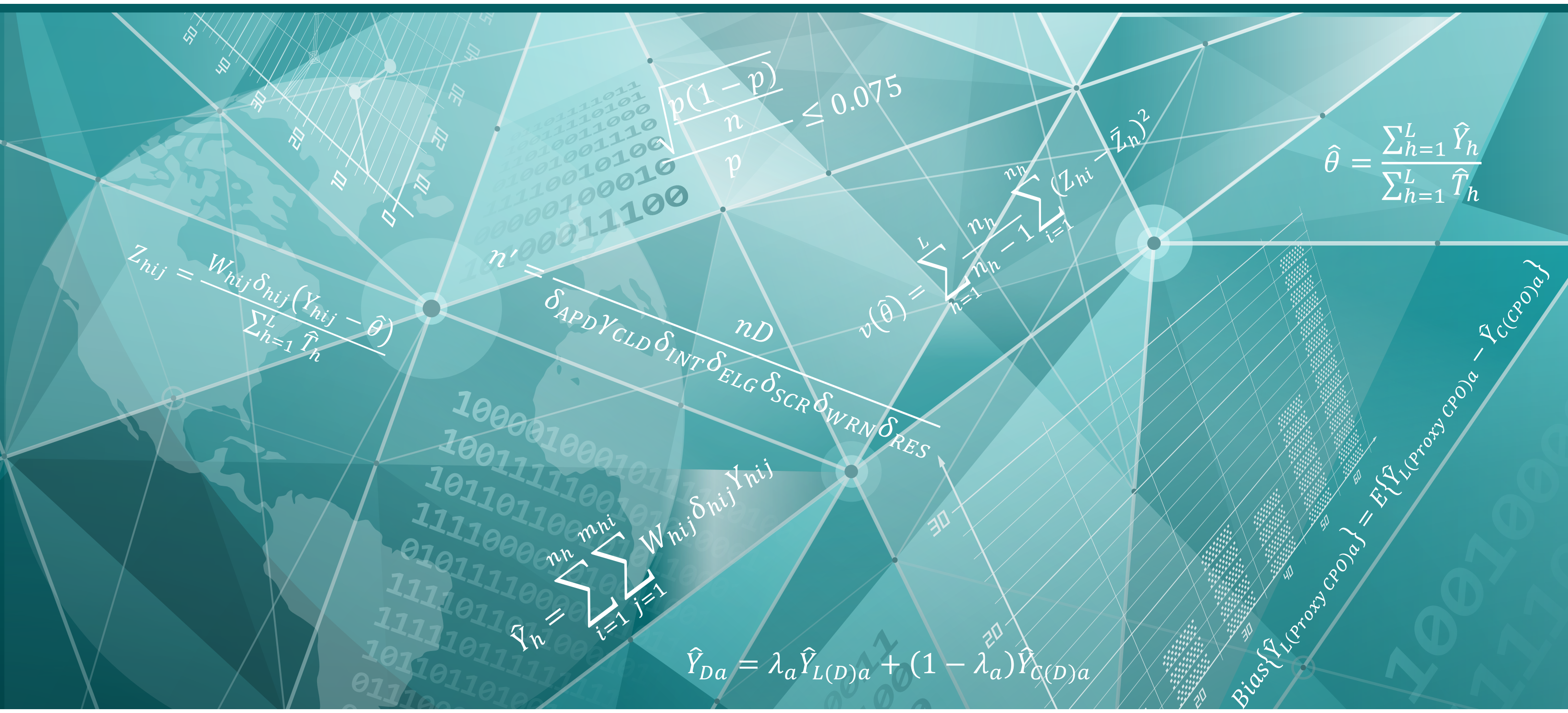
NATIONAL CENTER FOR HEALTH STATISTICS

Vital and Health Statistics

Series 2, Number 185 December 2020



An Investigation of Nonresponse Bias and Survey Location Variability in the 2017–2018 National Health and Nutrition Examination Survey

Data Evaluation and Methods Research

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

NCHS reports can be downloaded from: https://www.cdc.gov/nchs/products/index.htm.

Suggested citation

Fakhouri THI, Martin CB, Chen TC, Akinbami LJ, Ogden CL, Paulose-Ram R, et al. An investigation of nonresponse bias and survey location variability in the 2017−2018 National Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Stat 2(185). 2020.

For sale by the U.S. Government Publishing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20401–0001
Printed on acid-free paper.

NATIONAL CENTER FOR HEALTH STATISTICS

Vital and Health Statistics

Series 2, Number 185 December 2020

An Investigation of Nonresponse Bias and Survey Location Variability in the 2017–2018 National Health and Nutrition Examination Survey

Data Evaluation and Methods Research

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
December 2020

National Center for Health Statistics

Brian C. Moyer, Ph.D., *Director*

Amy M. Branum, Ph.D., *Acting Associate Director for Science*

Division of Health and Nutrition Examination Surveys

Ryne Paulose-Ram, M.A., Ph.D., *Acting Director*

Lara J. Akinbami, M.D., *Acting Associate Director for Science*

Contents

Text Figures

Contents—Con.

Text Tables

Appendix Table

An Investigation of Nonresponse Bias and Survey Location Variability in the 2017–2018 National Health and Nutrition Examination Survey

by Tala H.I. Fakhouri, Ph.D., M.P.H, Crescent B. Martin, M.P.H, M.A., Te-Ching Chen, Ph.D., Lara J. Akinbami, M.D., Cynthia L. Ogden, Ph.D., Ryne Paulose-Ram, Ph.D, M.A., Minsun K. Riddles, Ph.D., Wendy Van de Kerckhove, M.A., Shelley Brock Roth, M.S., Jason Clark, M.S., Leyla K. Mohadjer, Ph.D., and Robert E. Fay, Ph.D.

Abstract

Background

Over the past two decades, a steady decline in response rates on national face-to-face surveys has been documented, with steeper declines observed in recent years. The impact of nonresponse on survey estimates is inconsistent and depends on the correlation between response propensity and the survey estimates.

To better understand the impact of declining response rates on the 2017−2018 National Health and Nutrition Examination Survey (NHANES), potential nonresponse bias (NRB) was investigated. NRB was assessed using three approaches: (a) studying variation within the respondent set; (b) benchmarking and comparisons to external data; and (c) comparing alternative weighting adjustments. Because NHANES only samples 30 counties in every 2-year cycle, the sample of counties in any given cycle may be an outlier on some characteristics. Such sampling variability may compound the effects of NRB. For this reason, the representativeness of the 2017−2018 NHANES counties was examined by comparing: (a) the characteristics of the 2017−2018 sampled counties with those from prior cycles; (b) each sampled county with the average of all the counties in the sampling stratum from which that county was selected; and (c) the 2017−2018 counties with 5,000 other samples that could have been drawn under the same sample design using a simulation study.

The NRB analyses showed that the 2017−2018 NHANES sample had a lower proportion of college graduates and higher-income individuals compared with prior cycles. Additionally, the 2017−2018 NHANES counties had lower proportions of college graduates and lower mean incomes compared with counties from prior cycles and counties not selected in 2017−2018, which exacerbated the effects of NRB.

Weighting adjustments used in prior cycles were not sufficient to address the bias in the 2017−2018 NHANES. Instead, enhanced weighting adjustments for education and income reduced the bias resulting from nonresponse and location sampling variability.

Keywords: response rates • survey nonresponse • weighting adjustment • total survey error

Introduction

The National Health and Nutrition Examination Survey (NHANES) is a national survey designed to assess the health and nutritional status of children and adults living in the United States (1). NHANES is unique in that it combines in-home health interviews with standardized health examinations and biospecimen collections and is the only national survey that measures dietary intake and environmental exposures. NHANES findings have been used to set goals and track progress in reducing cholesterol levels, the prevalence of high blood pressure, and the risks of blood lead exposure in the United States, among other health outcomes, and NHANES data were used to develop the Centers for Disease Control and Prevention, or CDC, pediatric growth charts used in pediatricians' offices across the country (2). NHANES documented the rise in obesity and diabetes and produced the first population-based estimates of HIV infection and osteoporosis in the United States.

Like many other national, probability-based face-to-face surveys, NHANES has experienced a decline in response rates in recent years (3). The response rate to a survey is a valuable data quality measure and the most widely used indicator of survey quality. Nonresponse increases the risk of bias in survey estimates (4−6). However, there is no known threshold for a response rate at which bias is introduced.

Instead, nonresponse bias is a function of the correlation between the propensity to respond and survey outcomes (4−6). Importantly, within the same survey, different sample estimates can be subject to different nonresponse biases (4−6).

The 2017−2018 cycle of NHANES experienced another drop in response rates (3). This decline necessitated a thorough investigation into potential biases in survey estimates that may have resulted from unit nonresponse (i.e., when a sampled person [SP] does not participate in the survey). Additionally, because NHANES only samples 30 counties per 2-year cycle, the impact of survey location sampling variability on survey estimates was assessed (i.e., differences in the characteristics of the survey locations in the 2017−2018 cycle of NHANES compared with other cycles and with national estimates of high quality for benchmarking). The relatively small number of sampled counties in any given cycle of NHANES makes the survey susceptible to an outlier sample of counties (i.e., sampling variation). This sampling variation may exacerbate or lessen the effects of nonresponse, depending on the variation in the sociodemographic and health characteristics between the counties, the differential response rates across counties, and the correlation between survey participants' response propensities and the survey estimates. This report summarizes the findings from the investigation of nonresponse bias and survey location sampling variability in the 2017−2018 cycle of NHANES and presents the rationale behind the introduction of enhanced weighting adjustments that included additional calibrations to education and income to mitigate the impact of nonresponse bias and location sampling variability.

Methods

Survey Design

NHANES uses a complex, multistage, stratified probability sampling design. The primary sampling units (PSUs) for NHANES are single counties, or groups of tracts within counties, or small groups of contiguous counties that are sampled with probability proportional to size within strata. The 2015−2018 NHANES sample design had 14 major strata, each with 4 minor strata (7). The sample design applies to two consecutive 2-year cycles. For each 4-year sample design, 4 PSUs were chosen from each major stratum, one from each of its 4 minor strata, for a total of 56 PSUs plus an additional 4 PSUs chosen with certainty, for a grand total of 60 PSUs. Major strata had a range of 12 to 897 PSUs. The details of the sample design and the stratification scheme are described elsewhere (7).

The second sampling stage consists of selecting segments defined by census blocks or groups of blocks within each PSU. The third stage of sample selection consists of selecting dwelling units (DUs), including households and noninstitutional group quarters, within each segment. The fourth and final stage of sample selection consists of screening DUs to select persons for interview and examination. Persons are selected at different rates according to certain characteristics (age, sex, race and Hispanic origin, and income) to ensure a representative sample and to provide sufficient sample sizes for smaller groups to obtain reliable survey estimates.

Response Rates

Sample sizes and response rates for all 2-year survey cycles, overall and by age and sex, are provided on the NHANES website (8). Response rates are calculated for each stage of the survey: (a) screening of DUs to identify SPs; (b) household interviewing of SPs; and (c) examination of interviewed SPs. In the 2017–2018 NHANES cycle, the household screener response rate was 90.9%. From the responding households, 16,211 SPs were selected from 30 different PSUs. Of those selected, 9,254 completed the interview and 8,704 were examined. After adjustment for nonresponse to the screener, the final interview response rate was 51.9%, and the final examination response rate was 48.8%.

Nonresponse Bias Assessment

Nonresponse bias, which affects survey estimates, may be introduced when some sample members do not participate in a survey. This bias can be substantial when two conditions hold: (a) the response rate is relatively low, and (b) the difference between the characteristics of respondents and nonrespondents is relatively large. Unit nonresponse bias in NHANES 2017–2018 was investigated using three methodological approaches based on the Groves and Brick typology and as described in a recent report on nonresponse bias by the Federal Committee on Statistical Methodology (9,10): (a) studying variation within the respondent set; (b) comparing with external data sources and benchmarking (i.e., comparisons with other and authoritative survey estimates); and (c) comparing alternative postsurvey weighting adjustments. This process was iterative and the results from each set of analyses were used to inform subsequent weighting adjustments.

Variation Within the Respondent Set

For analyses evaluating the variation within the respondent set, three methods were used: (a) comparison of response rates across subgroups; (b) R indicator analyses; and (c) level of effort analyses. Base weights that account for the differential probabilities of selection by age, sex, race and Hispanic origin, and income were used for all three analyses (7).

Response rate comparisons across subgroups

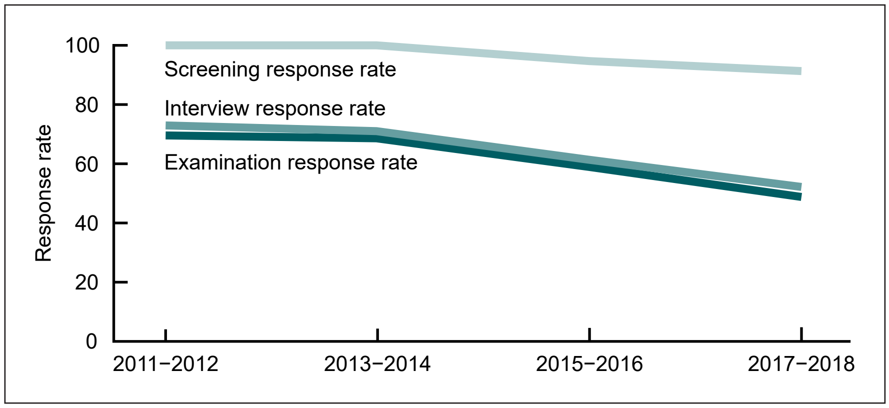
For this analysis, the differences in response rates by key subgroups of the target population were investigated (e.g., by age, race, sex, and urbanicity). A difference in response rates by subgroup indicates that additional postsurvey weighting adjustments using some of these characteristics could be warranted to mitigate bias.

Differences in the household interview response rates by 38 subgroup variables were examined (see the Appendix Table for the complete list of all 38 variables). The focus of this analysis was on differences in the household interview response rates because most of the nonresponse occurs between the screening and the household interview stages (Figure 1). The 38 variables include those describing the SP (e.g., age, sex, and race and Hispanic origin), the household where the SP lives (e.g., household size and presence of children in the household), and the census tract where the SP lives (e.g., median census tract income and median census tract education levels). The distribution of interview respondents was compared with the distribution of SPs that did not respond to the interview, and a second-order (Satterthwaite) Rao-Scott chi-square test of independence was performed to test the significance of the relationship between response status and each of the subgroup variables (11).

Representativeness indicator (R indicator) analysis

The representativeness indicator (R indicator) is a measure to assess data quality before any nonresponse adjustment weighting (12). It is a function of the variation in response propensities and takes a value from 0 to 1. A high value means low variation in propensity scores, signifying that the sample respondents are highly representative of the target population.

