

# National Health and Nutrition Examination Survey, 2017–March 2020 Prepandemic File: Sample Design, Estimation, and Analytic Guidelines

Data Evaluation and Methods Research



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

**Copyright information**

All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.

**Suggested citation**

Akinbami LJ, Chen TC, Davy O, Ogden CL, Fink S, Clark J, et al. National Health and Nutrition Examination Survey, 2017–March 2020 prepandemic file: Sample design, estimation, and analytic guidelines. National Center for Health Statistics. *Vital Health Stat* 2(190). 2022.  
DOI: <https://dx.doi.org/10.15620/cdc:115434>

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 190

May 2022

## **National Health and Nutrition Examination Survey, 2017–March 2020 Prepandemic File: Sample Design, Estimation, and Analytic Guidelines**

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  
May 2022

## **National Center for Health Statistics**

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

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

## **Division of Health and Nutrition Examination Surveys**

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

Cynthia L. Ogden, Ph.D., M.R.P., *Acting Associate Director for Science*

# Contents

---

- Acknowledgments . . . . . v
- Abstract . . . . . 1
- Introduction . . . . . 1
- Sample Design . . . . . 2
  - First Stage of Sampling: Primary Sampling Units . . . . . 2
  - Second Stage of Sampling: Segments . . . . . 3
  - Third Stage of Sampling: Dwelling Units . . . . . 4
  - Fourth Stage of Sampling: People . . . . . 4
- Construction of 2017–March 2020 Prepandemic Public-use Data File . . . . . 4
  - Combining 2019–March 2020 and 2017–2018 PSUs . . . . . 4
  - Development of PSU Adjustment Factor . . . . . 4
- Computing Sample Weights . . . . . 4
  - Base Weights . . . . . 5
  - Nonresponse Adjustment . . . . . 5
  - Trimming . . . . . 6
  - Calibration: Creation of Initial Weights and Additional Adjustment for Urbanicity . . . . . 6
  - Final Interview and MEC Examination Weights . . . . . 7
  - Subsample Weights . . . . . 8
- Assessment of Weighting Methodology . . . . . 8
  - Response Rates . . . . . 8
  - Comparison With ACS Estimates . . . . . 8
  - Simulation to Assess the Impact of Reassigning PSUs to Different Sample Design Strata . . . . . 10
- Variance Estimation . . . . . 10
  - Variance Units for Publicly Released Data . . . . . 12
- Analytic Guidelines . . . . . 13
  - Survey Weights . . . . . 13
  - Variance . . . . . 13
  - Combining Survey Cycles . . . . . 13
  - Subsamples . . . . . 14
  - Trend Analysis . . . . . 14
  - Demographic Variables That Were Modified or Not Released . . . . . 15
  - Restricted Data Access in the NCHS Research Data Center . . . . . 15
- Summary . . . . . 16
- References . . . . . 16
- Appendix I. Supporting Tables . . . . . 18
- Appendix II. Definitions of Terms . . . . . 25

# Contents—Con.

---

## Text Figures

- 1. Comparison of National Health and Nutrition Examination Survey 2017–March 2020 prepandemic weighted distributions to 2018 5-year American Community Survey samples for urban-rural designation and primary sampling unit population size for adults aged 20 and over . . . . . 9
- 2. Comparison of National Health and Nutrition Examination Survey 2017–March 2020 prepandemic weighted distributions to 2018 5-year American Community Survey samples for selected characteristics for adults aged 20 and over. . . . . 11
- 3. Distribution of differences between estimates from simulated partial samples and their average and estimates from the full NHANES sample for diagnosed diabetes and obesity among adults aged 20 and over, 2013–2016. . . . . 12

## Text Tables

- A. Sampling domains, by Hispanic origin and race, income, sex, and age: National Health and Nutrition Examination Survey, 2015–2018 and 2019–2022 . . . . . 3
- B. Age-related variables on the publicly released data files: National Health and Nutrition Examination Survey, 2007–March 2020 . . . . . 15

## Appendix Tables

- I. Race and Hispanic-origin and income group sampling fractions used to calculate primary sampling unit measure of sizes: National Health and Nutrition Examination Surveys, 2015–2018 and 2019–2022 . . . . . 18
- II. Final sampling rates and initial base weights: National Health and Nutrition Examination Surveys, 2017–2018 and 2019–March 2020 . . . . . 19
- III. Variables used to form nonresponse adjustment cells for weighting interview samples: National Health and Nutrition Examination Survey, 2017–March 2020. . . . . 21
- IV. Variables used to form nonresponse adjustment cells for weighting mobile examination center examination samples: National Health and Nutrition Examination Survey, 2017–March 2020 . . . . . 22
- V. Most common survey sample weights and their appropriate use: National Health and Nutrition Examination Survey, 2017–March 2020. . . . . 23

# Acknowledgments

---

The authors gratefully acknowledge the assistance of Joseph Afful, Margaret Carroll, Michele Chiappa, Cheryl Fryar, Craig Hales, and Ryne Paulose-Ram in the preparation of this report.

# National Health and Nutrition Examination Survey, 2017–March 2020 Prepandemic File: Sample Design, Estimation, and Analytic Guidelines

by Lara J. Akinbami, M.D., Te-Ching Chen, Ph.D., Orlando Davy, M.P.H., Cynthia L. Ogden, Ph.D., Steven Fink, M.A., Jason Clark, M.S., Minsun K. Riddles, Ph.D., and Leyla K. Mohadjer, Ph.D.

## Abstract

### Background

The National Health and Nutrition Examination Survey (NHANES) produces national estimates that are representative of the total noninstitutionalized civilian U.S. population. The NHANES sample is selected using a complex, four-stage sample design. NHANES sample weights are used by analysts to produce estimates of the health statistics that would have been obtained if the entire sampling frame (the noninstitutionalized civilian U.S. population) had been surveyed. Sampling errors should be calculated for all survey estimates to assess their statistical reliability. Variance approximation procedures are required to provide reasonable, approximately unbiased, and design-consistent variance estimates for complex sample surveys like NHANES.

The 2017–March 2020 files represent a unique public-use data release from NHANES. The coronavirus disease 2019 (COVID-19) pandemic required suspension of data collection in March 2020. As a result, the partially completed NHANES 2019–2020 cycle was not nationally representative. Therefore, the 2019–March 2020 data

were combined with the data from the 2017–2018 cycle to create the nationally representative 2017–March 2020 prepandemic data files.

### Objective

This report describes the creation of the NHANES 2017–March 2020 prepandemic data files, including the selection of the appropriate NHANES sample design (2015–2018) to create sample weights and variance units for public-use data files. Additionally, the development of a factor applied to the primary sampling units to adjust the 2017–March 2020 data to fit the NHANES 2015–2018 sample design is described. Analyses to assess representativeness of the target population were performed, and a simulation to replicate the impact of interrupted data collection using earlier NHANES cycles was undertaken. Analytic guidance specific to use for prepandemic data files is also included.

**Keywords:** sampling • weighting adjustment • variance estimation • NHANES

## Introduction

The National Health and Nutrition Examination Survey (NHANES) is conducted by the National Center for Health Statistics (NCHS) to provide information on the health and nutritional status of the noninstitutionalized civilian population of the United States. NHANES collects person-level demographic, health, and nutrition information from personal interviews and a standardized physical examination. The high-quality, standardized procedures for examinations and specimen collection occur in specialized mobile examination centers (MECs). Moving and operating the MECs present operational challenges that limit the number of locations where data can be collected each year.

Since 1999, NHANES has continuously collected health data from across the nation, with public-use data released in 2-year cycles. Data collection for the 2019–2020 cycle was suspended in March 2020 due to safety concerns from the coronavirus disease 2019 (COVID-19) pandemic and was not rescheduled for the remaining sites in 2020. As a result, the 2019–March 2020 data were not nationally representative and would not yield meaningful stand-alone estimates. Additionally, the 2019–March 2020 data were drawn from a 2019–2022 sample design, and the sample selected for 2021–2022 data collection based on that sample design was also dropped. Therefore, the partial 2019–March 2020 data were combined with data from the previous cycle (2017–2018) for survey content that was consistent across the two cycles to create nationally representative



2017–March 2020 prepandemic data files. Combining data from a full cycle and a partial cycle means that this public-use data release covers a longer data collection period than previous releases. Most 2017–March 2020 prepandemic data files are available for downloading from the NCHS website, like for prior data releases (<https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>). For survey content that was unique to the 2019–2020 cycle, 2019–March 2020 data are available as convenience samples in the Research Data Center (RDC) (<https://wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=LimitedAccess&CycleBeginYear=2019>).

This report describes the process used to combine the 2017–2018 and 2019–March 2020 data into 2017–March 2020 prepandemic files and the construction of sample weights and variance units to produce nationally representative estimates. In addition to standard weighting procedures that account for unequal probability of selection and nonresponse within sampled areas, additional survey weight adjustments accounted for the impact of partial data collection for the 2019–2020 cycle. The analytic implications of the methods used to create the 2017–March 2020 prepandemic files are also discussed in this report. These are the first public-use files without the possibility of separate analysis for each 2-year cycle. Additional documentation of the survey content, data collection procedures, and methods for assessing nonsampling errors is provided elsewhere (<https://www.cdc.gov/nchs/nhanes.htm>).

## Sample Design

The NHANES sample represents the noninstitutionalized civilian population living in the 50 states and the District of Columbia, including people living in noninstitutional group quarters, like college or university housing or adult residential treatment facilities. Since 1999, the sample design has consisted of multiyear, stratified, clustered four-stage samples, with data released in 2-year cycles. However, the 2019–2020 data collection was interrupted in March 2020 due to the COVID-19 pandemic. Consequently, the 2019–March 2020 sample is not nationally representative, and unbiased estimates cannot be reliably produced from this partially completed cycle. To provide nationally representative estimates, the 2019–March 2020 data were combined with the 2017–2018 data to create NHANES 2017–March 2020 prepandemic data files. These data files differ from previous files in two ways: The sample spans a 3.2-year period rather than a 2-year period; and data were drawn from primary sampling units (PSUs) selected using two different sample designs, the 2017–2018 data from the 2015–2018 sample design (1) and the 2019–March 2020 data from the 2019–2022 sample design. This section describes the similarities and differences between the 4-year sample designs for 2015–2018 and 2019–2022.

Both the 2015–2018 and 2019–2022 NHANES samples were drawn in four stages: 1) PSUs (counties, groups of tracts

within counties, or combinations of adjacent counties), 2) segments within PSUs (census blocks or combinations of blocks), 3) dwelling units (DUs; households) within segments, and 4) people within households. PSUs are sampled from all U.S. counties. Screening is conducted at the DU level to identify sampled persons (SPs), based on oversampling criteria.

## First Stage of Sampling: Primary Sampling Units

The NHANES PSUs were selected with probabilities proportionate to a measure of size (MOS) (1). Each PSU's MOS was determined by previously established criteria for obtaining survey estimates for subgroups determined by age group, sex, race and Hispanic origin, and income. MOS is a weighted average of population counts, where the weights are calculated to give relatively higher probabilities of selection to PSUs with higher proportions of people within the subgroups chosen for oversampling (2). The weights, or sampling fraction, used to assign the relative contribution from each race and Hispanic-origin group in the computation of MOS are listed in Appendix Table I. For a more detailed description of the calculation of PSU MOS, see the description of the 2011–2014 NHANES sample design (2). Some PSUs have MOS large enough that they are selected with certainty (1). The remaining PSUs are referred to as noncertainty PSUs. The certainty PSUs are removed from the county frame before noncertainty PSU selection.