Figure 1. Screening, interview, and examination response rates: National Health and Nutrition Examination Survey, 2011–2012 through 2017–2018



SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2011–2018.

The R indicator is calculated as follows:

$$\hat{R}(\hat{\rho}) = 1 - 2\sqrt{\frac{1}{\hat{N}-1}\sum_{i \in S} w_i \left(\hat{\rho}_i - \hat{\bar{\rho}}\right)^2}$$

where $\hat{\rho}_i$ is the estimated response propensity for unit i;

$$\hat{\bar{\rho}} = \hat{N}^{-1}\sum_{i \in S} w_i \, \hat{\rho}_i$$

S denotes the sample; w_i is the inverse of the selection probability of unit i; and

$$\hat{N} = \sum_{i \in S} w_i$$

Because the true response propensities are not known, they must be estimated. This was done by fitting a propensity model with available auxiliary data. To estimate response propensities, population-level logistic regression models were fit for the NHANES 2011–2012, 2013–2014, 2015–2016, and 2017–2018 cycles. The dependent variable was the response indicator (1: responded to both the interview and the examination, 0: otherwise) conditional on responding to the screener; as a result, this analysis does not address nonresponse to the screener. Due to the high conditional examination response rate, the combined interview and examination response was treated as a single response status variable for this analysis. In other words, the response to the interview alone was not analyzed separately.

For a consistent comparison among the four cycles, the following set of common auxiliary variables from the screener were included as predictors: SP's sex, SP's categorized age, household composition of SPs by age (i.e., a variable that combines number of SPs in a household with the age groups), household size, presence of children in the household, household reference person's sex, number of SPs in the household, and race and Hispanic origin (i.e., black, Hispanic, Asian, and white by income: white or other, low income, and white or other, not low income). Additionally, the following set of common auxiliary variables from the census were also used as predictors: census region, urbanicity, PSU stratum, census tract level median income, census tract level percentage of the workforce population that have a 30-minute or longer commute to work, and census tract level percentage of housing units with a monthly housing cost greater than $1,500.

To select predictors, a stepwise logistic regression model and a classification tree using the SAS procedure HPSPLIT were fit with response status as the dependent variable and a larger number of auxiliary variables as the potential predictors. The response propensity model included predictors that were selected in any of these two methods consistently over cycles. After the response propensities were estimated, the R indicators, appropriate standard errors (SEs), and 95% confidence bounds were calculated for each of the four survey cycles.

Level of effort analysis

The level of effort analysis assesses bias in outcome estimates by evaluating whether estimates change as effort to encourage an SP to participate increases, and more critically, whether the final estimate would have changed if additional contacts (beyond those already done) were made with the selected sample. To the extent that the hard-to-reach respondents (requiring higher level of effort) are similar to the nonrespondents, differences in outcome estimates between the hard-to-reach and easy-to-reach respondents could indicate nonresponse bias (13−15). Therefore, examining the differences between outcome estimates separately for these different types of respondents is a common approach for evaluating potential for bias when response rates are low.

For this analysis, the measure of effort is defined by the number of contact attempts made with an SP before the SP completes the interview. Contact attempts include all visits to the household, when the SP may or may not be present, including visits needed to complete the screener questionnaire. The number of total contact attempts for SPs who completed an interview in 2017−2018 ranged from 2 to 31. Because the number of respondents with two contact attempts was small, these cases are collapsed with respondents with three contact attempts. Respondents who required 4 to 15 contact attempts were divided into 12 separate groups by the number of contact attempts. Again, and to ensure that enough respondents were in each group, respondents with 16 or 17 contact attempts were grouped together, respondents with 18 to 20 contact attempts were grouped together, and respondents with 21 or more contact attempts were grouped together.

The following four key health outcomes were selected for the analysis: (a) obesity prevalence for adults aged 20 and over; (b) hypertension prevalence for adults aged 20 and over; (c) high total cholesterol prevalence for adults aged 20 years and over; and (d) diagnosed diabetes prevalence for adults aged 20 and over. Pregnant individuals were excluded from all analyses. Obesity in adults was defined as a body mass index (BMI) of greater than or equal to 30. BMI was calculated as weight in kilograms divided by height in meters squared, rounded to one decimal place. Hypertension was defined as systolic blood pressure greater than or equal to 130 mm Hg or diastolic blood pressure greater than or equal to 80 mm Hg, or currently taking medication to lower high blood pressure. All blood pressure readings were obtained during a single examination visit. After a 5-minute rest in a seated position, participants had up to three brachial systolic and diastolic blood pressure measurements taken 30 seconds apart. An average of up to three systolic and diastolic blood pressure readings was used for systolic and diastolic blood pressure values. High total cholesterol was defined as serum total cholesterol greater than or equal to 240 mg/dL. Adults aged 20 and over were classified as having diagnosed diabetes if they answered "yes" to the question, "Other than during pregnancy, have you ever been told by a doctor or health professional that you have diabetes or sugar diabetes?"

A set of base-weighted estimates and 95% confidence intervals were computed for each of the four key outcomes (obesity, hypertension, high total cholesterol, and diabetes): (a) the estimate based on the easiest-to-reach respondents who required two to three contact attempts for cooperation, (b) a sequence of cumulative estimates based on respondents who required up to four contact attempts, up to five contact attempts, etc., and (c) the estimate based on all respondents. To detect potential differences in health characteristics between easier-to-reach respondents and all respondents, a *t* test was conducted for each estimate using PROC SURVEYREG in SAS. The *t* test compared estimates from respondents who responded by the *i*th contact attempt with estimates based on all respondents, where i = 3, 4, 5, …, 15, 17, 20. For example, estimates from respondents who responded within the first five contact attempts were compared with estimates based on all respondents.

Comparisons With External Data Sources and Benchmarking

NHANES estimates for demographic characteristics and selected health conditions were compared with estimates from the American Community Survey (ACS) and the National Health Interview Survey (NHIS). The first set of comparisons was performed using the W1-adjusted NHANES sample weights. These weights were developed using the same weighting approach used in prior NHANES cycles, as described in detail in the following section. Differences between the W1-adjusted NHANES and ACS estimates may indicate the presence of nonresponse bias, and the results from these analyses were used to inform additional weighting adjustments used in the construction of W2-adjusted sample weights. The W2-adjusted NHANES estimates were then compared with ACS estimates again, and the findings were used to guide the development of the W3-adjusted weights, and so on. A total of four weighting adjustments are described in the following section of this report (i.e., W1, W2, W3, and W4). Briefly, the W1-adjusted weights were constructed using the same weighting approach as in the 2011−2012, 2013−2014, and 2015−2016 NHANES cycles. The W2-adjusted weights included additional adjustments for educational attainment, and the W3- and W4-adjusted weights included additional adjustments for income using two different approaches, as described in the next section. While the comparisons with ACS were used to inform these additional weighting adjustments, the comparisons with NHIS were used as a diagnostic tool to assess the impact of these additional weighting adjustments on key health outcomes and on trends over time, as described in later sections.

Comparisons were made at different time periods of data collection that correspond to NHANES cycles (i.e., 2011−2012, 2013−2014, 2015−2016, and 2017−2018). Some differences in the estimates may be the result of differences in survey design and administration. However, large differences may indicate the presence of nonresponse bias, especially if these differences are not consistent across time periods.

Comparisons with estimates from ACS

The weighted distribution of the NHANES sample was compared with ACS on education for U.S. adults aged 20 and over, and on income for all U.S. noninstitutionalized residents. Data for this analysis were from the 2011−2012, 2013−2014, 2015−2016, and 2017−2018 cycles of NHANES, and from the ACS 1-year Public Use Microdata Sample from 2011 through 2018, combined into 2-year periods. ACS is an ongoing large-scale survey that provides detailed population and housing information about the United States. Data collected include demographic, income, education, employment, insurance, and other general information about the respondents' communities. One difference between ACS and NHANES is that ACS is a multimode survey, while NHANES interviews are strictly conducted in person. In addition, the ACS target population is the resident population of the United States, whereas NHANES is limited to the civilian noninstitutionalized population. To account for this difference, persons in the military or living in institutional group quarters were excluded from the ACS estimates for the comparisons in this report.

For this analysis, income levels were defined using the family income-to-poverty level ratio (FIPR), which was calculated by dividing family income by a poverty threshold specific for family size and year. The U.S. Department of Health and Human Services' poverty guidelines were used as the poverty measure to calculate the FIPR (16). The cutoff for the lower income level was based on income eligibility for participation in the Special Supplemental Nutrition Program for Women, Infants, and Children, or WIC, which is less than or equal to 185% of the FIPR (16).

Because both NHANES and ACS are subject to error, variances of all estimates were taken into account when making comparisons and performing significance testing. SEs of the ACS estimates were calculated using replicate weights and the jackknife method (17). Taylor series linearization was used to compute variance estimates for NHANES. Two-sided *t* tests were used to test for significant differences. The SE in the *t* test was calculated using the formula for the SE of a difference for two independent samples.

Benchmarking to NHIS

For the benchmarking analysis, NHANES health insurance coverage, disability, and diagnosed diabetes estimates were compared with those from NHIS for the 2017–2018 cycle and for the three prior cycles starting in 2011. NHIS was chosen for this comparison because it produces comparable measures of these three estimates with higher precision. Both surveys are nationally representative household surveys of the civilian noninstitutionalized U.S. population. In addition, NHANES and NHIS questions used to assess health insurance coverage, disability, and diagnosed diabetes are identical. NHIS estimates for these measures are also used to inform Healthy People goals (18). NHIS also has a larger sample size and more PSUs than NHANES, so the NHIS estimates have lower sampling error. The NHIS sample design and methodology is described in detail on the NHIS website (19).

Adults aged 20 and over were classified as having diagnosed diabetes if they answered "yes" to the question, "Other than during pregnancy, have you ever been told by a doctor or health professional that you have diabetes or sugar diabetes?" This was the same question used for the level of effort analyses. Participants were classified as being uninsured if they answered "no" to the question, "Are you covered by health insurance or some other kind of health care plan?" Insurance status is analyzed among persons aged 0−64 years because almost all persons aged 65 and over are eligible for coverage under Medicare. Disability prevalence was based on a set of six disability questions first developed and used on ACS and adopted by several federal data collection systems, including NHANES starting in 2013 and NHIS starting in 2011. Adults aged 20 and over were classified as having a disability if they answered "yes" to having difficulty in one or more of six areas: serious difficulties seeing, hearing, concentrating, or walking; difficulty dressing or bathing; and difficulty doing errands alone.

For the comparisons with NHIS, the W1- and W4-adjusted sample weights were used to compute the NHANES estimates. Taylor series linearization was used to compute variance estimates that account for the complex sample design of each survey. For consistency with other published reports, estimates were age adjusted to the projected 2000 U.S. Census population using age groups 20–39, 40–59, and 60 and over (for diagnosed diabetes and disability) or using age groups 0–19, 20–39, and 40–64 years (for percentage uninsured) (20). Confidence limits were calculated using the Korn and Graubard method (21). Annual data from NHIS were pooled into 2-year periods corresponding with NHANES cycles.

Comparing Alternative Postsurvey Adjustments

The findings from studying the variation within the respondent set and from the comparisons with external data sources and benchmarking were used to inform four weighting adjustments for nonresponse, as was briefly described in the preceding section. These weighting adjustments expanded beyond what was traditionally used in the construction of the NHANES sample weights. The performance of these newly adjusted weights was evaluated by comparing NHANES sociodemographic and health characteristics with those from ACS and NHIS, and the variances of estimates generated by the different weights were contrasted.

The weighting process that has been used in previous NHANES cycles, with no additional adjustments (W1), was compared with three other weighting procedures with additional adjustment for: (a) education level (W2), (b) education level by raking and income using generalized regression (GREG) (W3), and (c) education level and income by raking (W4). Because education and income are correlated, the first consideration for an additional weight adjustment was to adjust for education alone and then review whether the differences in income between the final sample and the total population were reduced without further adjustment.

Initial weighting process (W1)

As was done in previous NHANES cycles, the initial weighting process was carried out in three steps, as described in detail in the NHANES 2015−2018 sample design and estimation procedures report (7). Briefly, the first step involved the computation of base weights to compensate for unequal selection probabilities. The second step adjusted for nonresponse. In the third step, the sample weights were poststratified to U.S. Census Bureau estimates of the U.S. population. These steps were performed for respondents to each stage of the survey: the screener, personal interview, and examination. Weights equal to the reciprocal of the SP's probability of selection were the starting point (or base weights) for the screener weight calculation. Those weights were then adjusted for nonresponse to the screener and poststratified. The resulting weights were the starting point for the calculation of the interview weights, which were then adjusted for nonresponse to the interview, inspected for extreme weights, and again poststratified. Finally, those poststratified interview weights were the starting point for the calculation of the examination weights. Those weights were adjusted for nonresponse to the examination, inspected for extreme weights, and poststratified (7). After the W1-adjusted weights were created, a series of nonresponse bias analyses were performed using these weights, as described in the prior section. The comparisons with ACS were used to assess whether additional adjustments were needed. At each stage (i.e., screener, interview, and exam), the nonresponse adjustment procedure consisted of creating groups (called adjustment cells) of SPs with similar response propensities, computing adjustment factors for each cell, and applying these factors to the survey weights. The creation of the nonresponse adjustment cells was informed by the findings from the analyses of variation within the respondent set. Nonresponse adjustment reduces bias if response rates and characteristics vary from cell to cell and respondents and nonrespondents sharing the same characteristics are in the same cell.

After the nonresponse adjustment, the final step of poststratification to known population totals was performed to compensate for undercoverage or overcoverage of certain demographic groups and for any residual differential nonresponse among these groups. Control totals for groups defined by age, sex, and race and Hispanic origin were obtained using population estimates from ACS. The ACS estimates have undergone poststratification to the U.S. Census Bureau's best estimates of the total noninstitutionalized civilian population of the United States, including those not counted in surveys or in the most recent decennial census.

Adjusting for education (W2)

For this weighting adjustment, base weights and nonresponse adjustments were conducted the same way they were for the initial weights (W1). The final step in the weighting procedure for each survey stage was previously called poststratification, which is a specific kind of calibration to a single set of population totals. To also adjust for education in W2, calibration was done iteratively to population totals by a combination of education level, race and Hispanic origin, and sex, as well as to population totals by a combination of age group, race and Hispanic origin, and sex in a process called raking. The details of these adjustments can be found in the NHANES 2015–2018 sample design and estimation procedures report (7).

Race, Hispanic origin, age, and sex were collected from all SPs in the screener, but education level was not, so the screener weights were still poststratified to race−Hispanic origin−age−sex demographic subgroups. Highest education level for an SP was collected in the interview, so the interview and examination weights could be calibrated (using 2-dimensional raking) to race−Hispanic origin−age−sex demographic subgroups and race−Hispanic origin−sex−education level subgroups for adults aged 20 and over.

Adjusting for education by raking and income by generalized regression estimation (W3)

As an alternative approach, a GREG adjustment was made for PSU-level income (W3). In previous weighting processes, the last step was calibration to demographic controls. The poststratification or raking compensates for demographic differences between the realized NHANES sample and known national totals. In other words, poststratification or raking generally improves the precision of NHANES estimates by offsetting demographic variability. Poststratification or raking is a generalization of ratio estimation, but unlike ratio estimation, it is implemented on categorical variables only.

As a closely related alternative to raking, GREG was investigated to attempt to compensate for some of this ecological variability at the county level. GREG is also a generalization of ratio estimation but unlike raking, it can adjust for a mixture of continuous and categorical variables. A general description of the method is described by Särndal (22). The application of the GREG method to domain estimation is described by Rao and Molina (23). GREG estimation has also been used for survey estimation (24,25).

Rather than a full-scale implementation of GREG to replace the raking, a two-step method was investigated: an initial GREG step at the PSU level, followed by raking to demographic controls at the person level. The GREG step adjusted the PSU weights to agree with national totals from the U.S. Census Bureau's Vintage 2018 postcensal estimates for the population age 20 and over on July 1, 2017 and from the 2013−2017 ACS 5-year estimates for the aggregate household income.

Adjusting for education and income by raking (W4)

As an alternative to a PSU-level adjustment for income, a raking approach using census tract level income was explored and used for the creation of sample weight W4. For this weighting adjustment, base weights and nonresponse adjustments were again conducted the same way as with the initial weights (W1). Similar to the adjustment for education level (W2), calibration was done iteratively to population totals by a combination of education level, race and Hispanic origin, and sex, as well as to population totals by a combination of age group, race and Hispanic origin, and sex using raking. To adjust the income by raking, census tract level mean household income was obtained from the 2013−2017 ACS for every tract in the United States. These tracts were then divided into income deciles, with an equal number of tracts (about 7,200) in each decile. The weights were adjusted so that the population estimate in each decile matched the U.S. population. Income is known to have relatively high item nonresponse and falsification rates (26,27), so area-level mean household income was used for calibration instead of person-level income.

Race, Hispanic origin, age, and sex are collected from all SPs in the screener, and area-level mean household income can be obtained for the area (census tract) where each SP lives, so the screener weights were calibrated (using 2-dimensional raking) to race−Hispanic origin−age−sex demographic subgroups and area-level household income. Highest education level for an SP was collected in the interview, so the interview and examination weights could be calibrated (using 3-dimensional raking) to race−Hispanic origin−age−sex demographic subgroups, race−Hispanic origin−sex−education level subgroups for adults aged 20 and over, and area-level household income.

Survey Location Sampling Variability Assessment

The first stage of the NHANES sample design consists of selecting one PSU per minor stratum, where sampling strata are created by grouping similar PSUs to improve efficiency (7). Millions of possible samples, with different combinations of PSUs, could be drawn from these strata, each producing estimates of the population values of interest. Because only 30 PSUs are selected for each cycle, NHANES is especially susceptible to survey location variability compared with surveys with more PSUs.

To investigate the impact of survey location sampling variability on survey outcomes, three analyses were conducted. First, the socioeconomic and health characteristics of the counties selected in the 2017−2018 cycle were compared with the same characteristics of counties selected in the 2011−2012, 2013−2014, and 2015−2016 cycles. This comparison used the same data source from the same time period for each set of counties so that the only difference in each cycle was the counties themselves and not changes over time. Second, the health and socioeconomic characteristics of each selected PSU in 2017−2018 were compared with the average characteristics of all PSUs in the stratum from which the selected PSU was sampled. Finally, the socioeconomic and health characteristics of the 2017−2018 NHANES sampled counties were compared with the distribution of the same characteristics of counties selected in a simulation of 5,000 other samples that could have been drawn from the same sample design.

County-level income and education level information were obtained from ACS. Specifically, estimates for mean household income and percentage of the population by education level were obtained for all 3,142 counties in the United States from the 5-year 2013−2017 ACS. County-level health data were obtained for all 3,142 counties in the United States County Health Rankings & Roadmaps Program (28), a collaboration between the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute. The 2019 County Health Rankings & Roadmaps data release was used for these analyses. Specifically, years of potential life lost before age 75 per 100,000 population (age adjusted) were calculated from the mortality data drawn from the National Center for Health Statistics National Vital Statistics System (29). Estimates of the percentage of the population that is uninsured are from the U.S. Census Bureau's Small Area Health Insurance Estimates Program (30), which produces estimates of health insurance coverage for all states and counties. Adult obesity and diabetes estimates are from the United States Diabetes Surveillance System (31), which provides county-level estimates using 3 years of data from the Behavioral Risk Factor Surveillance System and data from the U.S. Census Bureau's Population Estimates Program (32,33). The county-level estimates are based on indirect model-dependent estimates. Bayesian multilevel modeling techniques were used to obtain these estimates. Importantly, the purpose of these health-related analyses was not to compare NHANES health data with health data obtained from external sources. Instead, the goal was to use county-level health information from external data sources to compare the health of counties selected for the 2017–2018 cycle of NHANES with: (a) counties selected in prior cycles; (b) the average of the stratum to which each county belongs; and (c) counties that could have been selected under the NHANES 2015–2018 sample design to determine if the sample selected was toward the tail of the sampling distribution.

For the simulation for this third comparison, 5,000 separate random samples of PSUs were drawn under the NHANES 2015–2018 sample design. Socioeconomic and health characteristics were obtained for each sample, and the distribution of these characteristics was used for comparison with the actual 2017–2018 sample.

Results

Nonresponse Bias Analyses Results

Figure 1 shows screener, interview, and examination response rates for the last four cycles of NHANES from 2011–2012 through 2017–2018. The screener response rate decreased from near 100% in 2011–2012 to 90.9% in 2017–2018, a 7.4 percentage point decline. The response rates for both the interview and the examination stages of the survey also declined but at a greater rate. Between 2011–2012 and 2017–2018, the interview response rate declined 20.7 percentage points from 72.6% to 51.9%. The examination response rate also declined 20.7 percentage points, from 69.5% to 48.8%.

Variation Within the Respondent Set

Response rates by subgroups

Base-weighted response rates for characteristics of the SP, the household where the SP lives, and the census tract where the SP lives, were evaluated. Table A shows some variables that have a statistically significant association ($p < 0.05$) with interview response status: SP characteristics including sex, age, and race and Hispanic origin by income categories; household characteristics including size and the presence of children in the household; geographic location characteristics including PSU population size and other census-tract level variables including the proportion of adults aged 25 and over with a college degree and those who are uninsured. The complete table with all 38 variables, the accompanying subgroup response rates, and significance testing results can be found in the Appendix Table.

Table A. Significant results from the analysis of interview response rates, by potential weighting variables: National Health and Nutrition Examination Survey, 2017–2018

Variable	Base-weighted interview response rate
Sex ($p = 0.0001$)	
Male	54.3
Female	58.0
Age (years) ($p < 0.0001$)	
5 and under	67.6
6–19	63.9
20–59	53.8
60 and over	50.9
Race and Hispanic origin and income ($p < 0.0001$)	
Black, all income levels	63.9
Hispanic, all income levels	60.3
Asian, all income levels	49.0
White or other and low income	65.6
White or other and not low income	49.9
Children in household ($p < 0.0001$)	
No	50.1
Yes	61.7
Household size ($p = 0.0001$)	
1–2	51.9
3–4	56.0
5–6	60.3
7 or more	70.2
PSU population size ($p < 0.0001$)	
100,000 or Less	64.5
100,001–250,000	57.6
250,001–1,000,000	53.8
1,000,000 or more	48.0
Percentage of population aged 25 and over with a college education or higher ($p = 0.0001$)	
1st quartile	60.9
2nd quartile	58.7
3rd quartile	54.5
4th quartile	49.5
Median household income ($p < 0.0001$)	
1st quintile	64.7
2nd quintile	57.9
3rd quintile	57.3
4th quintile	53.7
5th quintile	45.8
Percentage of households that received food stamps in the past 12 months ($p < 0.0001$)	
1st quartile	48.0
2nd quartile	52.7
3rd quartile	58.3
4th quartile	62.6

See footnote at end of table.

Table A. Significant results from the analysis of interview response rates, by potential weighting variables: National Health and Nutrition Examination Survey, 2017–2018—Con.

Variable	Base-weighted interview response rate
Percentage of occupied housing units that are owned (p = 0.0248)	
1st quartile	56.9
2nd quartile	59.6
3rd quartile	56.2
4th quartile	51.9
Percentage of population who are uninsured (p = 0.0131)	
1st quartile	50.9
2nd quartile	55.9
3rd quartile	57.2
4th quartile	60.8
Percentage of population with disability (p = 0.0201)	
1st quartile	53.1
2nd quartile	53.1
3rd quartile	56.8
4th quartile	60.4

NOTES: PSU is primary sampling unit. p values are shown for the chi-square test of the relationship between response status and each of the subgroup variables.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2017–2018.

Interview response rates differed by SP characteristics such as sex, age, and race and Hispanic origin by income categories, and by household characteristics. For example, response rates generally decreased as the age of the SP increased, and SPs living in households with no children were less likely to respond than SPs living in households with children. Differences in interview response rates by geographic location characteristics were also noted. Specifically, interview response status was found to be significantly related to many of the census tract-level variables describing the characteristics of the area in which the SP resides. In particular, areas with higher educational attainment and income had lower response rates. For example, SPs living in areas with lower percentages of the population with a college degree were more likely to respond.

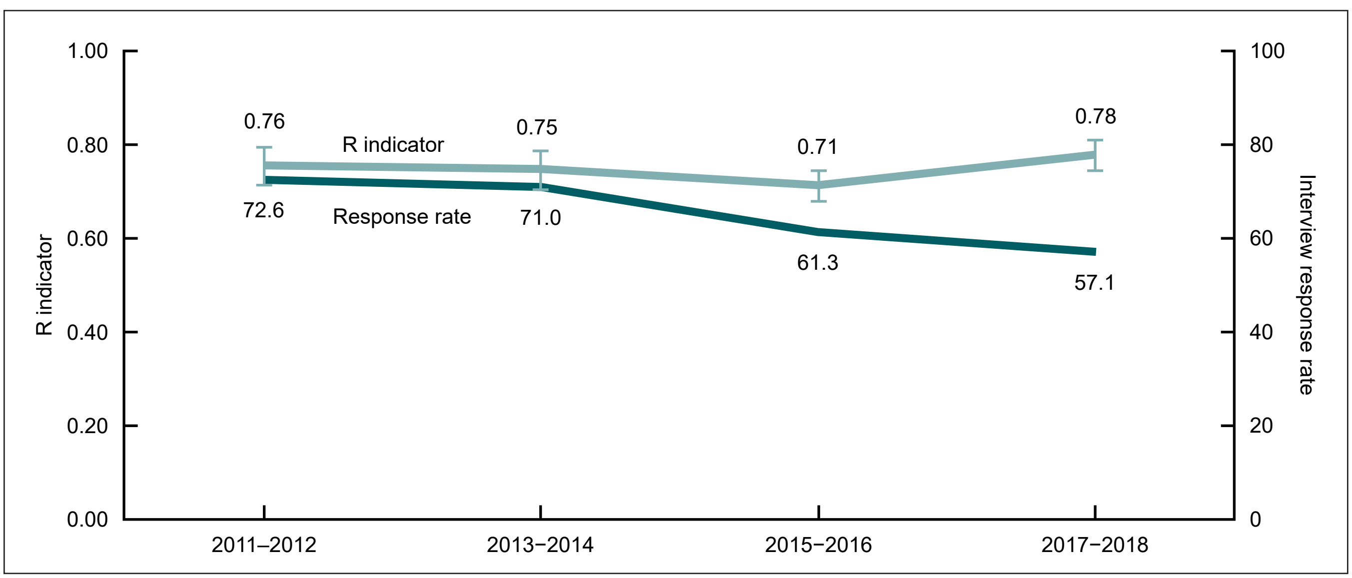
R indicator

The R indicators and 95% confidence bounds for NHANES 2011–2012, 2013–2014, 2015–2016, and 2017–2018 are shown in Figure 2, along with the base-weighted interview response rates. As described in the Methods section, an R indicator is a negative linear function of the standard deviation of estimated response probabilities and ranges from 0 to 1. A value close to 1 implies low variation in propensity scores, signifying respondents' high representativeness of the target population. For all four cycles, the R indicator was above 0.70. Although the response rate in 2017–2018 was significantly lower than in previous cycles, the R indicators were similar across cycles, providing an indication that the representativeness of the sample was not affected by the decline in response rate with respect to the auxiliary variables.