A major purpose for stratification for the NHANES PSU sample is to ensure that the selected PSUs are distributed as evenly as possible among the following area characteristics: overall health index, geography, urban-rural distribution, and population demographics. The NHANES 2011–2014 sample design was the first to incorporate health indicators as the initial step of the stratification scheme to form groups of states with similar levels of health indicators (2). The reason for the change was that stratifying by an outcome variable, or a variable highly correlated with an outcome variable, provides the greatest increase in precision. Both the 2015–2018 and 2019–2022 sample designs categorized all U.S. states into four health groups according to health index values (Group A states had the healthiest indices, and Group D states the least healthy indices) (1). Because the health indices change over time, some states were in different health groups in the 2019–2022 design than in the 2015–2018 design. Each state group was split into three or four major strata based on census region and the percentage of the population living in rural areas, for a total of 14 major strata. Each major stratum was split into 4 minor strata based on population demographics. One PSU was selected from each minor stratum for each 4-year sample design.

Both the NHANES 2015–2018 and 2019–2022 sample designs included oversampling of some population subgroups. The population subgroups chosen for oversampling directly determine the sampling domains used to select the sample

at all stages. The subgroups targeted for oversampling in both sample designs were:

- Hispanic people
- Non-Hispanic Black people
- Non-Hispanic, non-Black Asian people
- Non-Hispanic White people and people of other races and ethnicities at or below 185% of the federal poverty level
- Non-Hispanic White people and people of other races and ethnicities aged 0–11 years or 80 and over

The race and Hispanic-origin domains used in the sample design and weighting process differ from the categories used in variables in publicly released data files (RIDRETH3 and RIDRETH1). Race and Hispanic-origin information used for sampling and weighting is based on data obtained from the household screener to determine eligibility for inclusion in the survey. When collecting race and Hispanic-origin information at screening for the NHANES 2017–2018 and 2019–2020 data collection cycles, the Black category included people reporting non-Hispanic Black as a single race or in combination with other races, including Asian; the Asian screening category included non-Hispanic Asian people as a single race and in combination with other races, except Black. In contrast, the race and Hispanic-origin variables in publicly released data files are based on participant response to the household interview and include only single-race categories for non-Hispanic White, Black, and Asian groups. All participants reporting to belong to other or multirace groups are coded into the “other races, including multiracial” category.

Table A lists the set of 87 sampling domains in the NHANES 2015–2018 and 2019–2022 sample designs. The federal poverty levels are established 1 year before the start of each annual data collection. Each annual federal poverty level referenced remained in place for the calendar year of data

collection and was updated annually. In general, at least 4 years of data must be combined to obtain an acceptable level of reliability for most of the sampling domains given in Table A. Because data collection for 2017–March 2020 spanned fewer than 4 years, this data set may be combined with previous cycles, or some domains may need to be collapsed to produce adequate sample sizes for analysis.

The 2015–2018 and 2019–2022 sample designs did not have any methodological differences in the three sample stages within PSUs.

## Second Stage of Sampling: Segments

In the second sampling stage, each selected PSU was divided into segments that included one or more contiguous census blocks. Segments within PSUs were also sampled based on MOS (2), which were calculated similarly to the methods used for PSU MOS. The segment MOS is the sum of the MOS for each census block with the segment, and each segment must meet a minimum size to ensure it contains enough sample. Because census block data was obtained from the 2010 census, segment MOS was adjusted in some PSUs based on updated block-level housing counts to account for growth in segments due to new construction. Segment MOS was implicitly stratified by density of minority populations so that the race and Hispanic-origin distribution of the sample reflected the overall distribution

**Table A. Sampling domains, by Hispanic origin and race, income, sex, and age: National Health and Nutrition Examination Survey, 2015–2018 and 2019–2022**

Hispanic	Non-Hispanic Black	Non-Hispanic, non-Black Asian	Non-Hispanic White and other races and ethnicities <sup>1</sup>	
			Low income <sup>2</sup>	Non-low Income
All, age group (years)				
Under 1	Under 1	Under 1	Under 1	Under 1
1–2	1–2	1–2	1–2	1–2
3–5	3–5	3–5	3–5	3–5
Male, age group				
6–11	6–11	6–11	6–11	6–11
12–19	12–19	12–19	12–19	12–19
20–39	20–39	20–39	20–29	20–29
...	...	...	30–39	30–39
40–49	40–49	40–49	40–49	40–49
50–59	50–59	50–59	50–59	50–59
60 and over	60 and over	60 and over	60–69	60–69
...	...	...	70–79	70–79
...	...	...	80 and over	80 and over
Female, age group				
6–11	6–11	6–11	6–11	6–11
12–29	12–19	12–19	12–19	12–19
20–39	20–39	20–39	20–29	20–29
...	...	...	30–39	30–39
40–49	40–49	40–49	40–49	40–49
50–59	50–59	50–59	50–59	50–59
60 and over	60 and over	60 and over	60–69	60–69
...	...	...	70–79	70–79
...	...	...	80 and over	80 and over

... Category not applicable.

<sup>1</sup>Excludes non-Hispanic Black and non-Hispanic Asian people.

<sup>2</sup>People living in households at or below 185% of the federal poverty level.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2015–2022.

of the PSU. Segments were selected into the sample with probability proportional to the segment-level MOS.

### Third Stage of Sampling: Dwelling Units

A DU is the physical location where someone may live, and includes single-family homes, apartments, and noninstitutional group quarters like dormitories or shelters. All DUs within sampled segments were listed, and DUs were then sampled at rates designed to produce a national, approximately equal probability sample (2).

The exact number of DUs needed to identify the necessary number of SPs in each study location was unknown, so systematic subsamples of the entire DU screening sample were released in stages. Based on response rates from the initial subsample release, additional subsamples were released for screening as needed.

### Fourth Stage of Sampling: People

Selected DU addresses were screened to determine whether any person in the DU was eligible for selection (living in the DU as a primary residence, civilian, noninstitutionalized). For these people, race and Hispanic origin, sex, age category, and income were collected, and people within the DU were selected at rates established based on the target domain sample sizes. Sampling of people within screened households was developed to provide approximate self-weighting samples for each domain and to maximize the number of SPs per household.

## Construction of 2017–March 2020 Prepandemic Public-use Data File

### Combining 2019–March 2020 and 2017–2018 PSUs

The interruption in data collection for the 2019–2020 cycle created an extra challenge due to an adjustment made in 2019 to the order in which PSUs were visited. In the 2015–2018 sample design, data collected for each single year was nationally representative, because one PSU from each minor stratum was assigned to each year. However, the order of data collection in the PSUs in the 2019–2022 sample design was redistributed geographically to minimize travel time and maximize data collection time so that each 2-year cycle was nationally representative. That is, two PSUs from each major stratum were assigned to 2019–2020, but both PSUs could be assigned to a single year to improve operational efficiency. As a result, neither the 15 PSUs visited in 2019, nor the 18 PSUs visited in 2019–March 2020 were nationally representative or representative of any target population. To analyze the data already collected, the 2019–March 2020 data set was combined with the previous data set that was nationally representative by treating the

2019–2020 PSUs as a subset of a probability sample (or a partial probability sample) when combining it with the previous probability data set (3).

The first step in combining the NHANES 2017–2018 data and 2019–March 2020 prepandemic data was choosing a sample design that would allow for calculation of survey weights. Because populations change over time, major strata included different PSUs in the 2015–2018 design compared with the 2019–2022 design. There were two main possibilities to define a design for the combined 2017–March 2020 data: reassigning the 2017–2018 PSUs to major strata under the 2019–2022 sample design or reassigning the 2019–March 2020 PSUs to major strata under the 2015–2018 sample design. The latter option was chosen because some major strata in the 2019–2022 design would not have any sampled PSUs even after reassigning the 30 sampled 2017–2018 PSUs to the 2019–2022 sample design. The combined 2017–March 2020 prepandemic data file includes 48 PSUs: 30 from the 2017–2018 data file and 18 from the 2019–March 2020 data file, with representation of all 14 major strata from the 2015–2018 sample design.

### Development of PSU Adjustment Factor

Once the sample file was created, the weighting process needed to account for the uneven distribution of PSUs across strata. Because the strata from the different sample designs were defined differently (and therefore included different sets of PSUs), when the 2019–March 2020 PSUs were reassigned to the 2015–2018 strata, the number of PSUs per stratum in the 2017–March 2020 prepandemic sample ranged from two (in strata with two PSUs from 2017–2018 and zero PSUs from 2019–March 2020) to six (in strata with two PSUs from 2017–2018 and four PSUs from 2019–March 2020). A PSU adjustment factor was created for all 48 PSUs based on the ratio of the expected number of PSUs per stratum for a full 4-year cycle ( $n = 4$ ) to the completed PSUs per stratum ( $n = 2$  to 6). The PSU adjustment factors ranged from a high value of 2.00 to a low value of 0.67.

The PSU adjustment factors were then applied to the participant base weights as described in the next section and served to calibrate the 2017–March 2020 prepandemic data sample back to the intended design of an equal number of PSUs across strata. Note that several weighting adjustments are applied to the weights after the PSU adjustment factor, so participants with the highest weights may not necessarily have the highest PSU adjustment factors.

## Computing Sample Weights

Weighting the NHANES data produces estimates representative of the civilian, resident noninstitutionalized U.S. population. Sample weights can be considered measures of the number of people in the target population represented by each participant. Weighting accounts for

several features of the survey: the differential probabilities of selection for the sampling domains, survey nonresponse, and differences between the final sample distribution and the target population distribution. Each of the three data collection stages for NHANES (screening, interview, and examination) has a response rate. Consequently, sample weights are calculated for each of these stages.

Sample weighting is carried out in three steps. The first step involves the computation of base weights to compensate for unequal probabilities of selection within the sampling domains. The second step adjusts for nonresponse to reduce potential bias. In this step, weights are trimmed, if necessary, to reduce the impact of extreme weights on estimation. In the last step, the sample weights are calibrated to the reference population. Calibration is used to compensate for possible coverage differences from the eligible population and to reduce variances in the estimation procedure (4). The nonresponse and calibration steps are performed at each of the three data collection stages.

## Base Weights

The initial base weight for each participant within a sampling domain ( $k$ ) is the same for all other participants in that domain and equals the inverse of the sampling rate ( $r_k$ ) within the sampling domain (Table II). Although the definitions of the sampling domains are the same for the NHANES 2015–2018 and 2019–2022 sample designs, the initial base weights for each domain are different due to the changing population distributions and estimated response rates. The initial base weights were adjusted to account for: the proportion of DUs released for screening in a PSU,  $f_{i(release)}$ ; the increase in DU sample size needed in some PSUs,  $f_{i(inc)}$ ; and the factor to adjust for the incomplete sample,  $f_{i(stratum)}$ . The screening base weights are calculated as the product of the initial base weights and the three adjustment factors:

$$W_{i(base,screener)} = \frac{1}{r_k} (f_{i(release)} f_{i(inc)} f_{i(stratum)})$$

## Adjustment for number of sampled DUs released to the field

At the screening stage, not all DUs can be screened, and not all of those screened contain SPs. Consequently, a larger sample is deliberately drawn from each study location, and subsamples are released for screening as needed to obtain a relatively fixed sample size of completed examinations in each PSU. For both the NHANES 2015–2018 and 2019–2022 sample designs, the selected sample in each study location was 80% larger than was expected to be needed, but the actual sample released varied depending on the characteristics of the location. To adjust for this approach, a subsample factor was calculated for each study location as the inverse of the proportion of sampled DUs released for screening ( $R_i$ ):

$$f_{i(release)} = \frac{1}{R_i}$$

## Adjustment for increased DU sample size

Due to declining response rates and varying levels of growth and decline in different PSUs, the DU sample size was sometimes increased to ensure that enough SPs could be identified and examined in the PSU. To adjust for this approach, an increase factor was calculated for each study location as the inverse of the percentage increase in the DU sample ( $I_i$ ):

$$f_{i(inc)} = \frac{1}{I_i}$$

The increase in DU sample size could effectively be canceled out by the amount released. For example, if the DU sample size was increased to 125% ( $I_i = 1.25$ ), and then 80% of the sample was released ( $R_i = 0.8$ ), then the resulting combined factor  $f_{i(release)} \cdot f_{i(inc)} = 1$ .

## Adjustment for incomplete sample

As described in the section “Development of PSU Adjustment Factor,” a weighting adjustment was needed to account for the PSUs not completed in the 2019–2020 sample. The PSUs with data collected during 2019–March 2020 were reassigned to the strata from the NHANES 2015–2018 sample design. In a full 4-year sample, data would be collected from four PSUs in each stratum. Instead, the number of PSUs with data collected in each stratum ranged from two to six. To adjust for this imbalance, a PSU adjustment factor was applied to all PSUs with data collected during 2017–2018 and 2019–March 2020:

$$f_{i(stratum)} = \frac{4}{\text{Number of PSUs in stratum}}$$

The interview base weights were set to the screening final weights, which were the product of the screening base weights and weighting adjustment factors. Similarly, the MEC examination base weights were set to the interview final weights, which were the product of the interview base weights and interview weighting adjustment factors.

## Nonresponse Adjustment

If every selected household was screened and every SP agreed to complete the interview and the examination, then weighted estimates using screening base weights would produce approximately unbiased estimates of characteristics of the civilian noninstitutionalized U.S. population. However, some of the selected households were not screened, some of the SPs who were screened chose not to be interviewed, and some of the interviewed participants did not participate in the examination. To reduce the potential for nonresponse bias, base weights were adjusted for nonresponse at each stage of the survey (screening, interview, and examination).

The amount of information that can be used for these adjustments increases at each progressive stage; only the sampling information is available at the screening stage, while person-specific information from the interview is available to adjust MEC examination weights (Tables III and IV).

The nonresponse adjustment procedure consists of computing adjustment factors,

$$f_{i(NR)} = \frac{\text{Sum of stage base weights in the adjustment cell}}{\text{Sum of stage base weights of the participants in the adjustment cell}}$$

and applying these to the survey weights as:

$$w_{i(NR,stage)} = w_{i(base,stage)} f_{i(NR,stage)}$$

separately within nonresponse cells, where nonresponse cells are defined by categorical characteristics known for both participants and nonrespondents. Because little is known about households that do not complete the screener, the adjustment cells at the screening stage are just the segments, with the assumption that DUs in the same segment are similar. For the interview and examination nonresponse adjustments, a classification program is used to identify available variables most highly related to response propensity. Different variables are identified to form adjustment cells for the following age groups: 0–5, 6–19, 20–39, 40–59, and 60 years and over.

The use of these propensity-related variables differs by survey collection period. Variables used to form the nonresponse adjustment cells for interview weights are listed in Table III, and for MEC examination weights, in Table IV. Nonresponse adjustment reduces bias if response rates and survey characteristics vary from cell to cell, and if participants and nonrespondents sharing the same characteristics are in the same cell. An effect of nonresponse adjustment is that it increases the variability of the weights, which then increases the variance of estimates obtained from the data. When the nonresponse adjustment cells contain enough cases and the adjustment factors are not too large, the effect on variances is modest. A large adjustment factor in a cell is usually the result of a small number of participants in that cell. To avoid having nonresponse adjustments based on very small sample sizes, or having large nonresponse adjustment factors, nonresponse adjustment cells can be collapsed to form larger cells. Due to increasing nonresponse to the screener and interview, larger adjustment factors than in previous NHANES samples were allowed, which could result in higher variance for some survey estimates. The nonresponse adjustment factors for most cells were less than 2. The largest factors were 2.00 at the screening stage, 2.77 at the interview stage, and 1.44 at the MEC examination stage.

## Trimming

Weight-trimming procedures are used to reduce the impact of any extreme weights on estimation. Even a few unexpectedly large weights can markedly inflate the variance of survey estimates. However, trimming sample weights may introduce estimation bias (5), so trimming is not automatically used for all sample weights.

Nonresponse adjustment can contribute to extreme weights. To determine whether to trim weights for samples (or subsets of samples), the distribution of weights within each sampling domain was inspected. The threshold for the NHANES 2017–March 2020 prepandemic sample interview weights was defined as 4.75 times the sampling domain mean. Four weights exceeded this threshold. The values of these extreme weights were reduced to the threshold, and the weights of all cases in the same sampling domain were adjusted so that the sum of the weights in each sampling domain equaled the corresponding weighted sum before trimming. The threshold for the NHANES 2017–March 2020 prepandemic sample MEC examination weights was defined at five times the sampling domain mean, and 13 weights exceeded this threshold and were trimmed.

Trimmed sample weights were calculated as follows. Let  $t_i$  be the weight after trimming for  $SP_i$ , defined as:

$$t_i = \begin{cases} w_{i(NR)}, & \text{if } w_{i(NR)} \leq \text{threshold} \\ \text{threshold}, & \text{otherwise} \end{cases}$$

Then the trimming factor  $f_{i(TR)}$  was calculated as:

$$f_{i(TR)} = \frac{t_i}{w_{i(NR)}} \cdot \frac{\sum_{i=1}^{n_k} w_{i(NR)}}{\sum_{i=1}^{n_k} t_i}$$

where  $n_k$  is the sample size of the  $k$ th race–Hispanic-origin–income–sex–age sampling domain, and

$$w_{i(TR,stage)} = w_{i(NR,stage)} f_{i(TR,stage)}$$

## Calibration: Creation of Initial Weights and Additional Adjustment for Urbanicity

The final step in the weighting procedure for each survey stage is calibration to known population totals. Calibration can be done iteratively to multiple sets of population totals in a process called raking, which was done for the NHANES 2017–March 2020 prepandemic weights. Calibration compensates for undercoverage of certain demographic groups and for any residual differential nonresponse among these groups. Like nonresponse adjustment, calibration is done at the screening, interview, and examination stages.

As mentioned previously, a participant’s sample weight represents the number of people with similar characteristics in the target population. The sum of all participants’ weights in a demographic subgroup could be considered as the total

number of people that NHANES participants represent for this subgroup. Calibration adjusts the individual sample weights so that the sum of the sample weights within a demographic subgroup equals the population from an independent data source for that subgroup.

Similar to the adjustment factors for earlier steps, calibration involves applying a ratio adjustment to the survey weights. In this step, the denominator of the adjustment factor for a particular demographic subgroup is the sum of the nonresponse-adjusted and trimmed sample weights within the demographic subgroup, and the numerator  $N_C$  is the reference population control total for the demographic subgroup:

$$f_{i(C)} = \frac{N_C}{\text{Sum of nonresponse adjusted and trimmed weights of the demographic subgroup}}$$

The calibrated weights are then calculated as:

$$W_{i(C,stage)} = W_{i(TR,stage)} f_{i(C,stage)}$$

NHANES 2017–March 2020 prepandemic weights were initially raked following the same process as NHANES 2017–2018 (6), which will be referred to as W1. However, a review of the NHANES 2017–March 2020 prepandemic weights showed a need for additional calibration to urban-rural status. Urbanicity is known to be correlated with health (7–9), so this additional adjustment was applied to reduce the risk of bias in the survey estimates due to undercoverage. The revised weight will be referred to as W2 and is the final weight for each stage. A comparison of estimates using W1-adjusted NHANES sample weights and W2-adjusted weights with estimates from the American Community Survey (ACS) is shown in “Comparison With ACS Estimates.”

Race, Hispanic origin, age, and sex are collected from all SPs in the screener. Area-level, mean household income can be obtained for the area (census tract) where each SP lives, and urbanicity can be obtained for the county where each SP lives, so the 2017–March 2020 prepandemic screening weights were calibrated (using three-dimensional raking) to race–Hispanic-origin–age–sex demographic subgroups, area-level household income, and urbanicity. Because highest education level for an SP is collected in the household interview, this information is only known for interview participants. The 2017–March 2020 prepandemic interview and MEC examination weights were calibrated (using four-dimensional raking) to race–Hispanic-origin–age–sex demographic subgroups, race–Hispanic-origin–sex–education-level subgroups (for ages 20 years and over), area-level household income, and urbanicity.

For the NHANES 2017–March 2020 prepandemic weights, the 2018 1-year ACS was used as the source for the reference totals for each race–Hispanic-origin–age–sex subgroup and

each race–Hispanic-origin–sex–education-level subgroup. The 2018 5-year ACS was used to obtain the mean household income for each census tract, and the tracts were then divided into deciles. The first decile contained the 10% of tracts with the lowest mean household incomes, and the 10th decile contained the 10% of tracts with the highest mean household incomes. The 2018 5-year ACS was then used as the source for the reference totals for each income decile. The 2013 NCHS Urban–Rural Classification Scheme for Counties was used to determine the urban-rural status of each county in the United States (10). Four categories were used, including large central metro, large fringe metro, medium and small metro, and micropolitan and noncore. The 2018 5-year ACS was then used as the source for the reference totals for each category.

The ACS population counts were adjusted to match best estimates from the U.S. Census Bureau of the total noninstitutionalized civilian population of the United States, including people not counted in surveys or in the most recent decennial census. Calibration using ACS, consequently, brings the weighted totals up to the level of the presumed total noninstitutionalized civilian population in the United States. A detailed report on the design and methodology of ACS is available from the Census Bureau’s website (11).

A major effect of calibration is that it implicitly imputes survey characteristics for people who were missed by the survey due to errors in the sampling frame and adjusts for residual nonresponse. The underlying assumption for calibration is that missed people not covered by the survey have the same distribution of characteristics as surveyed people within the calibration cells. This assumption is an oversimplification; the missed people are likely to be different. However, in the absence of information on the characteristics of the missed people, calibration is a technique available for reducing bias due to undercoverage and residual nonresponse (4).

## Final Interview and MEC Examination Weights

The final sample weight for each participant at each stage is calculated as the product of the base weight, the nonresponse adjustment, the trimming adjustment (if needed), and the calibration adjustment. That is:

$$W_{i,stage} = W_{i(base,stage)} f_{i(NR,stage)} f_{i(TR,stage)} f_{i(C,stage)}$$

The final screening weight was calculated as:

$$W_{i,screener} = W_{i(base,screener)} f_{i(NR,screener)} f_{i(TR,screener)} f_{i(C,screener)}$$

The final weights from the screening stage are the base weights for the interview stage, and the final interview weight was calculated as:

$$\begin{aligned} w_{i, \text{interview}} &= w_{i(\text{base, interview})} f_{i(\text{NR, interview})} f_{i(\text{TR, interview})} f_{i(\text{C, interview})} \\ &= w_{i(\text{base, screener})} f_{i(\text{NR, screener})} f_{i(\text{TR, screener})} f_{i(\text{C, screener})} \\ &\quad f_{i(\text{NR, interview})} f_{i(\text{TR, interview})} f_{i(\text{C, interview})} \end{aligned}$$

The final weights from the interview stage are the base weights for the MEC examination stage, and the final MEC examination weight was calculated as:

$$\begin{aligned} w_{i, \text{MEC}} &= w_{i(\text{base, MEC})} f_{i(\text{NR, MEC})} f_{i(\text{TR, MEC})} f_{i(\text{C, MEC})} \\ &= w_{i(\text{base, screener})} f_{i(\text{NR, screener})} f_{i(\text{TR, screener})} f_{i(\text{C, screener})} \\ &\quad f_{i(\text{NR, interview})} f_{i(\text{TR, interview})} f_{i(\text{C, interview})} \\ &\quad f_{i(\text{NR, MEC})} f_{i(\text{TR, MEC})} f_{i(\text{C, MEC})} \end{aligned}$$

The interview weight should be used for analyses using household interview data only when no variables from the examination are included. The MEC examination weights should be used for analyses that include examination data (including the MEC interview and some laboratory data). Additionally, special survey components and subsamples required further adjustment of the MEC examination weights due to specific inclusion criteria (morning fasting sample). Component-specific weights, if needed, are released with the data for the component and described in the component’s documentation. Although the NHANES 2017–March 2020 prepandemic public-use data file contains more participants than previous 2-year cycles (so the weights are smaller than the weights for the 2017–2018 NHANES on average), the weights have wide ranges. Analysts should be aware of the potential influence of large weights, especially when extreme weights are associated with extreme data values. Large weights may also inflate variance.