Level of effort

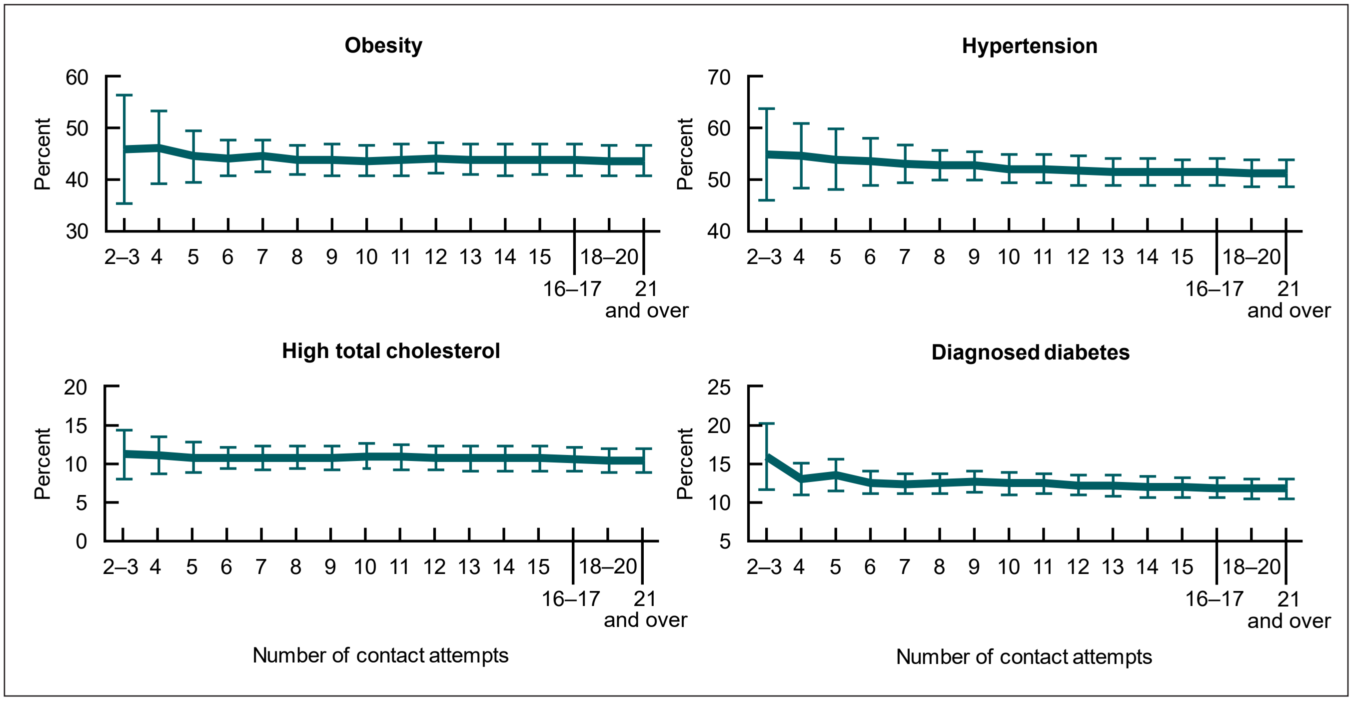
Changes in estimates for key health outcomes by the total, or cumulative, number of contact attempts are shown in Figure 3. These findings indicate that some differences were seen between the easier-to-reach respondents and the

Figure 2. R indicator and interview response rates: National Health and Nutrition Examination Survey, 2011–2012 through 2017–2018



NOTE: R indicator is representativeness indicator.
SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2011–2018.

Figure 3. Change in estimates for key health outcomes among adults aged 20 and over, by total number of contact attempts: National Health and Nutrition Examination Survey, 2017–2018



SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2017–2018.

harder-to-reach respondents. For example, the easier-to-reach adult respondents were more likely to have obesity, hypertension, or diabetes. However, the estimates became stable after approximately 8–10 contact attempts, which suggests that the characteristics of the harder-to-reach respondents had a minimal effect on the estimates of interest. To the extent that the harder-to-reach respondents were similar to the nonrespondents, these results indicate that nonresponse bias in the estimates may be minimal.

The analysis of the variation within the respondent set showed that response rates were significantly different when comparing several subgroups. The findings from this analysis were used to inform nonresponse adjustments as described in later sections. The R indicator analysis showed that representativeness of the sample overall was still at a similar level to previous cycles despite the declining response rates. The variation by level of effort showed that several key estimates stabilized after several contact attempts, indicating that nonrespondents (to the extent they are similar to the harder-to-reach respondents) would not have significantly changed the final weighted estimates.

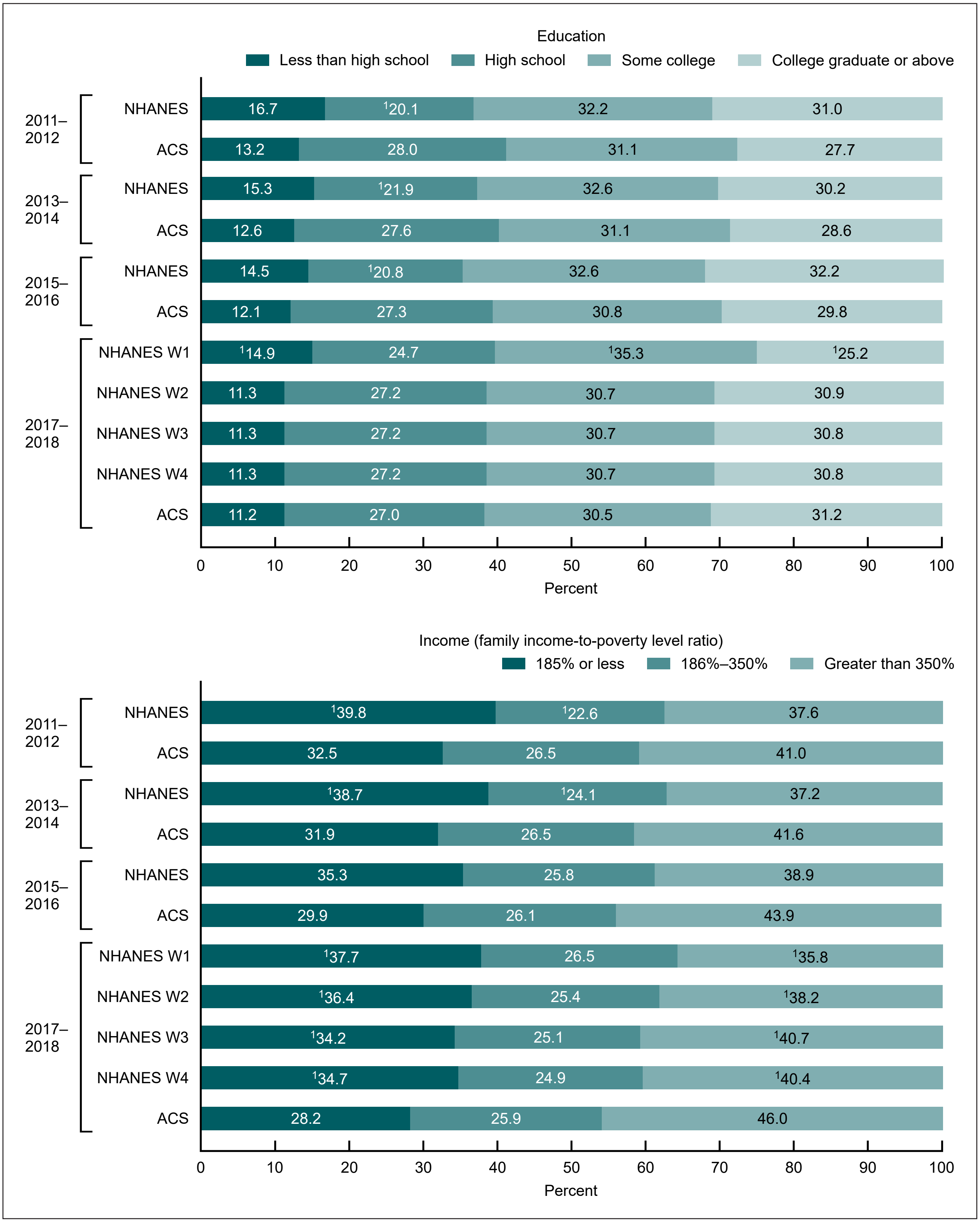
Comparisons With External Data Sources, Benchmarking, and Comparing Alternative Weighting Adjustments

Comparisons with estimates from the American Community Survey

The comparison of NHANES weighted estimates with those from ACS was iterative, and the results were used to guide additional weighting adjustments. First, NHANES education and income estimates were computed after the initial weighting process (W1), which used the same weighting approach that was used in prior NHANES cycles. Differences between NHANES estimates using the initial weights and ACS estimates for education and income may indicate the presence of bias or an imbalance in the sample. Based on the results from these initial comparisons, additional weighting adjustments were explored, and the newly developed weights were then used for additional rounds of comparisons with ACS. Results of the comparisons between the different weighting adjustments (W1 compared with W2, W3, and W4) against ACS estimates are shown in Figure 4.

Weighted NHANES estimates for education and income levels were compared with those from ACS separately for the 2011–2012, 2013–2014, 2015–2016, and 2017–2018 cycles. This comparison served two functions. The analysis was first performed using the initial (W1) weights for NHANES 2017–2018. As shown in Figure 4, the 2017–2018 NHANES sample had a lower percentage of college graduates and a higher percentage of adults with some college and adults

Figure 4. Comparison of National Health and Nutrition Examination Survey and American Community Survey estimates for education and income: 2011–2018



[1]Statistically significant difference at $p < 0.05$ between the NHANES and ACS estimate.
NOTES: NHANES is the National Health and Nutrition Examination Survey. ACS is American Community Survey. W is weight adjustment. Totals may not add to 100 due to rounding.
SOURCES: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2011–2018; U.S. Census Bureau, American Community Survey, 2011–2018.

with less than a high school degree compared with ACS. Specifically, the W1-adjusted NHANES estimate of 25.2% of adults aged 20 and over with a college degree or higher was statistically significantly smaller than the ACS estimate of 31.2%, and the W1-adjusted NHANES estimates of those with less than a high school degree (14.9%) and those with some college (35.3%) were significantly higher than the ACS estimates of 11.2% and 30.5%, respectively. Some differences in the estimates may be the result of differences in survey administration (i.e., ACS is a multimode survey and NHANES is an in-person survey). However, in 2017−2018, the difference for college graduates was larger and in the opposite direction than what was observed in previous cycles.

Similar comparisons were conducted for income levels using the FIPR. As shown in Figure 4, the 2017−2018 NHANES sample (W1) had a higher percentage of adults with lower income levels (37.7% at 185% FIPR or less) and a lower percentage of adults with higher income levels (35.8% at greater than 350% FIPR) compared with ACS (28.2% and 46.0%, respectively).

These findings indicated that the 2017−2018 NHANES initial weighting adjustments (W1) did not reduce bias to the same extent as was observed in previous cycles. For example, the absolute difference between the NHANES and ACS estimates for the percentage of college graduates was 3.3% in 2011−2012 (31.0% compared with 27.7%, respectively), 1.6% in 2013−2014 (30.2% compared with 28.6%, respectively), and 2.4% in 2015−2016 (32.2% compared with 29.8%, respectively) compared with 6.0% in 2017−2018 with the W1-adjusted estimate (25.2% versus 31.2%, respectively). For income, the absolute difference between the ACS and NHANES estimates for the highest income category (greater than 350% FIPR) was 3.4% in 2011−2012 (41.0% compared with 37.6%, respectively), 4.4% in 2013−2014 (41.6% compared with 37.2%, respectively), and 5.0% in 2015−2016 (43.9% compared with 38.9%, respectively) compared with 10.2% in 2017−2018 with the W1-adjusted estimate (46.0% compared with 35.8%, respectively). The proportion of missing values for income in the W1-weighted NHANES data increased from 6.4% in 2011−2012 to 11.4% in 2017−2018 (data not shown); therefore, observed differences could be the result of item nonresponse as well as unit nonresponse.

In general, the findings suggested that additional calibrations to education and income levels may be warranted and may help reduce the risk of nonresponse bias. Based on the results from these analyses, the initial sample weight (W1) was raked to ACS education levels (W2). A review of the income levels after the education adjustment indicated that the NHANES estimate of higher income population was still lower than expected, so further adjustments by income were needed. Different methods to calibrate by income were explored, including using a GREG adjustment for PSU-level income (W3) and using area-level mean household income deciles for raking (W4).

As shown in Figure 4, raking the 2017−2018 NHANES weights for education levels—as was done for sample weights W2, W3, and W4—balanced the NHANES sample to ACS estimates on education. In other words, no significant differences in education level estimates were observed between NHANES and ACS for 2017−2018 when using W2, W3, or W4.

Additional adjustments for income reduced, but did not eliminate, the imbalance between NHANES and ACS on income. For example, the final W4-weighted NHANES estimate of the percentage of the population with an FIPR of 185% or less was still significantly higher than the ACS estimate (34.7% compared with 28.2%, respectively), and conversely, the final NHANES estimate of the percentage of the population with FIPR greater than 350% was still significantly lower than the ACS estimate (40.4% compared with 46.0%, respectively). However, the differences were substantially reduced through the weighting adjustments, and were similar to prior cycles.

Both the W3 and W4 adjustments were effective in mitigating nonresponse bias related to education and income. However, the variance of the estimates after the W3 adjustment was greater than the variance of the estimates after the W4 adjustment. Figure 5 shows the SEs of 81 health estimates in domains formed by crossing age, race and Hispanic origin, and sex (e.g., non-Hispanic white male adult obesity or total adult diabetes) after the W3 adjustment and the W4 adjustment, relative to W1. If the SE of a W3 or a W4 estimate was equivalent to the SE of the W1 estimate, the data point on the scatter plot would fall on the solid diagonal line. Figure 5 shows that data points on the scatter plot for W3 are shifted farther up, and W3 has more points above the dashed line that passes through the origin with a slope of 1.3, which indicates that the W3 adjustment inflated the SEs to a greater extent than W4 (the dashed line indicates an increase of 30% from W1 SEs). On average, the SEs for estimates using W3 increased 7% from those using W1, whereas the SEs for estimates using W4 increased only 2%. These observations indicate that the variance of the estimates after the GREG adjustment (W3) was greater than the variance of the estimates after raking to income deciles (W4). A greater inflation in variance using W3 was anticipated because the W3 adjustment occurred at the PSU (county) level with a small number of PSUs, and the W4 adjustment occurred at the smaller tract level across PSUs. For this reason, the adjustment for education and income by raking (W4) was chosen to create the final set of 2017−2018 weights.

Benchmarking to the National Health Interview Survey

Figure 6 shows the results of the analysis comparing W1- and W4-adjusted NHANES health insurance coverage, diagnosed diabetes, and disability estimates with those from NHIS for 2011–2012, 2013–2014, 2015–2016, and 2017–2018.

The proportion of uninsured persons under age 65 was consistently higher for NHANES compared with NHIS in all four cycles; however, the difference was not statistically significant in 2017–2018 for either the W1 or the W4 NHANES estimates. Estimates of disability differed at each cycle, with higher prevalence estimates reported from NHANES in all cycles. While no significant differences were found in the prevalence of diagnosed diabetes for adults aged 20 and over between the two surveys in prior cycles, the W1 and W4 NHANES estimates were both significantly higher in 2017–2018 compared with NHIS, although a similar but not statistically significant difference was seen in the 2015–2016 NHANES when compared with the W4 NHANES 2017–2018 estimates.

The 2017–2018 W4-weighted NHANES estimates were closer to the NHIS estimates compared with the W1-weighted NHANES estimates for all three variables. These results provide additional evidence that enhanced weighting adjustments that included calibration to education and income, along with race and Hispanic origin, sex, and age groups reduced nonresponse bias.

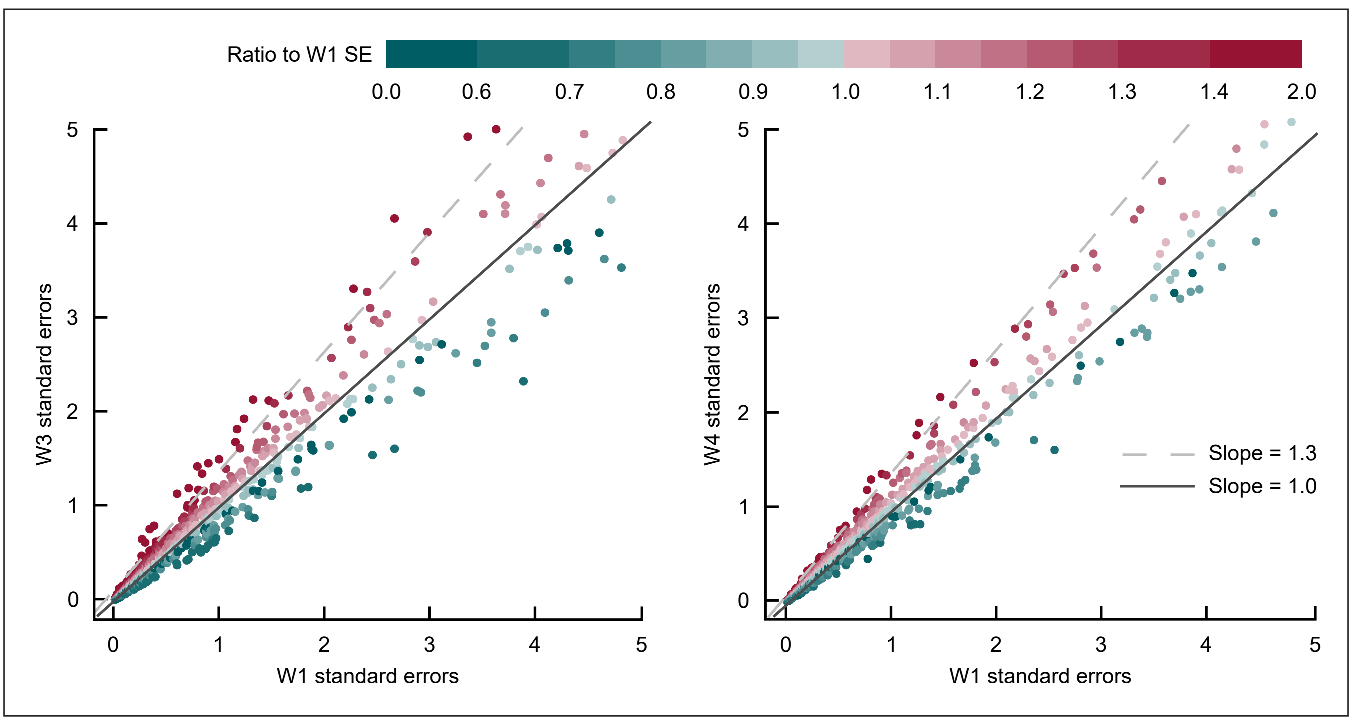
Survey Location Sampling Variability Results

The nonresponse bias analyses of the initial weights (W1) at the respondent level indicated that levels of educational attainment and income deviated from the three previous cycles to some degree. To examine whether the 2017–2018 PSU sample affected these deviations, additional analyses at the PSU level were conducted using the three approaches below.

2017–2018 primary sampling units compared with primary sampling units selected in prior cycles

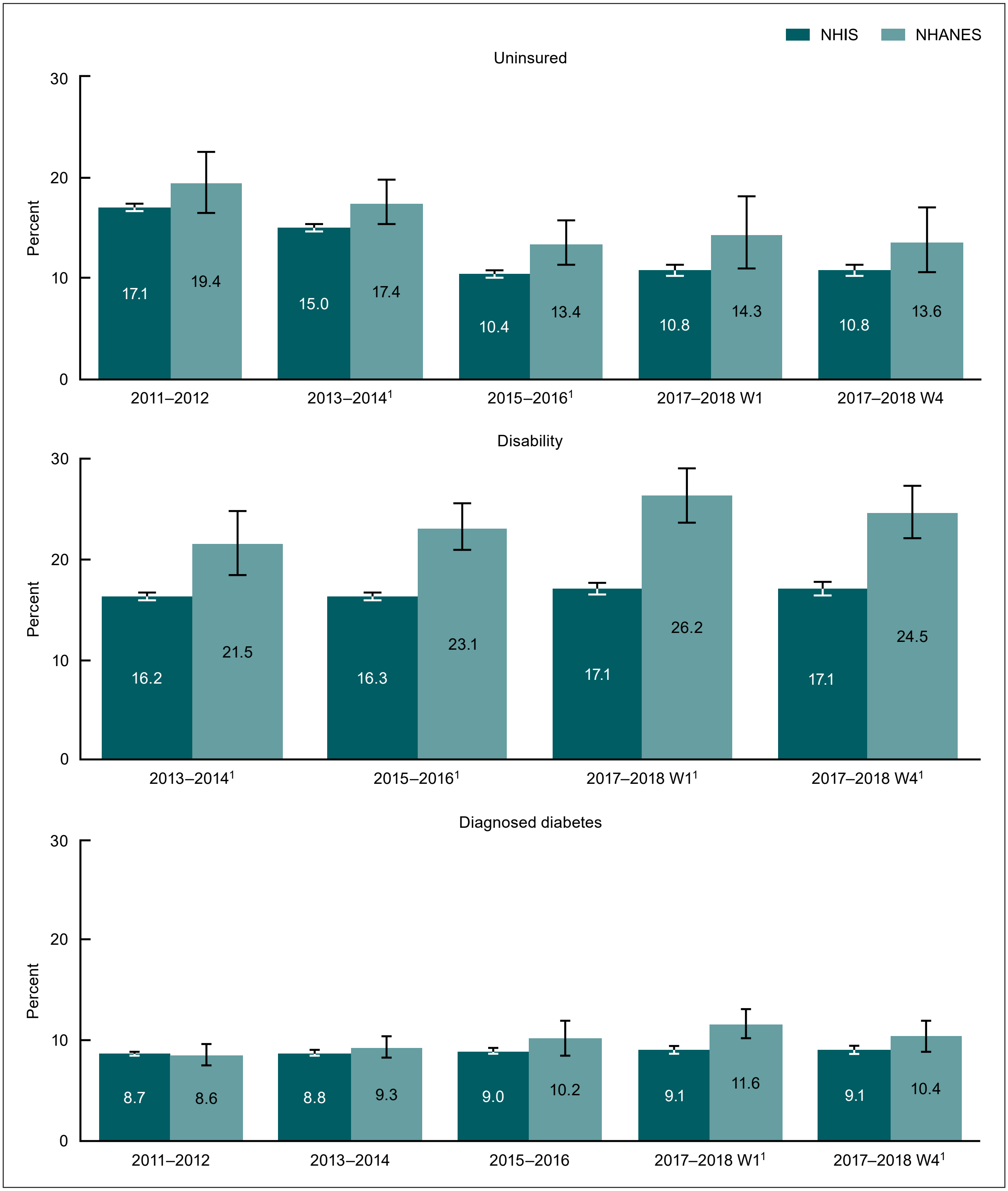
Table B shows the average percentage of the population with a college degree and the average mean household income of the 30 PSUs selected in each of the last four cycles of NHANES, starting in 2011. All of the county-level data used for this analysis were obtained from a single data source, the 5-year ACS released in 2017, so differences between cycles are due only to the PSUs in each cycle, not to changes over time. Both the unweighted and weighted estimates (i.e., weighted by PSU probability of selection) and the standard

Figure 5. Standard errors of 81 W3 and W4 estimates relative to W1 estimates: National Health and Nutrition Examination Survey, 2017–2018



NOTES: W is weight adjustment. SE is standard error. Red points are for estimates with higher SEs after the W3 or W4 adjustment, and green points are for estimates with lower SEs after W3 or W4 adjustment. The ratio of a W3 or W4 SE to the corresponding W1 SE determined the intensity of color, with lighter points where the ratio is close to one, and darker points where the ratio is further away.
SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2011–2018.

Figure 6. Comparison of National Health Interview Survey and National Health and Nutrition Examination Survey estimates for uninsured persons under age 65, adults with a disability, and adults with diagnosed diabetes: 2011–2018



[1]Statistically significant difference at $p < 0.05$ between the NHIS and NHANES estimate.
NOTES: W is weight adjustment. NHIS is National Health Interview Survey. NHANES is National Health and Nutrition Examination Survey. Estimates are age adjusted by the direct method to the 2000 projected U.S. population using age groups 20–39, 40–59, and 60 and over (for diagnosed diabetes and disability) or age groups 0–19, 20–39, and 40–64 years (for percent uninsured). Questions used for the assessment of disability differed between NHIS and NHANES in 2011–2012. Sample weights W1 and W4 were used in computing the NHANES estimates.
SOURCES: National Center for Health Statistics, National Health Interview Survey and National Health and Nutrition Examination Survey, 2011–2018 for uninsured status and diagnosed diabetes and 2013–2018 for disability.

Table B. Average of the percentage of the population aged 25 and over with a college degree and average of mean income for the 30 primary sampling units selected in each of the last four cycles of the National Health and Nutrition Examination Survey, 2011–2018

NHANES cycle	Mean household income (U.S. dollars)			Percent of the population with a college degree		
	Unweighted average	Weighted average	SE	Unweighted average	Weighted average	SE
2011–2012	86,760	66,838	3,540	33.2	22.0	2.3
2013–2014	84,333	66,355	5,400	31.6	21.5	3.5
2015–2016	82,311	69,287	2,953	31.8	23.6	3.3
2017–2018	78,189	60,602	3,034	27.1	17.6	1.3

NOTES: NHANES is National Health and Nutrition Examination Survey. SE is standard error. The weighted estimates were weighted by primary sampling unit probability of selection.

SOURCE: American Community Survey, 2017, 5-year estimates.

error of the weighted estimate are shown. If the sample of 30 PSUs selected in the 2017−2018 cycle was representative, the average percentage of the population with a college degree and the average mean household income of the sampled PSUs should be similar to those selected in prior cycles.

The unweighted and weighted averages of mean household income were lower for the 30 PSUs in the 2017−2018 cycle compared with prior cycles. Likewise, the percentage of the adult population aged 25 and over with a college degree or higher was lower in 2017−2018 (27.1% unweighted, 17.6% weighted) compared with previous cycles. While these differences were not statistically significant, the larger differences between the 2017−2018 cycle and the previous cycles support evidence from the preceding analysis that showed weighting adjustments by income and education were needed.

Table C shows the health characteristics for the 30 PSUs selected in each of the last four cycles of NHANES. The county-level data used for this analysis were obtained from the 2019 County Health Rankings & Roadmaps data set as described in the methods section (28), so differences between cycles are due only to the PSUs sampled in each cycle, not to changes over time. For each of the previous four cycles of NHANES, the average years of potential life lost (YPLL), the average percentage of adults who were uninsured, who had obesity, and who were diagnosed with diabetes for all 30 counties per cycle were estimated. No significant differences were detected in these health outcomes, both weighted and unweighted, for the 30 PSUs selected in each of the previous three cycles of NHANES compared with the 2017−2018 cycle, except for the prevalence of diabetes where the average of the 30 PSUs in the 2015−2016 cycle was statistically significantly lower than in 2017−2018 (i.e., 10.3% compared with 11.7%).

Primary sampling unit-level characteristics compared with the sampling stratum average

Because demographic data were obtained from ACS for every county, the average percentage of the population aged 25 and over with a college degree and the average of mean household income were computed for all the counties within each of the 56 strata in the 2015−2018 sample design. The same analysis was done for the strata in the 2011−2014 sample design. Each stratum average was compared with the value for the county that was selected from the stratum.

Table C. Differences in the average health characteristics of the 30 primary sampling units selected in each of the last four National Health and Nutrition Examination Survey cycles: 2011–2018

NHANES cycle	Years of potential life lost			Adult uninsured			Adult obesity			Adult diabetes		
	Unweighted mean	Weighted mean	SE	Unweighted mean	Weighted mean	SE	Unweighted mean	Weighted mean	SE	Unweighted mean	Weighted mean	SE
	Number			Percent								
2011–2012	6,611	7,961	258	10.0	11.8	1.3	28.3	32.9	1.4	9.8	11.2	0.5
2013–2014	6,711	8,689	840	9.8	9.8	0.8	28.0	32.3	1.2	9.7	10.9	0.6
2015–2016	6,841	7,175	472	11.5	10.7	1.3	28.0	31.9	0.9	9.8	10.3	0.4
2017–2018	7,368	8,379	451	10.5	10.1	1.2	29.3	32.8	0.9	10.7	[1]11.7	0.4

[1]Statistically significant difference at $p < 0.05$ between the 2015–2016 and 2017–2018 estimate.

NOTES: NHANES is National Health and Nutrition Examination Survey. SE is standard error. The weighted estimates were weighted by primary sampling unit probability of selection. Years of potential life lost before age 75 per 100,000 population (age adjusted) were calculated from 2015–2017 mortality data from the National Center for Health Statistics' National Vital Statistics System; insurance coverage estimates are from the U.S. Census Bureau's Small Area Health Insurance Estimates Program; adult obesity and adult diagnosed diabetes estimates are from the Centers for Disease Control and Prevention's Diabetes Interactive Atlas.

SOURCES: National Center for Health Statistics, National Vital Statistics System, Mortality, 2015–2017; U.S. Census Bureau, Small Area Health Insurance Estimates Program; and Centers for Disease Control and Prevention, Diabetes Interactive Atlas.

A difference greater than zero indicated that the selected PSU had a higher percentage of the population with a college degree or higher mean household income than the average of the PSUs in the stratum from which this PSU was selected. A difference less than zero indicated the value for the selected PSU was lower than the average of the PSUs in the stratum.