## Subsample Weights

Some laboratory and examination components are performed for a subsample of NHANES participants. For example, some, but not all, NHANES 2017–March 2020 prepandemic participants aged 12 years and over were selected to fast for 8–23 hours and participate in a MEC examination the following morning. The subsamples selected for these components were chosen randomly with a specified sampling fraction (one-half of the examined age group) and according to the protocol for that component. Each subsample is selected to be a nationally representative sample of the target population and has its own designated sample weight that accounts for the additional probability of selection into the subsample component and any additional nonresponse to the component.

Subsample weights are included in the respective component data files. Because these weights differ from the MEC examination weights, subsample weights must be used

for statistical estimation of measures collected only in that sample and for analyses that include those measures (see Table V for a list of special component sample weights and information regarding their appropriate use.) See the survey protocol and documentation (<https://www.cdc.gov/nchs/nhanes.htm>) for more detail on specific laboratory tests and health measurements completed for a subsample of participants.

## Assessment of Weighting Methodology

The COVID-19 pandemic that interrupted data collection required adapting previous weighting procedures to produce a file that could be used for nationally representative estimates for 2017–March 2020 (see “Development of PSU Adjustment Factor”). These adaptations were based on methods used previously in other settings and applications (3). Additionally, falling response rates in NHANES, as seen across national surveys, have required more in-depth and extensive analyses to assess the potential effects of nonresponse bias on outcome statistics (6,12,13). The enhanced weighting measures incorporated for the 2017–2018 NHANES data to address nonresponse and location sampling variability were retained for the 2017–March 2020 data (6). To assess these weighting adjustments, the NHANES 2017–March 2020 prepandemic estimates for demographic characteristics were compared with estimates from ACS, the largest representative survey of the target population.

## Response Rates

Sample sizes and response rates for all 2-year cycles and the overall 2017–March 2020 prepandemic data set are available from the NHANES website (13). Response rates were calculated for each stage of the survey: screening of DUs to identify SPs; household interviewing of SPs; and examination of interviewed SPs. In the 2017–March 2020 prepandemic data set, the household screener response rate was 88.7%. From the responding households, 27,066 SPs were selected from 48 different PSUs. Of those selected, 15,560 participants completed the interview, and 14,300 were examined. After adjustment for nonresponse to the screener, the final interview response rate was 51.0%, and the final examination response rate was 46.9%.

## Comparison With ACS Estimates

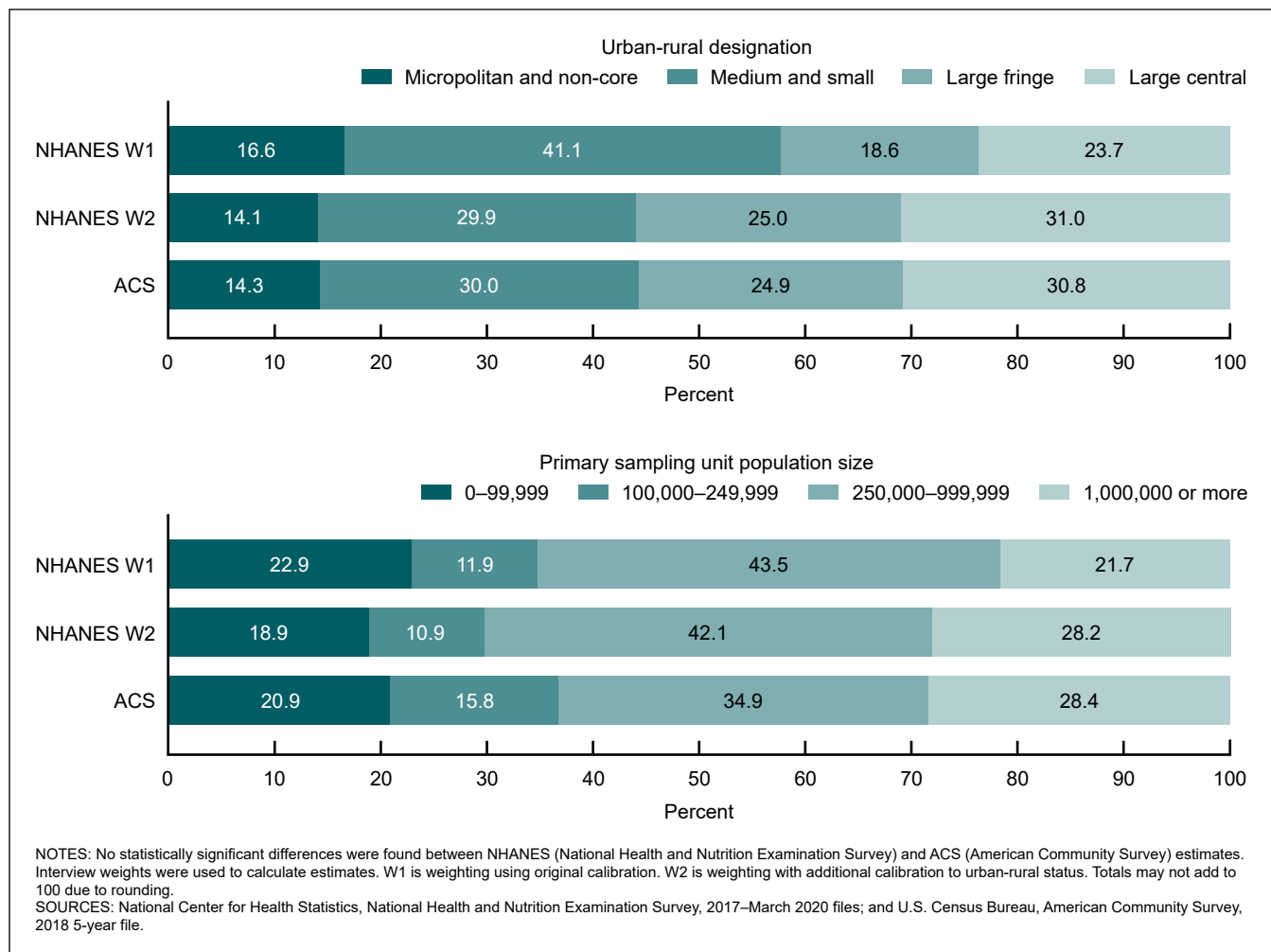
ACS is an ongoing survey conducted by the U.S. Census Bureau that provides detailed population and housing information about the United States and represents the largest nationally representative household survey conducted, with 3.5 million completed interviews per year (<https://www.census.gov/programs-surveys/acs/>). 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, people in the military or living in institutional group quarters were excluded from the ACS estimates for the comparisons in this report. The 2018 5-year ACS data were the most recent available at the time of the evaluation. Because both NHANES and ACS are subject to error, variances of all estimates were considered when making comparisons and performing significance testing. Two-sided *t* tests were used to evaluate statistical significance, and the standard error was calculated using the formula for the standard error of a difference for two independent samples for all ACS comparisons except urbanicity. Counties were categorized into four categories based on 2013 NCHS urban-rural codes. ACS distributions were weighted based on county noninstitutionalized population counts, and NHANES distributions were weighted based on SP weights. Goodness-of-fit tests were used to test the difference between ACS and NHANES urban-rural distributions.

The set of calibrations developed for the NHANES 2017–2018 weighting process (6)—race–Hispanic–origin–

age–sex demographic subgroups, race–Hispanic–origin–sex–education–level subgroups, and area-level income deciles—were used in the calibration step for the initial 2017–March 2020 sample weights, designated as W1. However, it was noted that the distribution of the W1-weighted NHANES sample across urban-rural designations did not match the ACS distribution. Although the differences were not statistically significant, they were large enough to warrant an additional calibration dimension. Additional calibration to urban-rural designations was undertaken to recalculate the weights and produce W2. This additional calibration of the survey weights resulted in the weighted NHANES 2017–March 2020 prepandemic sample distribution matching well to the 2018 5-year ACS estimates for urban-rural designation (Figure 1). Consequently, W2 were considered the final survey weights. However, some differences remained in the distributions among categories of PSU population size between the W2-weighted NHANES estimates (calculated using interview weights) and ACS estimates (Figure 1).

**Figure 1. Comparison of National Health and Nutrition Examination Survey 2017–March 2020 prepandemic weighted distributions to 2018 5-year American Community Survey samples for urban-rural designation and primary sampling unit population size for adults aged 20 and over**





W1- and W2-weighted NHANES estimates (interview weights) for education level, marital status, health insurance coverage, and income levels (categories according to family-income-to-poverty ratio) were compared with ACS estimates (Figure 2). Little change was observed between W1- and W2-weighted NHANES estimates for the characteristics examined. For education and marital status, no statistically significant differences were observed in the percentages of adults in each category between the ACS sample and the W1- and W2-weighted NHANES samples. However, differences were seen between NHANES and ACS for health insurance coverage and income status. A significantly higher percentage of adults reported having no health insurance for both the W1- and W2-weighted NHANES estimates compared with ACS estimates. A higher percentage of adults were classified as low income (family-income-to-poverty ratio less than or equal to 185%) for W1- and W2-weighted NHANES estimates compared with ACS estimates after excluding missing responses (12.3% for NHANES, 1.2% for ACS). The differences between 2017–March 2020 NHANES and ACS income estimates were similar or smaller than for previous cycles (6). Some observed differences between NHANES and ACS estimates may come from differences in survey administration: ACS is a multimode survey and NHANES is conducted in person.

## Simulation to Assess the Impact of Reassigning PSUs to Different Sample Design Strata

To evaluate the effects of the PSU adjustment factors on the NHANES estimates, a simulation was conducted using data from previous cycles. The 2017–March 2020 NHANES prepandemic data were created by combining the 30 PSUs in 2017–2018 (from the 2015–2018 design) with the 18 completed PSUs in 2019–2020 (from the 2019–2022 design). To approximate these conditions, 300 independent samples of 18 PSUs were selected randomly from among the 30 PSUs from the 2015–2016 cycle (from the 2015–2018 design), and each of the 300 samples was combined with the 30 PSUs from the 2013–2014 cycle (from the 2011–2014 design).

There were three key differences between the 2011–2014 and 2015–2018 designs: 1) the state of California was a separate health state group in 2011–2014 but not in 2015–2018; 2) the low-income sampling criteria was changed from a family-income-to-poverty ratio of at or below 130% to at or below 185%; and 3) 13 major strata with two certainty PSUs were included for each annual sample in the 2011–2014 design, but 14 major strata with one certainty PSU were included for each annual sample in the 2015–2018 design (1).

In each of the 300 iterations in the simulation, the random sample of 18 PSUs from 2015–2016 was combined with the 30 PSUs from the 2013–2014 cycle by reassigning them to major strata under the 2011–2014 sample design and applying PSU adjustment factors created through the same

methodology used for the 2017–March 2020 prepandemic files. The PSU adjustment factors ranged from a high of 2.00 to a low of 0.80, but the range varied in each iteration (compared with a range of 2.00 to 0.67 for the 2017–March 2020 files). A classification tree algorithm was used to adjust sample weights for nonresponse for the full 2013–2016 data set (with 60 PSUs) and each of the 300 partial data sets (with 48 PSUs). For each health outcome, the difference for each simulation and the true estimate was calculated to establish a distribution of the differences across the 300 simulations.

The first panel in Figure 3 presents the distribution of the 300 differences (one from each iteration) for the prevalence of diagnosed diabetes in adults, and the second panel for the prevalence of obesity in adults. The gray dashed lines represent the averages of each set of the 300 differences. Differences in estimates between the full 4-year sample and partial sample were small for overall estimates on average. For diagnosed diabetes in adults aged 20 and over, the mean difference among the 300 simulations and the actual NHANES 2013–2016 estimates was -0.13 percentage points. The mean difference between the simulations and the actual obesity prevalence estimates was also -0.13 percentage points.

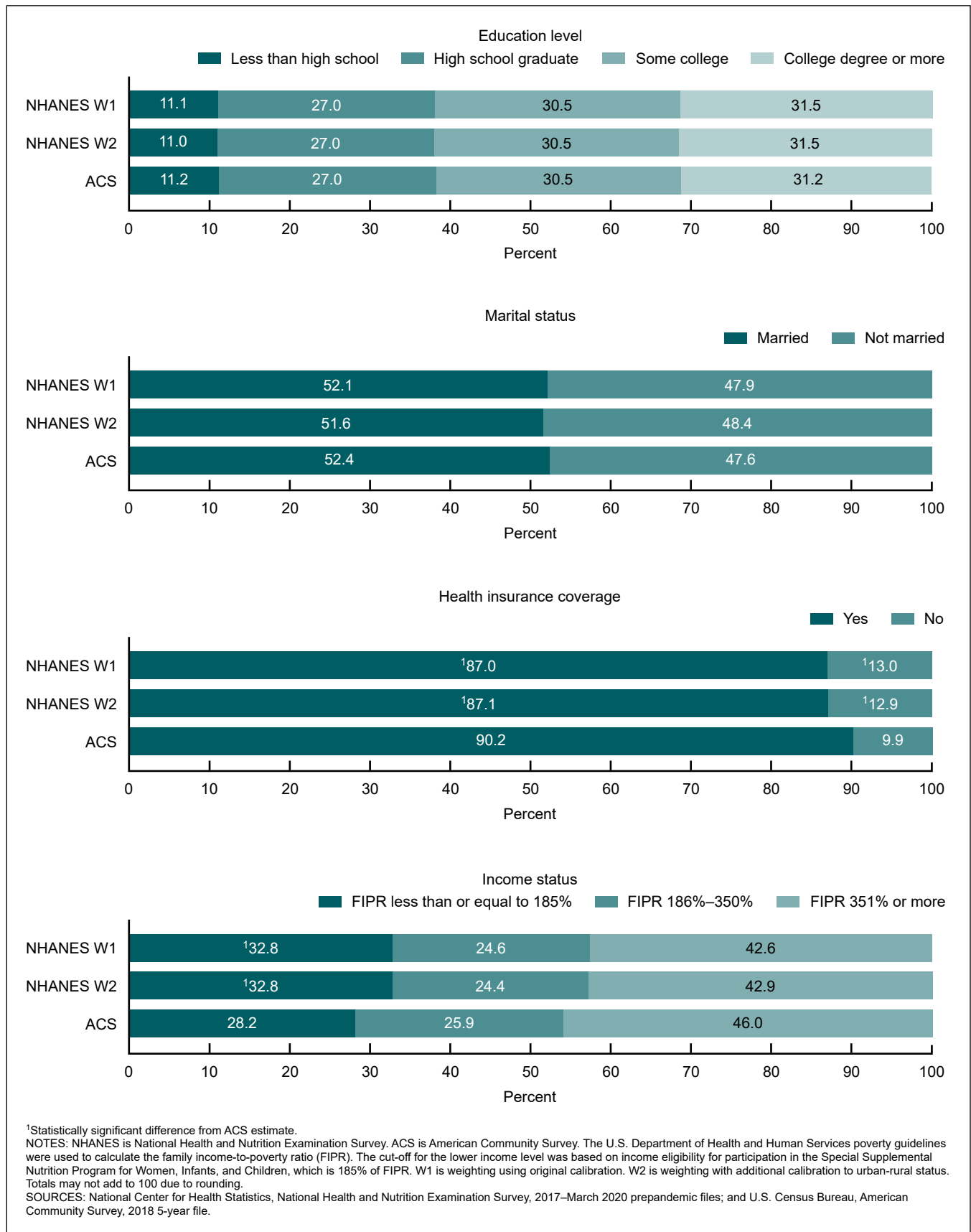
However, note that the averages from the simulation are based on various combinations of PSUs, including some where the 2015–2016 sample of 18 PSUs more closely resembles national representation than the data collected during 2019–2020. The small differences in Figure 3 should not imply that the 2017–March 2020 prepandemic data would be expected to provide estimates that would have been obtained if data had been collected from the entire 2019–2020 sample. That is, the simulation was done with 2013–2016 data, for which differences in PSUs between completed cycles was known. Because the 2019–2020 cycle was not completed, it is not possible to determine how much a completed 2019–2020 cycle would have differed from the full 4-year estimate. However, the simulation study suggests that the chance of 2017–March 2020 prepandemic data producing estimates that are significantly different from the counterfactual full 4-year estimates is low.

## Variance Estimation

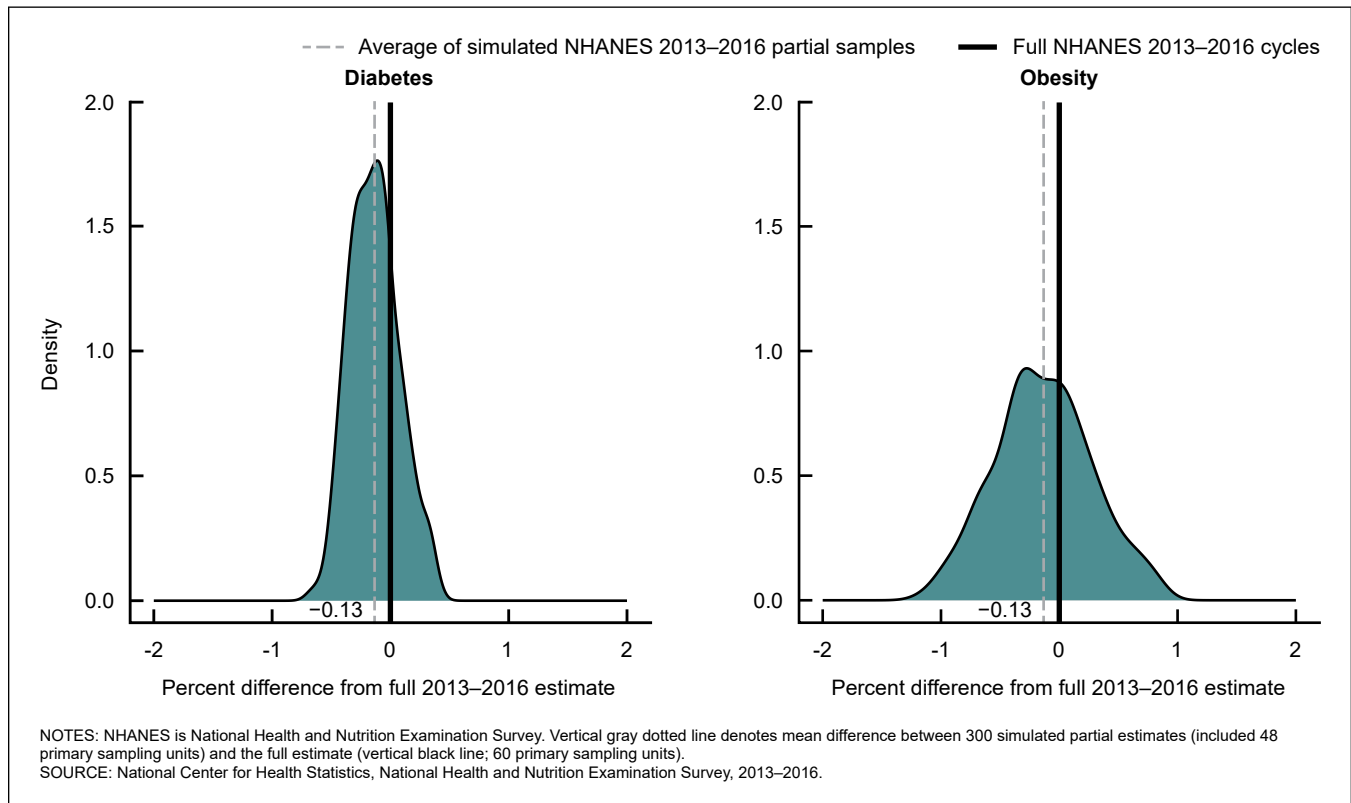
This section discusses design-based methods of variance estimation for complex sample survey data and describes the creation of variables needed for variance estimation on the public-use data files for the NHANES 2017–March 2020 prepandemic data. Sampling errors should be calculated for all survey estimates to help determine the statistical reliability of those estimates.

For complex sample surveys, exact mathematical formulas for variance estimates are not available. Variance approximation procedures are necessary to provide reasonable, approximately unbiased, and design-consistent estimates of variance. These routines require special programs that

**Figure 2. Comparison of National Health and Nutrition Examination Survey 2017–March 2020 prepandemic weighted distributions to 2018 5-year American Community Survey samples for selected characteristics for adults aged 20 and over**



**Figure 3. Distribution of differences between estimates from simulated partial samples and their average and estimates from the full NHANES sample for diagnosed diabetes and obesity among adults aged 20 and over, 2013–2016**



account for the survey design. Standard statistical software routines that assume a simple random sample should not be used for computing variances for NHANES estimates.

Although the NHANES 2017–March 2020 prepandemic sample is nationally representative, it contains data from fewer PSUs than 4-year data sets made by combining 2-year cycles (48 compared with 60). This smaller number of PSUs may create challenges for variance estimation. Sample sizes for some specific race–Hispanic-origin–income–sex–age domains may be small. Additionally, direct design-based variance estimates may be unstable for some measures.

Two variance approximation procedures that account for the complex sample design and allow for the computation of design effects are replication methods and Taylor series linearization. Currently, NCHS recommends using the Taylor series linearization method for variance estimation in analyses of NHANES data for 2-year cycles or combined cycles (including the NHANES 2017–March 2020 prepandemic public-use data files). SUDAAN, Stata, R, and SAS survey procedures can be used to obtain variance estimates using this method.

### Variance Units for Publicly Released Data

Noncertainty PSUs are grouped into major strata defined by state-level health-related variables. In any 2-year sample, two noncertainty PSUs are sampled from each major stratum. These strata are used as variance strata to estimate sampling error in the Taylor series linearization method. In the 2017–March 2020 prepandemic public-use data files, each stratum had two to six noncertainty PSUs sampled during 2017–March 2020. Additional variance strata were created for the PSUs with two or three sampled PSUs, generally defined as variance units, in each stratum. Certainty PSUs are not selected within strata. Variance strata for these PSUs are formed based on the relative size of the PSU compared with the other PSUs. Consequently, some of these variance strata may have one PSU split into multiple variance units, while other variance strata may comprise three PSUs for variance estimation, depending on the number and size of the certainty locations that year.

Risks of data disclosure that may compromise participants’ confidentiality include the small number of PSUs in the sample, geographic data, and other area characteristics in the data files. As a result, masked variance units (MVUs) are provided for use with the public-use data files to reduce the chance of matching PSUs in the sample to PSUs in the geographic areas, while minimizing the bias in the variance

caused by altering the PSU structure. Collectively, the MVUs formed by the noncertainty and certainty locations can be used for variance estimations and to estimate sampling error. Though they are not the “true” design PSUs, MVUs produce variance estimates that closely approximate the variances that would have been estimated using the true design PSUs.

MVUs have been created for all 2-year survey cycles from NHANES 1999–2000 through 2017–2018, and have also been created for the NHANES 2017–March 2020 prepandemic public-use data files. The MVUs can be used for analyzing any 2-year cycle data set or any combined cycles data set. Analysts can compute replicate weights for variance estimation based on MVUs.

## Analytic Guidelines

Although a thorough evaluation was conducted to assess the reliability of national estimates produced using the 2017–March 2020 prepandemic data files, some considerations should be made before performing data analyses. Previously published guidelines (14), including accounting for the complex survey design for variance estimation, choosing appropriate survey weights for analysis of subsamples, and handling missing data, still apply. Analysts should also consult the NCHS guidelines for presentation of proportions for guidance on statistical reliability of prevalence estimates (15). Data file documentation is available through the link next to the data file on the NHANES website and is always the most current source of information about the variables in each data file (16). Additional considerations specific to the 2017–March 2020 prepandemic file are discussed in the next section.