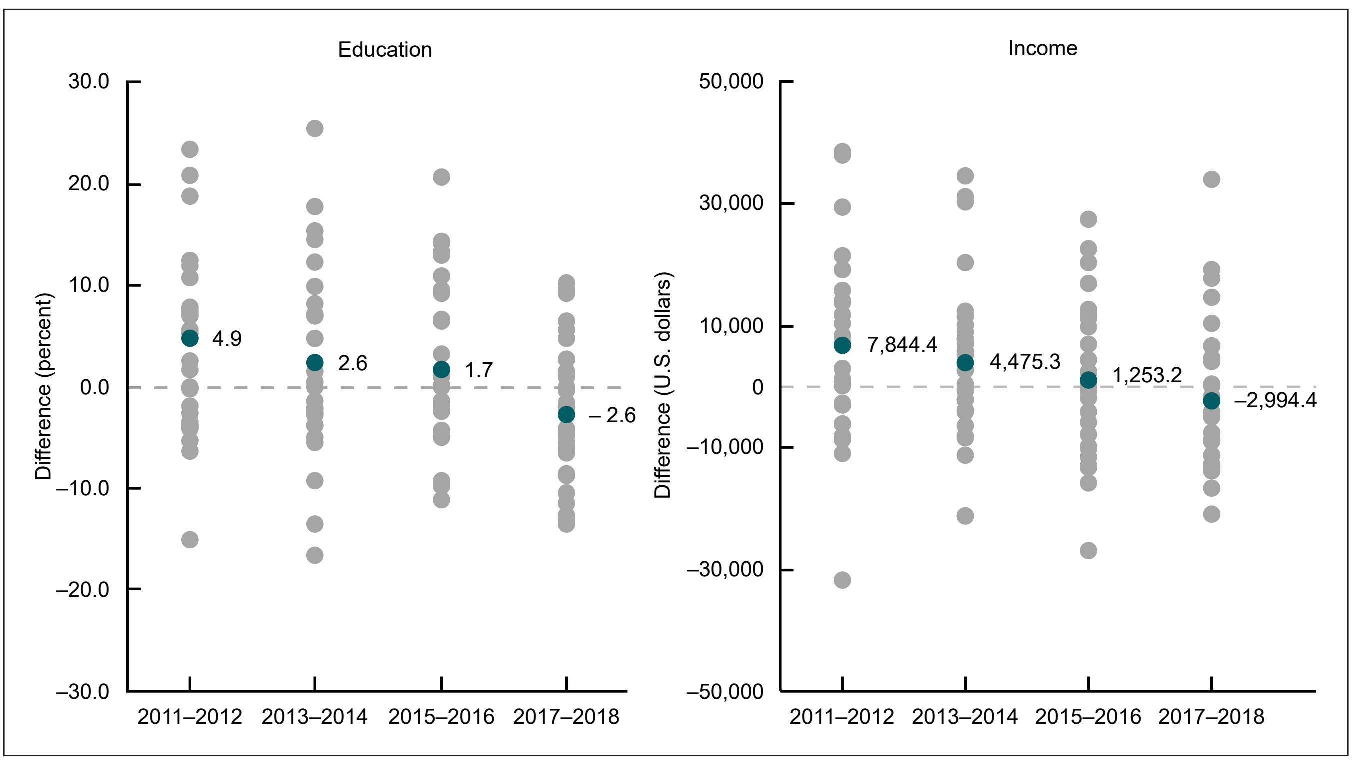
Figure 7 shows the differences between selected PSU and stratum average for household income and education levels for the PSUs in the last four cycles of NHANES. The differences for each comparison are shown as gray circles, and green circles represent the average of the differences for the cycle. The results show that the 2017–2018 selected sample of PSUs had lower household incomes and a lower percentage of the population with a college degree than would be expected. On the other hand, selected samples in prior cycles had higher household incomes and percentages of the population with a college degree than would be expected. However, it should be noted that the average of the difference for the combined 2015–2016 and 2017–2018 cycles, which are the two halves of the 2015–2018 sample design, is closer to zero than each cycle individually.

Figure 8 shows the differences in health outcomes between the selected PSU and the stratum average for the last four cycles of NHANES. YPLL (a proxy for premature mortality), the proportion of uninsured persons, the proportion of adults with obesity, and the proportion of adults with diabetes were used for these analyses. The differences between a selected PSU and the stratum average are shown as gray circles, and green circles represent the average of the differences per cycle. The results show that, on average, the 2017–2018 selected sample had a greater number of YPLL and a greater percentage of adults with obesity and diagnosed diabetes than in prior cycles. However, these differences were small and within a reasonable range given the random variation associated with sample selection that was expected under the NHANES 2015–2018 sample design, and these differences were accounted for when alternative weighting adjustments were explored.

Primary sampling unit-level characteristics compared with simulated samples

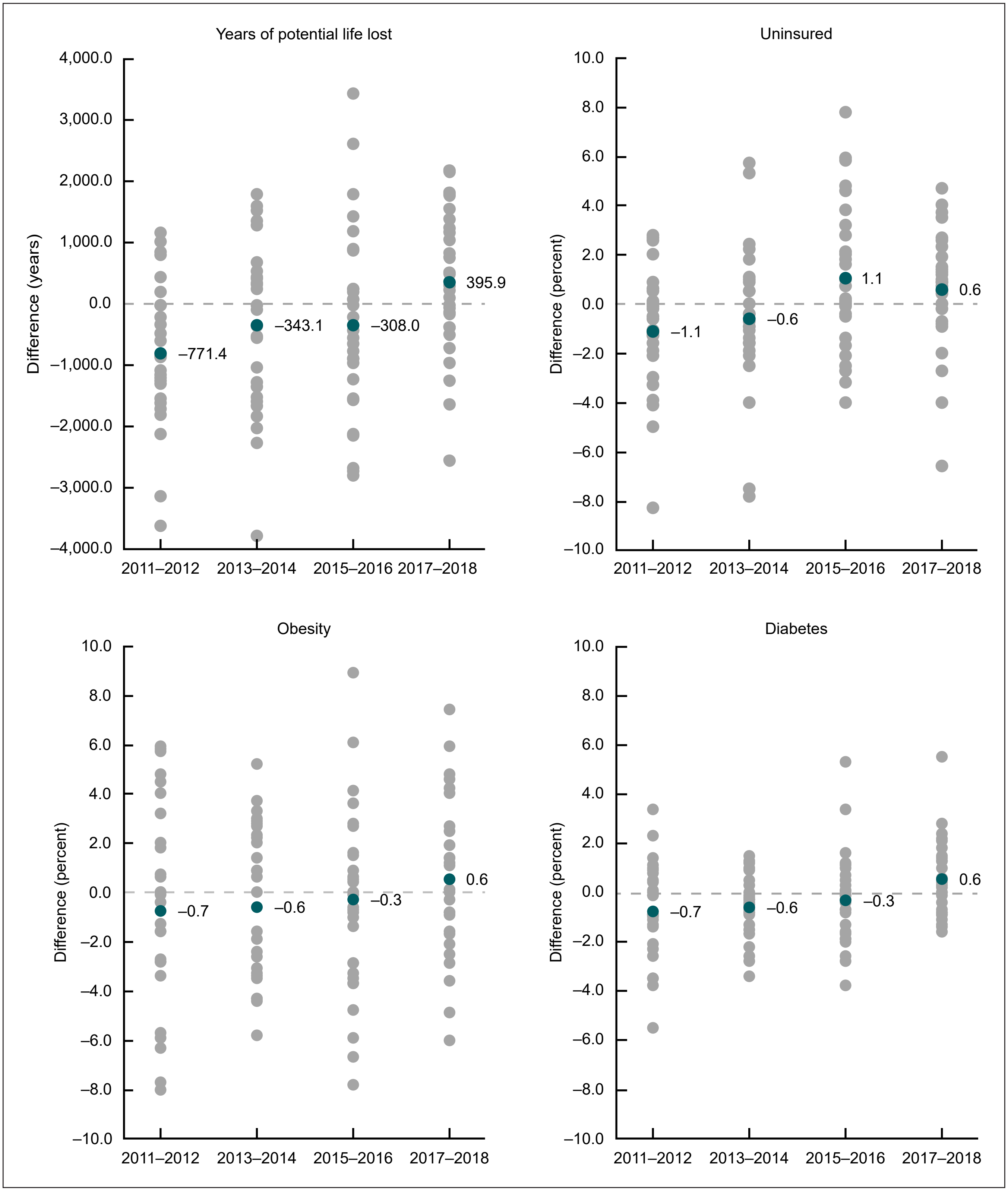
As a final analysis of the PSU-level characteristics in the 2017–2018 NHANES sample, 5,000 samples of PSUs were drawn using the 2015–2018 NHANES sample design. The demographic and health characteristics previously analyzed were weighted using the inverse of the PSU probability of selection and calculated for each of the 5,000 samples. The distributions of these characteristics are shown in Figure 9. For example, most of the samples had a mean household income estimate between $55,000 and $75,000, with a few

Figure 7. Differences in education and income levels for each of the 30 primary sampling units per cycle compared with all other primary sampling units in the same sampling stratum: National Health and Nutrition Examination Survey, 2011–2018



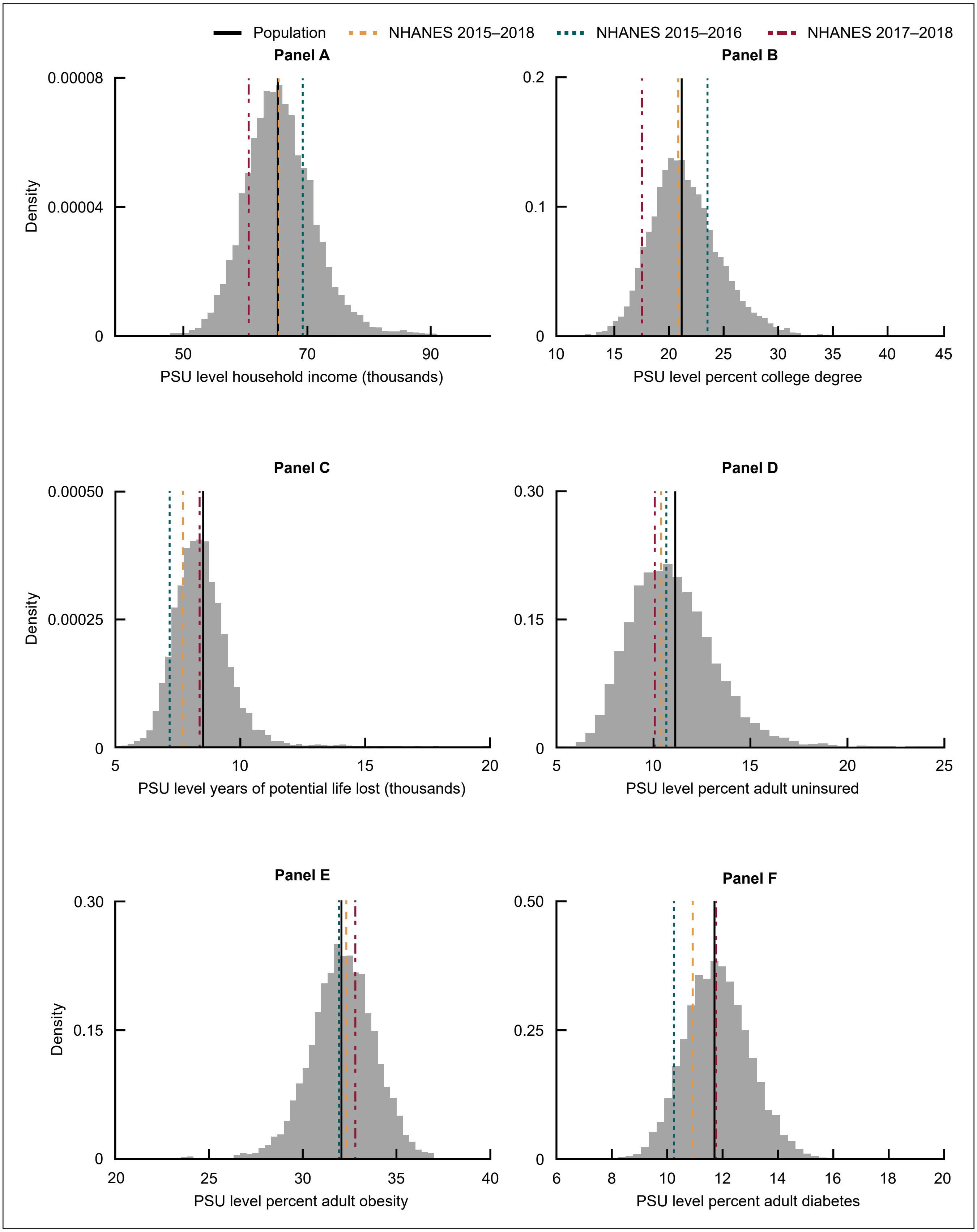
NOTES: The gray colored dots represent the differences between a selected primary sampling unit (PSU) and the average of all PSUs in the stratum from which that PSU was sampled. The green dots represent the average of all the differences per cycle. For education, the percentage of adults aged 25 and over with a bachelor's degree or higher is shown.
SOURCE: American Community Survey, 2017, 5-year estimates.

Figure 8. Differences in health outcomes for each of the 30 National Health and Nutrition Examination Survey primary sampling units per cycle compared with all other primary sampling units in the same sampling stratum: 2011–2018



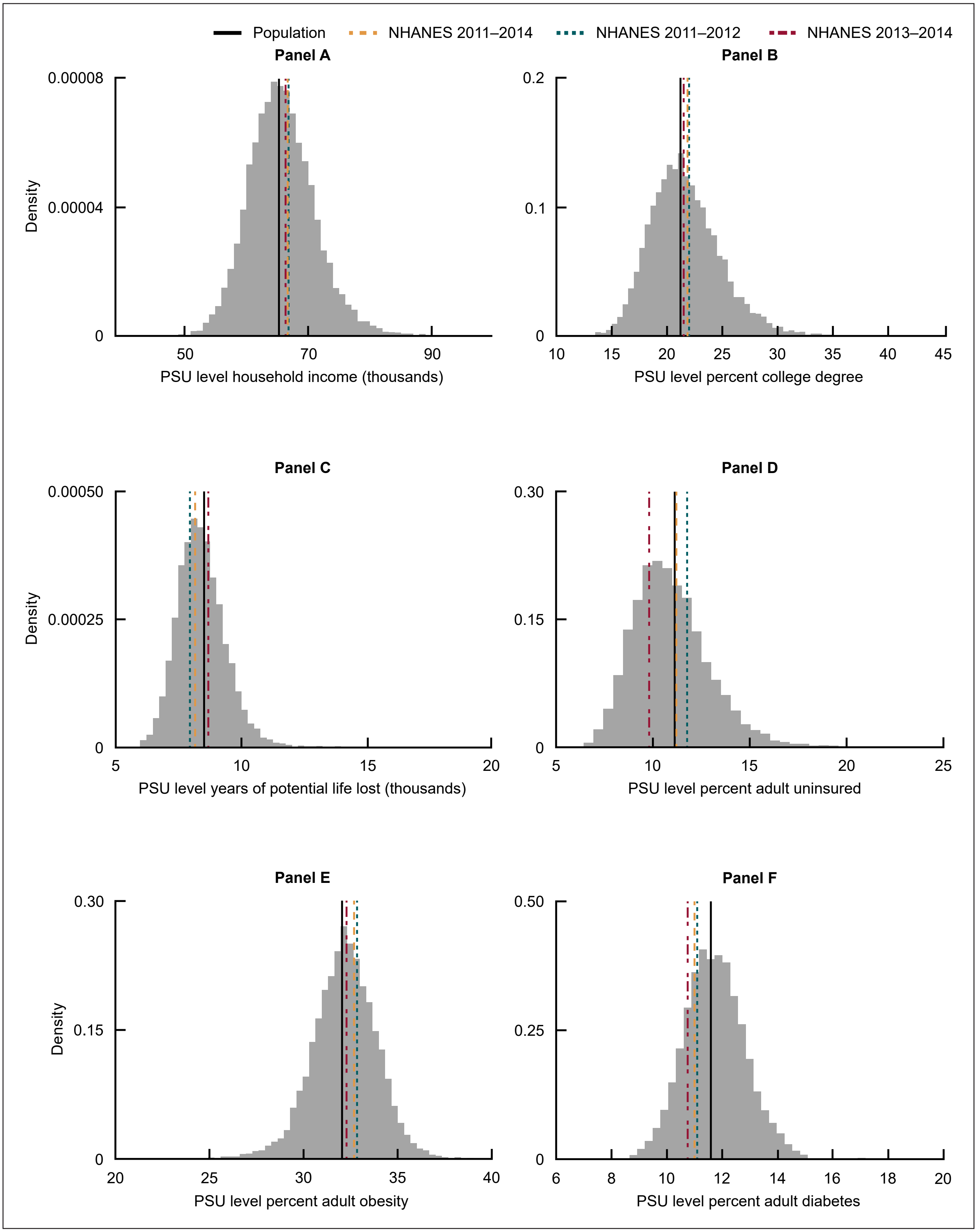
NOTES: The gray colored dots represent the differences between a selected primary sampling unit (PSU) and the average of all PSUs in the stratum from which that PSU was sampled. The green dots represent the average of all the differences per cycle. Years of potential life lost before age 75 per 100,000 population (age adjusted) were calculated from 2015–2017 mortality data from the National Center for Health Statistics' National Vital Statistics System; insurance coverage estimates are from the U.S. Census Bureau's Small Area Health Insurance Estimates Program; adult obesity and adult diagnosed diabetes estimates are from the Centers for Disease Control and Prevention Diabetes Interactive Atlas.
SOURCES: National Center for Health Statistics, National Vital Statistics System, Mortality, 2015–2017; U.S. Census Bureau, Small Area Health Insurance Estimates Program; and Centers for Disease Control and Prevention, Diabetes Interactive Atlas.