## Survey Weights

The sample weights that appear on the data file should be used to calculate estimates using the combined 2017–March 2020 prepandemic data. The weights are designed to produce nationally representative estimates for the entire period covered by the 2017–March 2020 prepandemic files. It is not possible to calculate separate weights for the 2019–March 2020 sample given that these data do not follow the 2019–2022 sample design due to an interruption in data collection before the 2-year cycle was completed.

## Variance

NHANES is designed to produce reliable health statistics for many subdomains of the general population because health characteristics can vary by age, race and Hispanic origin, sex, income status, or geographical location. To achieve sufficient sample size in these subdomains, some are oversampled at a high rate relative to the sampling rate for the more populous domains. When subdomains are combined for

analysis, a wide range of weights may occur, which will lead to increased variance in the analytic results. Analysts should be aware of the range of weights within the subgroup being analyzed and the resulting potential increase in variance.

The variances of estimates from the 2017–March 2020 prepandemic file are generally smaller compared with the 2017–2018 files due to increased sample size achieved by combining data cycles. However, for some estimates and some demographic subgroups, the 2017–March 2020 files may produce larger than expected variance estimates due to the increased variation in the sampling weights and the increased variation in underlying variables.

Compared with previous 4-year cycles, the 2017–March 2020 estimates may have increased variance due to a relatively smaller sample size (27,066 SPs) and fewer PSUs (48 study locations). Additionally, the factors used to adjust those 48 study locations to represent the whole nation can lead to increased variance (17). Analysts should be aware that differences in characteristics between PSUs can cause high variances for specific analytic variables of interest.

## Combining Survey Cycles

Since 1999, NHANES data have been released in 2-year cycles. The 2017–March 2020 prepandemic data represent a 3.2-year period. The NHANES sample design makes it possible to combine two or more cycles to increase the sample size, including combining the 2017–March 2020 prepandemic data with previous cycles. When combining data cycles, it is important to:

- Be aware of sample design changes between cycles.
- Use the file documentation to verify that data items collected in the cycles that are to be combined are comparable in question wording, methods, and inclusion or exclusion criteria.
- Examine the validity of the assumption that no trend in the estimate exists over the period being combined.

Previous analytic guidance states that combining cycles into samples covering at least 4 years is recommended to ensure sufficient sample size for analysis of subsamples or outcomes with low prevalence. The 2017–March 2020 prepandemic data file covers 3.2 years compared with 2 years for other NHANES data files. Because the period differs from earlier cycles, the survey weights should be adjusted when 2017–March 2020 data files are combined with other 2-year cycles to reflect the longer period and larger population represented by the 2017–March 2020 files. New multicycle sample weights can be calculated based on the sample weights of the combined survey cycles with the following formulas:

Combining two survey cycles, 2015–2016 and 2017–March 2020 (5.2 years):

If  $SDDSRYR = 9$  then  $MEC52Y = (2/5.2) \cdot WTMEC2YR$ ;

If SDDSRVYR = 66 then MEC52Y = (3.2/5.2) • WTMECPRP.

Combining three survey cycles, 2013–2014, 2015–2016, and 2017–March 2020 (7.2 years):

If SDDSRVYR in (8,9) then MEC72Y = (2/7.2) • WTMEC2YR;

If SDDSRVYR = 66 then MEC72Y = (3.2/7.2) • WTMECPRP.

The sum of the combined multiyear sample weights should be reasonably close to an independent estimate of the midpoint population. The guidance for combining cycles also applies to subsamples.

## Subsamples

Combining data sets with subsample weights requires several considerations. Some subsamples, like environmental subsamples, are mutually exclusive (no overlapping participants) and cannot be combined in the same data release cycle or across cycles. Other subsamples do have some overlap and may be combined; for example, each environmental subsample has overlap with the fasting subsample. These could be combined within the same data release cycle. These could also be combined across cycles if the overlapping sample size is adequate, and the methods of data collection and definitions of measures are the same across cycles. However, NHANES does not provide sample weights for these types of combined data sets.

To combine two or more subsamples for analysis, random overlap is needed between the subsamples, and appropriate weights need to be recalculated. NCHS does not provide specific recommendations about how to create combined sample weights for overlapping subsamples, but calibration approaches to adjusting weights may be found within the SUDAAN software with the procedures WTADJUST and WTADJX (18,19). Adjustments to create sample weights for overlapping subsamples can be made by adjusting the sample weights to match population totals within adjustment cells defined by race and Hispanic origin, sex, and age group, demographic characteristics that are present in both subsamples. Additional adjustments specific to the analysis may also be made to the sample weights using characteristics common to both subsamples. The selection and categorization of the variables used for adjustment will depend on the size of the combined sample and the purpose of the analysis. Because the overlapping subsamples are typically small, coarser adjustment cells than those used for the original sample weight creation are often required.

Note that sample weight adjustments using the public-use data files would require one of the released race and Hispanic-origin variables (RIDRETH3 or RIDRETH1) available in the file for the calculations. As described previously, sample weights created by NCHS use the sampling categorization of race and Hispanic origin, which differs from the public-use data file categorization.

Because all NHANES participants are eligible for the MEC 24-hour dietary recall interview, analysts should treat this component like any other full sample (interviewed or examined sample) when combining with a subsample, and use the sample weights of the smallest subpopulation that includes all the analytic variables, in this case, the subsample weights. However, analysts should consider recalculating the subsample weight to account for the day of the week of reported consumption, which was considered in the calculation of the 24-hour dietary weights but is not reflected in the subsample weight.

## Trend Analysis

Because no national estimates can be made from the 2019–March 2020 data, comparisons or examination of trends between 2017–2018 and 2019–March 2020 data are not possible. To prevent data analysts from separating the two cycles, changes have been made to respondent sequence identification numbers, and PSU and strata variables have been masked.

When analyzing trends that include 2017–March 2020 data and data from previous 2-year cycles, analysts should account for unequally spaced intervals. For example, if observed time points are 2011–2012, 2013–2014, 2015–2016, and 2017–March 2020, then the interval midpoints for each of these cycles (2012, 2014, 2016, and 2018.6) could be used to represent time points in a trend model (5).

PSU-level adjustments to survey weights were designed for overall estimates and not specific subgroups. Any trend comparisons for subgroups (for example, by age, sex, race and Hispanic origin) between the 2017–March 2020 prepandemic file and previous NHANES cycles should be interpreted with caution. The magnitude and direction of the differences between two cycles may vary by subgroup. This could cause changes in the magnitude or the direction of the trend within a certain subgroup, relative to overall changes. When conducting analyses and interpreting results, analysts should consider the historical context of the trends in addition to the methodological approach to create the 2017–March 2020 prepandemic file.

Although the sample designs for NHANES 2015–2018 and NHANES 2019–2022 were generally the same (including the same definitions of sampling domains by race and Hispanic origin, income, age, and sex, and the same stratification scheme for noncertainty PSUs), the details of these designs differed (including different sampling rates for each domain and different groupings of PSUs by stratum due to updated response rate estimates, population estimates, and demographic and health characteristics). Because of these differences (and more significant changes in earlier sample designs), data users should be aware of potential inefficiencies when combining older data (1999–2006, 2007–2010, 2011–2014) with 2015–2016 and 2017–March 2020 data.

## Demographic Variables That Were Modified or Not Released

Because some demographic information, in combination with other information, can create disclosure risks, the following variables were modified or not included in the 2017–March 2020 prepandemic public-use data file release.

- **Age:** As for 2011–2016 cycles, the 2017–March 2020 demographic file includes the variable age in years at screening (RIDAGEYR) for all participants. Age at examination and age in months for children may be useful for some analyses. However, because exact age, in combination with other information, can create disclosure risks, this variable was not included in the 2017–March 2020 release (Table B).
- **Education:** Education level for children and youth aged 6–19 years was not included in the 2017–March 2020 release. Education level for adults aged 20 and over is included as for previous cycles.
- **Income:** Family and household income is not included in the 2017–March 2020 release. The ratio of family income to the federal poverty level is included as for previous cycles.
- **Marital status:** Participants who reported being divorced or separated were combined into a single category; this information was released as separate categories in previous data cycles.
- **Military service:** This information is not included in the 2017–March 2020 data release.

## Restricted Data Access in the NCHS Research Data Center

In addition to the many NHANES 2017–March 2020 prepandemic public-use data files available for downloading on the NCHS website, special data sets for this cycle with restricted access are available only through the NCHS RDC, like data releases in previous 2-year cycles. Additionally, data for variables newly collected in 2019–March 2020 that are not comparable with data from 2017–2018 will be released through RDC as convenience samples with no sample weights. RDC contains data sets with (a) data items that were collected from a single-year sample (before 2019) or collected for any period other than a public-release 2-year cycle (identified as such in the component description); (b) data merged geographically to some other contextual data files (often supplied by the data user); and (c) data items determined to be too sensitive or too detailed to be released to the public due to confidentiality restrictions. NHANES data linked with administrative data (Medicare, Medicaid or Children’s Health Insurance Program, and Social Security Administration) are only available through RDC.

If a special data file involves subsampling, then special subsample weights that reflect the number of calendar years in the data file and the rate of subsampling were created for that file. For all special data files, appropriate documentation is provided in RDC to describe the necessary sample weights.

Unmasked PSU and variance stratum codes (which differ from the MVU codes in the public-use files) can be provided for variance estimation for restricted data files in RDC. These unmasked variance codes are necessary for studies that use geographically defined variables or for studies that

**Table B. Age-related variables on the publicly released data files: National Health and Nutrition Examination Survey, 2007–March 2020**

Variable	Description	2007–2008 data files	2009–2010 data files	2011–2012 data files	2013–2014 data files	2015–2016 data files	2017–March 2020 data files
RIDAGEYR	Age in years at screening (for people aged 0–80 years)	Yes	Yes	Yes	Yes	Yes	Yes
RIDAGEMN	Age in months at screening (for people aged 0–79 years)	Yes	Yes	Yes—for children aged 24 months and under	Yes—for children aged 24 months and under	Yes—for children aged 24 months and under	Yes—for children 24 aged months and under
RIDAGEEX	Age in months at MEC examination (for people aged 0–79 years)	Yes	Yes	No	No	No	No
RIDEXAGM	Age in months at MEC examination (for people aged 0–19 years at screening)	No	No	Yes	Yes	Yes	No
RIDEXAGY	Age in years at MEC examination (for people aged 2–19 years at screening)	No	No	Yes	No	No	No

NOTE: MEC is mobile examination center.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2007–March 2020.

geographically merge NHANES data with external data sets. Additionally, Fay's adjusted balanced repeated replication weights (20) for 2-year data (and the 2017–March 2020 prepandemic public-use data files), and jackknife weights for single-year data (through 2018) with unmasked variance units are available through RDC. If replication methods are to be used for combined survey years or cycles with unmasked variance units, then replicate weights must be computed by the analyst in RDC. Providing the unmasked PSU and stratum codes does not create disclosure risk because of the access restrictions in place at RDC.

More information on RDC and lists of commonly used restricted NHANES data files and variables are available from the NCHS website at: <https://www.cdc.gov/rdc/>. Other restricted variables not listed here may be available by request. Information on proposals for use of stored specimens is available from the NHANES website: <https://www.cdc.gov/nchs/nhanes/biospecimens/biospecimens.htm>.

## Summary

The COVID-19 pandemic required suspension of the NHANES 2019–2020 field operations. Consequently, data collection for the NHANES 2019–2020 cycle was not completed, and the collected data are not nationally representative. The partial 2019–March 2020 data were combined with the full data set from the previous cycle (2017–2018) to create nationally representative 2017–March 2020 prepandemic data files. A PSU-level adjustment factor was created to equalize the contribution of each stratum to the total survey sample and applied to participant base weights. The performance of interview weights was assessed by comparing the demographic characteristics of the weighted NHANES 2017–March 2020 prepandemic sample to nationally representative estimates from the 2018 5-year ACS.

Additionally, a simulation was created using data for 2013–2016, where two important health estimates from a partial data set, with PSU adjustments calculated with the same methodology as 2017–March 2020, were similar to those from the full NHANES 2013–2016 data set. Although a thorough evaluation was conducted to assess the reliability of using this data set to produce national estimates, some considerations should be made before these data are analyzed, including implications of the 3.2-year period for combining the 2017–March 2020 data files with previous cycles and analyzing trends.

## References

1. Chen TC, Clark J, Riddles MK, Mohadjer LK, Fakhouri THI. National Health and Nutrition Examination Survey, 2015–2018: Sample design and estimation procedures. National Center for Health Statistics. *Vital Health Stat* 2(184). 2020.
2. Johnson CL, Dohrmann SM, Burt VL, Mohadjer LK. National Health and Nutrition Examination Survey: Sample design, 2011–2014. National Center for Health Statistics. *Vital Health Stat* 2(162). 2014.
3. Krenzke T, Mohadjer L. Application of probability-based link-tracing and nonprobability approaches to sampling out-of-school youth in developing countries. *J Surv Stat Methodol* 9(4):1–26. 2020.
4. Kalton G, Flores-Cervantes I. Weighting methods. *J Off Stat* 19(2):81–97. 2003.
5. Potter F. A study of procedures to identify and trim extreme sampling weights. In: 1990 Proceedings of the American Statistical Association, Survey Research Methods Section. Alexandria, VA: American Statistical Association. 1990.
6. Fakhouri THI, Martin CB, Chen TC, Akinbami L, 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.
7. National Center for Health Statistics. Health, United States, 2001: With urban and rural health chartbook. Hyattsville, MD. 2001.
8. Agency for Healthcare Research and Quality. National healthcare quality and disparities report: Chartbook on rural health care. 2017.
9. Singh GK, Siahpush M. Widening rural-urban disparities in life expectancy, U.S., 1969–2009. *Am J Prev Med* 46(2):e19–29. 2014.
10. Ingram DD, Franco SJ. 2013 NCHS urban–rural classification scheme for counties. National Center for Health Statistics. *Vital Health Stat* 2(166). 2014.
11. U.S. Census Bureau. American Community Survey design and methodology. 2014. Available from: [https://www2.census.gov/programs-surveys/acs/methodology/design\\_and\\_methodology/acs\\_design\\_methodology\\_report\\_2014.pdf](https://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/acs_design_methodology_report_2014.pdf).
12. 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. 2017.
13. National Center for Health Statistics. NHANES response rates and population totals. Available from: <https://www.cdc.gov/nchs/nhanes/ResponseRates.aspx>.
14. National Center for Health Statistics. National Health and Nutrition Examination Survey: Analytic guidelines, 2011–2014 and 2015–2016. 2018.
15. Parker JD, Talih M, Malec DJ, Beresovsky V, Carroll M, Gonzalez Jr JF, et al. National Center for Health Statistics data presentation standards for proportions. National Center for Health Statistics. *Vital Health Stat* 2(175). 2017.

16. National Center for Health Statistics. National Health and Nutrition Examination Survey. Available from: <https://www.cdc.gov/nchs/nhanes/index.htm>.
17. Stierman B, Afful J, Carroll MD, Chen TC, Davy O, Fink S, et al. National Health and Nutrition Examination Survey 2017–March 2020 prepandemic data files—Development of files and prevalence estimates for selected health outcomes. National Health Statistics Reports; no 158. Hyattsville, MD: National Center for Health Statistics. 2021. DOI: <https://dx.doi.org/10.15620/cdc:106273>.
18. Brown G. Survey weight adjustment using PROC WTADJUST from SUDAAN V10. In: 2010 American Association for Public Opinion Research Annual Meeting. Chicago, IL: RTI International. 2010.
19. Kott P. WTADJX is coming: Calibration weighting in SUDAAN when unit nonrespondents are not missing at random and other applications. 2011. RTI International.
20. Judkins D. Fay's method for variance estimation. *J Off Stat* 6(3):223–39. 1990.



# Appendix I. Supporting Tables

**Table I. Race and Hispanic-origin and income group sampling fractions used to calculate primary sampling unit measure of sizes: National Health and Nutrition Examination Surveys, 2015–2018 and 2019–2022**

Race and Hispanic origin	Sampling fraction values	
	2015–2018	2019–2022
Hispanic . . . . .	0.000206	0.000266
Non-Hispanic Black . . . . .	0.000287	0.000382
Non-Hispanic, non-Black Asian . . . . .	0.000453	0.000572
Non-Hispanic, low income, White and other races and ethnicities <sup>1</sup> . . . .	0.000184	0.000182
Non-Hispanic, non-low income, White and other races and ethnicities <sup>1</sup> . . . . .	0.000086	0.000147

<sup>1</sup>Excludes non-Hispanic Black and non-Hispanic Asian people.

NOTES: For more information on the calculation of the sampling fraction values, see “National Health and Nutrition Examination Survey: Sample Design, 2011–2014” ([https://www.cdc.gov/nchs/data/series/sr\\_02/sr02\\_162.pdf](https://www.cdc.gov/nchs/data/series/sr_02/sr02_162.pdf)). See “Sample Design” in this report for more information on how the data sources were combined.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Surveys, 2015–2018 and 2017–March 2020.

**Table II. Final sampling rates and initial base weights: National Health and Nutrition Examination Surveys, 2017–2018 and 2019–March 2020**

Race and Hispanic-origin–income–sex–age sampling domain <sup>1</sup>	2017–2018		2019–2020	
	Numerator of sampling rate <sup>2</sup>	Initial base weight	Numerator of sampling rate <sup>2</sup>	Initial base weight
<b>Non-Hispanic Black</b>				
Male and female:				
Under 1	1.00	1,814.82	1.00	1,520.18
1–2	0.84	2,158.28	1.00	1,520.18
3–5	0.60	3,005.41	0.75	2,014.54
Male:				
6–11	0.62	2,915.84	0.78	1,942.86
12–19	0.50	3,638.12	0.59	2,571.11
20–39	0.39	4,623.63	0.46	3,312.77
40–49	0.50	3,644.75	0.61	2,488.96
50–59	0.48	3,762.67	0.57	2,673.96
60 and over	1.00	1,814.82	0.85	1,798.64
Female:				
6–11	0.64	2,826.73	0.72	2,110.97
12–19	0.51	3,524.72	0.56	2,720.17
20–39	0.33	5,434.63	0.38	3,965.75
40–49	0.40	4,506.67	0.43	3,517.80
50–59	0.37	4,842.46	0.44	3,453.26
60 and over	0.67	2,703.47	0.60	2,517.27
<b>Hispanic</b>				
Male and female:				
Under 1	0.93	1,950.64	1.00	1,520.18
1–2	0.47	3,865.98	0.62	2,466.15
3–5	0.34	5,286.79	0.47	3,267.67
Male:				
6–11	0.34	5,299.42	0.44	3,448.29
12–19	0.29	6,155.91	0.32	4,742.44
20–39	0.25	7,347.30	0.28	5,405.46
40–49	0.28	6,398.99	0.35	4,348.67
50–59	0.42	4,317.43	0.43	3,563.98
60 and over	1.00	1,814.82	0.85	1,785.00
Female:				
6–11	0.38	4,806.88	0.43	3,544.75
12–19	0.32	5,660.18	0.34	4,522.28
20–39	0.23	7,960.54	0.27	5,734.65
40–49	0.28	6,374.51	0.28	5,355.72
50–59	0.39	4,696.32	0.37	4,079.26
60 and over	0.88	2,055.93	0.64	2,361.66
<b>Non-Hispanic, non-Black Asian</b>				
Male and female:				
Under 1	0.95	1,917.73	1.00	1,520.18
1–2	1.00	1,814.82	1.00	1,520.18
3–5	1.00	1,814.82	1.00	1,520.18
Male:				
6–11	1.00	1,814.82	1.00	1,520.18
12–19	1.00	1,814.82	1.00	1,520.18
20–39	0.63	2,874.44	0.71	2,148.72
40–49	0.80	2,260.34	0.83	1,840.10
50–59	1.00	1,814.82	1.00	1,520.18
60 and over	1.00	1,814.82	1.00	1,520.18
Female:				
6–11	1.00	1,814.82	1.00	1,520.18
12–19	1.00	1,814.82	1.00	1,520.18
20–39	0.60	3,004.83	0.70	2,172.98
40–49	0.68	2,687.58	0.69	2,195.31
50–59	0.77	2,370.10	0.95	1,607.40
60 and over	1.00	1,814.82	1.00	1,520.18

See footnotes at end of table.

**Table II. Final sampling rates and initial base weights: National Health and Nutrition Examination Surveys, 2017–2018 and 2019–March 2020—Con.**

Race and Hispanic-origin–income–sex–age sampling domain <sup>1</sup>	2017–2018		2019–2020	
	Numerator of sampling rate <sup>2</sup>	Initial base weight	Numerator of sampling rate <sup>2</sup>	Initial base weight
Non-Hispanic White or other races and ethnicities <sup>3</sup> , low income				
Male and female:				
Under 1	1.00	1,814.82	1.00	1,520.18
1–2	0.60	3,033.38	0.60	2,525.40
3–5	0.37	4,849.75	0.41	3,672.84
Male:				
6–11	0.40	4,502.70	0.38	4,015.25
12–19	0.33	5,419.49	0.33	4,665.57
20–29	0.22	8,324.09	0.20	7,580.63
30–39	0.30	6,059.67	0.25	6,116.56
40–49	0.33	5,431.33	0.30	4,993.21
50–59	0.33	5,552.45	0.24	6,287.06
60–69	0.38	4,776.97	0.23	6,638.41
70–79	0.84	2,153.25	0.34	4,467.06
80 and over	1.00	1,814.82	0.61	2,498.17
Female:				
6–11	0.45	4,022.33	0.37	4,056.54
12–19	0.38	4,761.14	0.31	4,936.79
20–29	0.13	13,588.21	0.13	11,559.79
30–39	0.19	9,309.00	0.16	9,595.32
40–49	0.26	7,019.96	0.21	7,152.74
50–59	0.22	8,247.26	0.20	7,581.81
60–69	0.28	6,469.73	0.16	9,427.68
70–79	0.31	5,780.92	0.21	7,401.21
80 and over	0.52	3,479.10	0.32	4,816.97
Non-Hispanic White or other races and ethnicities <sup>3</sup> , non-low income				
Male and female:				
Under 1	0.78	2,332.29	1.00	1,520.18
1–2	0.50	3,597.49	0.61	2,491.20
3–5	0.29	6,214.22	0.42	3,578.17
Male:				
6–11	0.26	7,012.11	0.43	3,514.05
12–19	0.17	10,528.23	0.29	5,298.19
20–29	0.11	16,229.23	0.18	8,529.52
30–39	0.11	16,493.76	0.16	9,495.47
40–49	0.10	17,502.40	0.16	9,350.76
50–59	0.09	20,224.32	0.14	10,936.48
60–69	0.10	18,446.86	0.13	11,269.16
70–79	0.19	9,634.63	0.20	7,474.43
80 and over	0.45	4,067.35	0.64	2,374.27
Female:				
6–11	0.26	7,043.13	0.43	3,536.70
12–19	0.19	9,501.80	0.31	4,964.45
20–29	0.14	13,400.71	0.15	9,971.57
30–39	0.11	17,060.86	0.13	11,392.59
40–49	0.09	19,706.19	0.14	10,547.06
50–59	0.08	21,360.31	0.13	11,984.57
60–69	0.09	19,617.21	0.11	13,449.93
70–79	0.18	10,014.45	0.22	6,888.80
80 and over	0.39	4,655.36	0.58	2,624.42

<sup>1</sup>Age in years.

<sup>2</sup>Corresponds to a 180% sample; sampling rates are calculated by dividing the numerator by 1,815 for 2017–2018, and by 1,520 for 2019–2020.