Figure 9. Distribution of demographic and health characteristics from 5,000 simulations of the 2015–2018 National Health and Nutrition Examination Survey sample compared with the actual sample



NOTES: PSU is primary sampling unit. NHANES is National Health and Nutrition Examination Survey.
SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2015–2018.

Figure 10. Distribution of demographic and health characteristics from 5,000 simulations of the 2011–2014 National Health and Nutrition Examination Survey sample compared with the actual sample



NOTES: PSU is primary sampling unit. NHANES is National Health and Nutrition Examination Survey.
SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2011–2014.

outliers beyond those levels. The 2017−2018 PSU sample had an average household income estimate of $60,602, which was lower than the 5,000 simulation population average estimate of $65,284 (Panel A). However, the 2017−2018 sample estimate was greater than 17.2% of the estimates in the simulation of 5,000 samples. Furthermore, the 2015−2016 sample had an average household income estimate of $69,287, and the 4-year 2015−2018 estimate was very close to the population estimate. Additionally, the 2017−2018 sample college degree estimate was greater than 8.9% of the estimates in the simulation of 5,000 samples (Panel B). On the other hand, the 2017−2018 sample obesity and diabetes estimates were greater than 65.2% and 50.8% of the estimates, respectively, in the simulation of 5,000 samples (Panels E and F).

For comparison, the same simulation analysis was performed for the 2011−2014 NHANES sample design and shown in Figure 10. Estimates from the 2-year cycles in that 4-year design are generally closer to each other and closer to the population values. For example, the 2011−2012 and 2013−2014 PSU samples had average household income estimates of $66,838 and $66,355, respectively (Panel A). Although the estimates are slightly higher than the population estimate of $65,284, they are much closer to the population estimate than the estimates from 2-year cycles of NHANES 2015−2018, as described earlier.

Discussion

Since 2011, NHANES response rates have been steadily falling, with more accelerated declines reported in recent years, especially in counties that have populations with higher levels of education and higher mean household incomes. To better understand the impact of these declines on the accuracy of the 2017−2018 NHANES estimates, nonresponse bias analyses were conducted and accompanied by an investigation of the effect of the selected survey locations in the 2017−2018 cycle on key estimates.

For the nonresponse bias analyses, several methodological approaches based on the Groves and Brick typology were used (9). While each nonresponse bias analysis has limitations, the use of multiple analyses can help corroborate findings and can provide insight into the patterns and potential for bias. These methods can also identify the most impactful variables for nonresponse weighting adjustments. First, the variation within the respondent set was examined by looking at the differences in response rates by subgroup characteristics at three levels: the SP level, the household level, and the geographic level. Among other variables, response rates were significantly associated with the sex and age of the SP, the number of people living in a household, and the educational attainment and income level of the area. These findings were suggestive of bias in the respondent sample, before any nonresponse adjustments, to the extent that these characteristics are related to health outcomes. As a result of these findings, some of these variables were used directly, or were correlated with other variables that were used in the final weighting adjustments for the 2017−2018 cycle of NHANES (7). For example, race and Hispanic origin, sex, and population size of the PSU, among others, were used to form nonresponse adjustment cells for the interview weights as was described in detail in the NHANES 2015–2018 sample design and estimation procedures report (7). Therefore, any potential bias due to these differences in response status was mitigated by the final weight adjustments (W4-adjusted weights).

Next, the R indicators were derived and used to gauge the representativeness of survey respondents. The R indicator analysis, which used base weights, did not show any evidence that the nonresponse bias before any weighting adjustment for NHANES 2017−2018 was larger than that for the earlier survey years, despite the declining response rates. In other words, the R indicators showed no significant differences across cycles. Although the R indicator is highly dependent on the variables that are included in the model and does not directly reflect bias in survey outcomes, some research has shown an association between the R indicator and nonresponse bias (12).

Next, level of effort analyses were used to examine any differences between easy-to-reach and hard-to-reach respondents. Hard-to-reach respondents are assumed to be a proxy for nonrespondents and differences between the easy- versus hard-to-reach respondents could indicate nonresponse bias. While easier-to-reach respondents had a higher prevalence of obesity, hypertension, or diabetes compared with harder-to-reach respondents, the cumulative estimates stabilized around the 8th contact attempt (i.e., additional contacts past the 8th attempt had a minimal effect on the estimates). These results indicate that nonrespondents likewise would have had a minimal effect on the estimates, to the extent that they are similar to hard-to-reach respondents.

The second approach to investigating nonresponse bias was comparing the weighted NHANES sample with ACS. The comparisons with ACS were focused on education and income because these two socioeconomic factors are known to be causally associated with health (34,35) and were strongly predictive of survey response, which makes these two variables strong candidates for nonresponse adjustments that may help directly address any potential bias in NHANES estimates. The comparisons after each adjustment were used to inform additional weighting adjustments. The W1-adjusted NHANES estimates for education and income were first compared with ACS. This initial sample weight (W1) was constructed using the same weighting approach as in previous cycles of NHANES. These comparisons showed that the 2017−2018 NHANES sample had a lower percentage of college graduates and a higher percentage of persons with less than a high school degree compared with ACS. Additionally, the NHANES sample had a lower proportion of higher-income persons compared with ACS. While some of the education and income differences between NHANES and ACS were observed in prior cycles, the magnitude of these differences was more pronounced in the 2017−2018 cycle. These findings indicated that the 2017−2018 NHANES initial sample weights (W1) did not reduce bias on these socioeconomic factors as much as the weights created in previous cycles and that additional weighting adjustments for education and income were needed.

As a result of the findings from the comparisons with ACS, a series of alternative weighting adjustments to reduce bias in the 2017−2018 NHANES cycle were explored. After the initial weighting process was applied (W1), the weights were adjusted to education (W2), and then further adjusted to income using two distinct approaches (W3 and W4). W3 used a GREG adjustment for PSU-level income, and W4 was raked to deciles of area-level mean household income. Both the W3 and W4 adjustments were effective in mitigating nonresponse bias by education and income. However, the variance of the estimates after the W3 adjustment was greater than the variance of the estimates after the W4 adjustment. For this reason, the adjustment for education and income by raking (W4) was chosen to create the final set of 2017−2018 weights.

The biases observed on income and education before the final weighting adjustments were exacerbated by the sample of counties that were randomly selected in 2017−2018. NHANES only samples 15 counties per year (i.e., 30 counties per cycle), making the survey susceptible to an outlier sample of counties in any given cycle. The 2017−2018 NHANES counties were atypical as a group—the sample of counties, on average, had lower proportions of college graduates and lower mean household incomes compared with prior cycles. The selected counties also had lower educational attainment and income levels compared with the average of each PSU's sampling stratum. Furthermore, in a simulation of 5,000 samples of counties under the same 2015−2018 sample design, the 2017−2018 counties were on the lower side of the distribution of possible samples on education and income, which could affect county-level health characteristics that are associated with education and income. For example, the sample of counties in the 2017−2018 cycle were below the 25th percentile on education and income. These observed differences in the socioeconomic characteristics of the 2017−2018 counties also tracked with poorer county-level health, which suggests that the NHANES participants from the 2017−2018 cycle were more likely to be from counties with poorer socioeconomic characteristics and health. While these ecological differences cannot be used to make an inference on the individual level, it suggests that the imbalances in the NHANES sample are a result of a combination of nonresponse bias that is related to socioeconomic factors and survey location sampling variability that resulted in a random sample of counties with lower levels of education and income.

Finally, the performance of the final sample weights (W4) was assessed by comparing NHANES estimates of insurance coverage, disability, and diagnosed diabetes with those from NHIS. The comparisons were not restricted to the 2017−2018 cycle, but also included prior cycles. These temporal trend analyses serve as a check that any differences observed between the two surveys were not unique to this last cycle but were consistent across the years. The differences between the surveys were consistent in the four cycles. The larger differences in the W1-adjusted NHANES estimates in 2017−2018 were attenuated by the education and income adjustments with W4. These observations provide evidence that differences between the two surveys were not due to bias in NHANES but likely arose from differences in survey administration and operations. One key difference between NHIS and NHANES is that NHANES includes standardized physical examinations in mobile examination centers as well as home interviews, which can potentially influence respondents' answers to the interview questions. A survey's "social setting" (i.e., the location of survey administration such as the respondent's home, a clinic, or a school, etc.) has been shown to influence survey responses in some instances (36−38), and this has been documented in NHANES previously (39). In addition, as a survey, NHIS is also subject to nonresponse and potential nonresponse bias. Therefore, differences between the surveys cannot be directly attributed to nonresponse bias in NHANES.

This extensive investigation suggests that bias in the outcome statistics was mostly reduced through the enhanced weighting adjustments, although the elimination of bias is impossible in any survey with missing data due to nonresponse. Differences between NHANES respondents and nonrespondents should not impact final 2017−2018 survey estimates any differently than in previous NHANES cycles.

References

1. National Center for Health Statistics. National Health and Nutrition Examination Survey. 2020. Available from: https://www.cdc.gov/nchs/nhanes/index.htm.
2. Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, et al. 2000 CDC growth charts for the United States: Methods and development. National Center for Health Statistics. Vital Health Stat Series 11(246). 2002.
3. Williams D, Brick JM. Trends in U.S. face-to-face household survey nonresponse and level of effort. J Surv Stat Methodol 6(2):186−211. 2018.
4. Tourangeau R, Brick JM, Lohr S, Li J. Adaptive and responsive survey designs: A review and assessment. J Royal Stat Soc Ser A Stat Soc 180(1):203−23. 2017.
5. Peytchev A. Consequences of survey nonresponse. Ann Am Acad Pol Soc Sci 645(1):88−111. 2013.

6. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: A meta-analysis. Public Opin Q 72(2):167–89. 2008.

7. Chen TC, Clark J, Riddles MK, Mohadjer LK, Fakhouri THI. The National Health and Nutrition Examination Survey, 2015–2018: Sample design and estimation procedures. National Center for Health Statistics. Vital Health Stat 2(184). 2020.

8. National Center for Health Statistics. National Health and Nutrition Examination Survey: Response rates and population totals. 2020. Available from: https://wwwn.cdc.gov/nchs/nhanes/responserates.aspx.

9. Groves RM, Brick JM. Practical tools for nonresponse bias studies. Short Course Materials. Joint Program in Survey Methodology. 2005.

10. Miller P, Fakhouri TH, Earp M, Piscopo KD, Frenk SM, Christopher E, Madans J. A systematic review of nonresponse bias studies in federally sponsored surveys. 2020. Available from: https://nces.ed.gov/fcsm/pdf/A_Systematic_Review_of_Nonresponse_Bias_Studies_Federally_Sponsored_SurveysFCSM_20_02_032920.pdf.

11. Rao JNK, Scott AJ. On simple adjustments to Chi-square tests with sample survey data. Annals Stat 15(1):385–97. 1987.

12. Schouten B, Cobben F, Bethlehem J. Indicators for the representativeness of survey response. Surv Methodol 35(1):101–113. 2008.

13. Curtin R, Presser S, Singer E. The effects of response rate changes on the index of consumer sentiment. Public Opin Q 64(4):413–28. 2000.

14. Lin IF, Schaeffer NC. Using survey participants to estimate the impact of nonparticipation. Public Opin Q 59(2):236–58. 1995.

15. Maitland A, Lin A, Cantor D, Jones M, Moser RP, Hesse BW, et al. A nonresponse bias analysis of the Health Information National Trends Survey (HINTS). J Health Commun 22(7):545–53. 2017.

16. U.S. Department of Health and Human Services. HHS poverty guidelines for 2020. Available from: https://aspe.hhs.gov/poverty-guidelines.

17. Torrieri N. American Community Survey design and methodology (January 2014). 2014. Available from: https://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/acs_design_methodology_report_2014.pdf.

18. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. Healthy People 2020. 2020. Available from: https://www.healthypeople.gov/2020.

19. National Center for Health Statistics. National Health Interview Survey. Available from: https://www.cdc.gov/nchs/nhis/index.htm.

20. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People Statistical Notes, no. 20. Hyattsville, MD: National Center for Health Statistics. 2001.

21. Korn EL, Graubard BI. Confidence intervals for proportions with small expected number of positive counts estimated from survey data. Surv Methodol 24(2):193–201. 1998.

22. Särndal CE, Swensson B, Wretman J. Model assisted survey sampling. New York: Springer-Verlag. 1992.

23. Rao JNK, Molina I. Small area models. In: Rao JNK, Molina I, editors. Small area estimation. Hoboken, NJ: Wiley. 75–96. 2015.

24. Estevao V, Hidiroglou MA, Särndal CE. Methodological principles for a generalized estimation system at Statistics Canada. J Off Stat 11(2):181–204. 1995.

25. Statistics Canada. Census of population 2016: Sampling and weighting technical report. 2018. Available from: https://www12.statcan.gc.ca/census-recensement/2016/ref/98-306/98-306-x2016001-eng.pdf.

26. Czajka JL, Denmead G. Income data for policy analysis: A comparative assessment of eight surveys. 2008. Available from: https://www.mathematica.org/our-publications-and-findings/publications/income-data-for-policy-analysis-a-comparative-assessment-of-eight-surveys.

27. Angel S, Disslbacher F, Humer S, Schnetzer M. What did you really earn last year?: Explaining measurement error in survey income data. J R Stat Soc Ser A Stat Soc 182(4):1411–37. 2019.

28. University of Wisconsin Population Health Institute. County health rankings & roadmaps. 2020. Available from: https://www.countyhealthrankings.org.

29. National Center for Health Statistics. National Vital Statistics System. Available from: https://www.cdc.gov/nchs/nvss/index.htm.

30. United States Census Bureau. Small area health insurance estimates program. Available from: https://www.census.gov/programs-surveys/sahie.html.

31. Centers for Disease Control and Prevention. United States diabetes surveillance system. Available from: https://gis.cdc.gov/grasp/diabetes/diabetesatlas.html.

32. Centers for Disease Control and Prevention. Behavioral risk factor surveillance system. Available from: https://www.cdc.gov/brfss/index.html.

33. United States Census Bureau. Population estimates program. Available from: https://www.census.gov/programs-surveys/popest.html.

34. Schultz WM, Kelli HM, Lisko JC, Varghese T, Shen J, Sandesara P, et al. Socioeconomic status and cardiovascular outcomes: Challenges and interventions. Circulation 137(20):2166–78. 2018.

35. Galea S, Ahern J, Karpati A. A model of underlying socioeconomic vulnerability in human populations: Evidence from variability in population health and implications for public health. Soc Sci Med 60(11):2417–30. 2005.

36. Brener ND, Eaton DK, Kann L, Grunbaum J, Gross LA, Kyle TM, Ross JG. The association of survey setting and mode with self-reported health risk behaviors among high school students. Public Opin Q 70(3):354–74. 2006.

37. Bowling A. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Oxf) 27(3):281–291. 2005.

38. Wildman RC. Effects of anonymity and social setting on survey responses. Public Opin Q 41(1):74–79. 1977.

39. Zajacova A, Dowd JB. Reliability of self-rated health in U.S. adults. Am J Epidemiol 174(8):977–83. 2011.

Appendix. Supporting Table

Table. Interview response rates, by potential weighting variables: National Health and Nutrition Examination Survey, 2017–2018

Variable	Base-weighted interview response rate
Sex ($p = 0.0001$)	
Male	54.3
Female	58.0
Age (years) ($p < 0.0001$)	
5 and under	67.6
6–19	63.9
20–59	53.8
60 and over	50.9
Race and Hispanic origin and income ($p < 0.0001$)	
Black, all income levels	63.9
Hispanic, all income levels	60.3
Asian, all income levels	49.0
White or other and low income	65.6
White or other and not low income	49.9
Children in household ($p < 0.0001$)	
No	50.1
Yes	61.7
Household size ($p = 0.0001$)	
1–2	51.9
3–4	56.0
5–6	60.3
7 or more	70.2
Number of SPs in household ($p = 0.0101$)	
1	56.9
2	51.6
3	56.0
4	57.8
5 or more	61.3
Household reference person's sex ($p < 0.0001$)	
Male	51.8
Female	61.3
Household composition of SPs by age ($p < 0.0001$)	
1 SP in HH, under 16	62.1
1 SP in HH, 16 or over	56.1
More than 1 SP in HH, all under 16	66.7
More than 1 SP in HH, all 16 or over	48.5
More than 1 SP in HH, mixed ages	61.9
Census region ($p = 0.1895$)	
Northeast	49.7
Midwest	60.5
South	57.4
West	54.8

See footnotes at end of table.

Table. Interview response rates, by potential weighting variables: National Health and Nutrition Examination Survey, 2017–2018—Con.

Variable	Base-weighted interview response rate
Health-based state groups ($p = 0.0696$)	
Best health (AZ, CA, HI, MA, NH, NY, OR, UT, VT, WA)	52.3
2nd best health (CO, CT, DE, FL, ID, ME, MN, MT, NJ, NM, NV, PA, RI)	54.0
3rd best health (AK, DC, GA, IA, IL, MD, MI, ND, NE, OK, SD, TX, VA, WI, WY)	59.5
4th best health (AL, AR, IN, KS, KY, LA, MO, MS, NC, OH, SC, TN, WV)	58.1
PSU population size ($p < 0.0001$)	
100,000 or less	64.5
100,001–250,000	57.6
250,001–1,000,000	53.8
1,000,000 or more	48.0
Percentage of population aged 25 and over with a college education or higher ($p = 0.0001$)	
1st quartile	60.9
2nd quartile	58.7
3rd quartile	54.5
4th quartile	49.5
Median household income ($p < 0.0001$)	
1st quintile	64.7
2nd quintile	57.9
3rd quintile	57.3
4th quintile	53.7
5th quintile	45.8
Percentage of households that received food stamps in the past 12 months ($p < 0.0001$)	
1st quartile	48.0
2nd quartile	52.7
3rd quartile	58.3
4th quartile	62.6
Percentage of occupied housing units that are owned ($p = 0.0248$)	
1st quartile	56.9
2nd quartile	59.6
3rd quartile	56.2
4th quartile	51.9
Percentage of population who are uninsured ($p = 0.0131$)	
1st quartile	50.9
2nd quartile	55.9
3rd quartile	57.2
4th quartile	60.8
Percentage of population with disability ($p = 0.0201$)	
1st quartile	53.1
2nd quartile	53.1
3rd quartile	56.8
4th quartile	60.4
Percentage of population aged 18 and over with disability ($p = 0.0050$)	
1st quartile	52.4
2nd quartile	52.6
3rd quartile	58.7
4th quartile	59.7

See footnotes at end of table.

Table. Interview response rates, by potential weighting variables: National Health and Nutrition Examination Survey, 2017–2018—Con.

Variable	Base-weighted interview response rate
Percentage of population that is black alone or in combination with one or more races ($p = 0.0272$)	
1st quartile	59.0
2nd quartile	52.2
3rd quartile	53.4
4th quartile	60.3
Percentage of population that is Asian alone or in combination with one or more races ($p < 0.0001$)	
1st quartile	62.0
2nd quartile	56.1
3rd quartile	53.6
4th quartile	50.7
Percentage of population that is Hispanic ($p = 0.1481$)	
1st quartile	58.9
2nd quartile	54.0
3rd quartile	54.8
4th quartile	56.8
Percentage of population born in the United States ($p = 0.0026$)	
1st quartile	55.4
2nd quartile	51.5
3rd quartile	54.2
4th quartile	61.8
Sampling segment median age ($p = 0.2048$)	
1st quartile	59.7
2nd quartile	56.3
3rd quartile	56.6
4th quartile	53.0
Percentage of population under age 18 ($p = 0.3039$)	
1st quartile	52.9
2nd quartile	55.4
3rd quartile	55.5
4th quartile	59.3
Percentage of households that are family households ($p = 0.0723$)	
1st quartile	58.3
2nd quartile	58.1
3rd quartile	56.8
4th quartile	51.3
Percentage of households with three or more people ($p = 0.0296$)	
1st quartile	59.3
2nd quartile	54.0
3rd quartile	58.9
4th quartile	52.2
Percentage of workforce population that have a 30-minute or longer commute to work ($p = 0.1418$)	
1st quartile	58.5
2nd quartile	57.3
3rd quartile	54.1
4th quartile	52.2
Percentage of population with family-income-to-poverty ratio of 1.50 or higher ($p < 0.0001$)	
1st quartile	62.0
2nd quartile	59.3
3rd quartile	54.8
4th quartile	46.6

See footnotes at end of table.

Table. Interview response rates, by potential weighting variables: National Health and Nutrition Examination Survey, 2017–2018—Con.

Variable	Base-weighted interview response rate
Per capita income ($p < 0.0001$)	
1st quartile	62.4
2nd quartile	59.4
3rd quartile	52.4
4th quartile	46.6
Percentage of housing units occupied ($p = 0.1049$)	
1st quartile	58.5
2nd quartile	57.3
3rd quartile	56.3
4th quartile	51.2
Percentage of occupied housing units with 0.5 or less occupants per room ($p = 0.6082$)	
1st quartile	57.5
2nd quartile	57.3
3rd quartile	54.9
4th quartile	55.5
Percentage of occupied housing units with one to three rooms ($p = 0.1582$)	
1st quartile	53.6
2nd quartile	56.5
3rd quartile	58.6
4th quartile	55.2
Percentage of housing units that are one unit detached structures ($p = 0.0283$)	
1st quartile	58.0
2nd quartile	57.1
3rd quartile	58.3
4th quartile	51.6
Percentage of housing units with 50 or more units in the structure ($p = 0.4849$)	
1st quartile	55.6
2nd quartile	58.0
3rd quartile	57.2
4th quartile	54.2
Percentage of housing units built in 2010 or later ($p = 0.6182$)	
1st quartile	55.0
2nd quartile	57.2
3rd quartile	57.6
4th quartile	55.1
Percentage of occupied housing units without a mortgage ($p = 0.4225$)	
1st quartile	57.7
2nd quartile	54.0
3rd quartile	55.2
4th quartile	57.6
Percentage of households with a monthly housing cost greater than $2,000 ($p < 0.0001$)	
1st quartile	62.7
2nd quartile	58.7
3rd quartile	51.0
4th quartile	46.6

See footnotes at end of table.

Table. Interview response rates, by potential weighting variables: National Health and Nutrition Examination Survey, 2017–2018—Con.

Percentage of population in group quarters (p = 0.1013)	
1st quartile	56.9
2nd quartile	51.1
3rd quartile	55.3
4th quartile	59.0

NOTES: SP is sampled person. HH is household. PSU is primary sampling unit. p values are shown for the chi-square test of the relationship between response status and each of the subgroup variables.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2017–2018.

Vital and Health Statistics Series Descriptions

Active Series

Series 1. **Programs and Collection Procedures**
Reports describe the programs and data systems of the National Center for Health Statistics, and the data collection and survey methods used. Series 1 reports also include definitions, survey design, estimation, and other material necessary for understanding and analyzing the data.

Series 2. **Data Evaluation and Methods Research**
Reports present new statistical methodology including experimental tests of new survey methods, studies of vital and health statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. Reports also include comparison of U.S. methodology with those of other countries.

Series 3. **Analytical and Epidemiological Studies**
Reports present data analyses, epidemiological studies, and descriptive statistics based on national surveys and data systems. As of 2015, Series 3 includes reports that would have previously been published in Series 5, 10–15, and 20–23.

Discontinued Series

Series 4. **Documents and Committee Reports**
Reports contain findings of major committees concerned with vital and health statistics and documents. The last Series 4 report was published in 2002; these are now included in Series 2 or another appropriate series.

Series 5. **International Vital and Health Statistics Reports**
Reports present analytical and descriptive comparisons of U.S. vital and health statistics with those of other countries. The last Series 5 report was published in 2003; these are now included in Series 3 or another appropriate series.

Series 6. **Cognition and Survey Measurement**
Reports use methods of cognitive science to design, evaluate, and test survey instruments. The last Series 6 report was published in 1999; these are now included in Series 2.

Series 10. **Data From the National Health Interview Survey**
Reports present statistics on illness; accidental injuries; disability; use of hospital, medical, dental, and other services; and other health-related topics. As of 2015, these are included in Series 3.

Series 11. **Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey**
Reports present 1) estimates of the medically defined prevalence of specific diseases in the United States and the distribution of the population with respect to physical, physiological, and psychological characteristics and 2) analysis of relationships among the various measurements. As of 2015, these are included in Series 3.

Series 12. **Data From the Institutionalized Population Surveys**
The last Series 12 report was published in 1974; these reports were included in Series 13, and as of 2015 are in Series 3.

Series 13. **Data From the National Health Care Survey**
Reports present statistics on health resources and use of health care resources based on data collected from health care providers and provider records. As of 2015, these reports are included in Series 3.

Series 14. **Data on Health Resources: Manpower and Facilities**
The last Series 14 report was published in 1989; these reports were included in Series 13, and are now included in Series 3.

Series 15. **Data From Special Surveys**
Reports contain statistics on health and health-related topics from surveys that are not a part of the continuing data systems of the National Center for Health Statistics. The last Series 15 report was published in 2002; these reports are now included in Series 3.

Series 16. **Compilations of Advance Data From Vital and Health Statistics**
The last Series 16 report was published in 1996. All reports are available online; compilations are no longer needed.

Series 20. **Data on Mortality**
Reports include analyses by cause of death and demographic variables, and geographic and trend analyses. The last Series 20 report was published in 2007; these reports are now included in Series 3.

Series 21. **Data on Natality, Marriage, and Divorce**
Reports include analyses by health and demographic variables, and geographic and trend analyses. The last Series 21 report was published in 2006; these reports are now included in Series 3.

Series 22. **Data From the National Mortality and Natality Surveys**
The last Series 22 report was published in 1973. Reports from sample surveys of vital records were included in Series 20 or 21, and are now included in Series 3.

Series 23. **Data From the National Survey of Family Growth**
Reports contain statistics on factors that affect birth rates, factors affecting the formation and dissolution of families, and behavior related to the risk of HIV and other sexually transmitted diseases. The last Series 23 report was published in 2011; these reports are now included in Series 3.

Series 24. **Compilations of Data on Natality, Mortality, Marriage, and Divorce**
The last Series 24 report was published in 1996. All reports are available online; compilations are no longer needed.

For answers to questions about this report or for a list of reports published in these series, contact:

Information Dissemination Staff
National Center for Health Statistics
Centers for Disease Control and Prevention
3311 Toledo Road, Room 4551, MS P08
Hyattsville, MD 20782

Tel: 1–800–CDC–INFO (1–800–232–4636)
TTY: 1–888–232–6348
Internet: https://www.cdc.gov/nchs
Online request form: https://www.cdc.gov/info
For e-mail updates on NCHS publication releases, subscribe online at: https://www.cdc.gov/nchs/email-updates.htm.

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road, Room 4551, MS P08
Hyattsville, MD 20782–2064

For more NCHS Series Reports, visit:
https://www.cdc.gov/nchs/products/series.htm

Series 2, No. 185
CS320876