<sup>3</sup>Excludes non-Hispanic Black and non-Hispanic Asian people.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Surveys, 2017–2018 and 2019–March 2020.

**Table III. Variables used to form nonresponse adjustment cells for weighting interview samples: National Health and Nutrition Examination Survey, 2017–March 2020**

Variables considered for nonresponse, by age group (years)	Categories of variables cross-classified to form nonresponse adjustment cells
<b>0–5</b>	
Sex of household reference person	Male, female
Household composition	One sampled person in household under age 16 years, more than one sampled person in household all under age 16, more than one sampled person in household of mixed ages
Race and Hispanic origin and income level of sampled person	Non-Hispanic Black; Hispanic; non-Hispanic Asian; non-Hispanic White or other races and ethnicities, low income; non-Hispanic White or other races and ethnicities, non-low income
Census region	Northeast, Midwest, South, West
Household size	1–2, 3–4, 5–6, 7 or more
<b>6–19</b>	
Sex of household reference person	Male, female
Race and Hispanic origin and income level of sampled person	Non-Hispanic Black; Hispanic; Non-Hispanic Asian; Non-Hispanic White or other races and ethnicities, low income; non-Hispanic White or other races and ethnicities, non-low income
Household size	1–2, 3–4, 5–6, 7 or more
Census region	Northeast, Midwest, South, West
Sex of sampled person	Male, female
Population of the primary sampling unit	Less than 100,000; 100,000 or more to less than 250,000; 250,000 or more to less than 1,000,000; 1,000,000 or more
Urban or rural based on NCHS CODE13 <sup>1</sup>	Large central metro counties in MSA of 1 million population; large fringe metro counties in MSA of 1 million or more; medium and small metro counties in MSA of less than 999,999 population; micropolitan counties in micropolitan statistical area and noncore counties not in micropolitan statistical area
Tract-level percentage of population born in the United States	First quartile, second quartile, third quartile, fourth quartile
<b>20–39</b>	
Census region	Northeast, Midwest, South, West
Sex of household reference person	Male, female
Race and Hispanic origin and income level of sampled person	Non-Hispanic Black; Hispanic; non-Hispanic Asian; non-Hispanic White or other races and ethnicities, low income; non-Hispanic White or other races and ethnicities, non-low income
Sex of sampled person	Male, female
Household composition	One sampled person in household at least age 16 years, more than one sampled person in household all at least age 16, more than one sampled person in household of mixed ages
<b>40–59</b>	
Census region	Northeast, Midwest, South, West
Household composition	One sampled person in household at least age 16 years, more than one sampled person in household all at least age 16, more than one sampled person in household of mixed ages
Race and Hispanic origin and income level of sampled person	Non-Hispanic Black; Hispanic; non-Hispanic Asian; non-Hispanic White or other races and ethnicities, low income; non-Hispanic White or other races and ethnicities, non-low income
Population of the primary sampling unit	Less than 100,000; 100,000 or more to less than 250,000; 250,000 or more to less than 1,000,000; 1,000,000 or more
Sex of household reference person	Male, female
Sex of sampled person	Male, female
<b>60 and over</b>	
Race and Hispanic origin and income level of sampled person	Non-Hispanic Black; Hispanic; non-Hispanic Asian; non-Hispanic White or other races and ethnicities, low income; non-Hispanic White or other races and ethnicities, non-low income
Household composition	One sampled person in household at least age 16 years, more than one sampled person in household all at least age 16, more than one sampled person in household of mixed ages
Population of the primary sampling unit	Less than 100,000; 100,000 or more to less than 250,000; 250,000 or more to less than 1,000,000; 1,000,000 or more
Sex of household reference person	Male, female

<sup>1</sup>Based on the 2013 NCHS Urban–Rural Classification Scheme for Counties ([https://www.cdc.gov/nchs/data\\_access/urban\\_rural.htm](https://www.cdc.gov/nchs/data_access/urban_rural.htm)); also see reference 10 in this report.

NOTES: The group “non-Hispanic White or other races and ethnicities” excludes non-Hispanic Black and non-Hispanic Asian people. MSA is metropolitan statistical area.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2017–March 2020.

**Table IV. Variables used to form nonresponse adjustment cells for weighting mobile examination center examination samples: National Health and Nutrition Examination Survey, 2017–March 2020**

Variables considered for nonresponse, by age group (years)	Categories of variables cross-classified to form nonresponse adjustment cells
0–5	
Census region	Northeast, Midwest, South, West
Household composition	One sampled person in household under age 16 years, more than one sampled person in household all under 16, more than one sampled person in household of mixed ages
Sex of household reference person	Male, female
6–19	
Household composition	One sampled person in household under age 16 years, one sampled person in household over age 16, more than one sampled person in household all under age 16, more than one sampled person in household all over age 16, more than one sampled person in household of mixed ages
Sex of sampled person	Male, female
Census region	Northeast, Midwest, South, West
Sex of household reference person	Male, female
Tract-level percentage of population aged 18 and over with a disability	First quartile, second quartile, third quartile, fourth quartile
20–39	
Sex of sampled person	Male, female
40–59	
Census region	Northeast, Midwest, South, West
Sex of household reference person	Male, female
Self-reported height and weight indicated sampled person was obese	No, yes
Sex of sampled person	Male, female
60 and over	
Census region	Northeast, Midwest, South, West
Household composition	One sampled person in household at least age 16 years, more than one sampled person in household all at least age 16, more than one sampled person in household of mixed ages
Race and Hispanic origin and income level of sampled person	Non-Hispanic Black; Hispanic; non-Hispanic Asian; non-Hispanic White or other races and ethnicities <sup>1</sup> , low income; non-Hispanic White or other races and ethnicities <sup>1</sup> , non-low income
Sex of household reference person	Male, female
Self-reported height and weight indicated sampled person was obese	No, yes

<sup>1</sup>Excludes non-Hispanic Black and non-Hispanic Asian people.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2017–March 2020.

**Table V. Most common survey sample weights and their appropriate use: National Health and Nutrition Examination Survey, 2017–March 2020**

Sample weights	Application	Notes
Full samples		
Interview weights (WTINTRP)	Use when analyzing data from the home interview only. Do not use if the analysis includes variables that were also collected on examined persons in the mobile examination center (MEC).	...
MEC examination weights (WTMECPRP)	Use when analyzing data from the MEC examination. Do not use if the analysis includes variables collected as part of one of the dietary interviews or part of one of the subsamples (for example, fasting or environmental).	...
Dietary day 1 sample weights (WTDRD1PP)	Use when analyzing data from sample with completed day 1 24-hour dietary recall or the Flexible Consumer Behavior Survey telephone follow-up module.	Dietary day 1 sample weights were adjusted based on race and Hispanic origin, age group, sex, and day 1 weekday-weekend categories (Weekdays, Friday, Saturday, and Sunday). To account for day of the week consumption, the weights are equally distributed to each of the 7 days of the week. So the weights for the four categories are 4/7, 1/7, 1/7, and 1/7 of the total weights.
Dietary day 1 and day 2 completed sample (WTDR2DPP)	Use when analyzing data from the smaller sample with 2 days of completed 24-hour dietary recall.	Dietary 2-day sample weights were adjusted based on race and Hispanic origin, age group, sex, and day 1 and day 2 weekday-weekend categories (weekday-weekday, weekday-weekend, weekend-weekday, and weekend-weekend). To account for day of the week consumption, the weights are equally distributed to each of the 7 days of the week for both day 1 and day 2. So, the weights for the four categories are 16/49, 12/49, 12/49, and 9/49 of the total weights.
One-half subsamples		
Fasting subsample weights (WTSAFPRP)	Use when analyzing the plasma fasting glucose, insulin levels, triglycerides, and low-density lipoprotein cholesterol (lipids) only for examined people assigned to and meeting the criteria for the fasting subsample.	Fasting subsample weights were for participants aged 12 years and over who were examined in a morning session and fasted 8–23 hours. Diagnosed diabetes is a separate poststratification category (to the interview weights).
Folates subsample weights (WTSFPRP)	Use when analyzing whole blood folate and serum total folate.	Folates subsample weights were for participants aged 1 year and over who were examined in a MEC in 2017–2020. The subsample includes: all female participants aged 12–49 years; and one-half sample of all the other participants.
Folates subsample weights RDC (WTSFOL)	Use when analyzing whole blood folate and serum total folate.	Folates subsample weights were for participants aged 1 year and over who were examined in a MEC in 2017–2020. The subsample includes: all participants aged 1 year and over from 2019 through 2020; all female participants aged 12–49 years from 2017 through 2018; and one-half sample of all the other 2017–2018 participants.
Volatile organic compound (VOC) subsample weights (WTSVOCPR)	Use when analyzing data from VOC one-half laboratory subsample for examined people assigned to and meeting the criteria for this subsample.	VOC subsample weights were for participants aged 12 years and over.

See footnotes at end of table.

**Table V. Most common survey sample weights and their appropriate use: National Health and Nutrition Examination Survey, 2017–March 2020—Con.**

Sample weights	Application	Notes
One-third subsamples		
Environmental chemicals A weights (WTSAPRP)	Use when analyzing data from the one-third laboratory environmental subsample A for examined people assigned to and meeting the criteria for this subsample.	Environmental subsample A, B, C weights were each for a one-third subsample of examined participants aged 6 years and over. Participants were assigned to one of the three mutually exclusive 1/3 environmental subsamples based on the sampling scheme. The analytes in each of the three subsamples varied, depending on what was included in the cycle. The names Subsample A, Subsample B, and Subsample C are for convenience and not representative of any ordering. The proper subsample weights attached in the data set should be used to analyze. Because the same analytes might be in different subsamples, it is important to check the weight.
Environmental chemicals B weights (WTSBPRP)	Use when analyzing data from the one-third laboratory environmental subsample B for examined people assigned to and meeting the criteria for this subsample.	See environmental chemical A weights.
Environmental chemicals C weights (WTSCPRP)	Use when analyzing data from the one-third laboratory environmental subsample C for examined people assigned to and meeting the criteria for this subsample.	See environmental chemical A weights.

... Category not applicable.

SOURCE: National Center for Health Statistics, National Health and Nutrition Examination Survey, 2017–March 2020.

# Appendix II. Definitions of Terms

---

**Calibration**—A statistical method that adjusts sample estimates of totals or percentages of target characteristics (for example, sex, race and Hispanic origin, and age) to equal the target population totals or percentages. The adjustment cells for calibration are like the domains used for sample selection but can include variables that were not used for the original sampling frame.

**Civilian noninstitutionalized population**—Includes all people living in households and noninstitutional group quarters who are not active members of the military. This is the target population for NHANES.

**Demographic subgroup**—A demographic group defined by race and Hispanic origin, age, and sex, which might be a single sample domain or a combination of multiple sample domains.

**Domain**—A demographic group of analytic interest (analytic domain). Analytic domains may also be sampling domains if a sample design is created to meet goals for those specific demographic groups. For the National Health and Nutrition Examination Survey (NHANES), sampling domains are defined by race and Hispanic origin, income, age, and sex. See Sampling domain.

**Dwelling unit (DU)**—Also called a “housing unit.” A house, apartment, mobile home or trailer, group of rooms, or single room occupied as separate living quarters or, if vacant, intended for occupancy as separate living quarters. Separate living quarters are those in which the occupants live separately from other people in the building, and that have direct access from outside the building or through a common hall. In this report, the term generally means those DUs that are eligible for the survey (excluding institutional group quarters), or that could become eligible (for example, vacant at the time of sampling but could be occupied when screening begins).

**Household**—All the people who live in a housing unit as their usual place of residence.

**Masked variance units (MVUs)**—A collection of secondary sampling units aggregated into groups for the purpose of variance estimation and designed to prevent disclosure of the identity of the selected primary sampling units (PSUs). For NHANES, rather than use the units as sampled, some pseudo units are created. The resulting units produce variance estimates that closely approximate the “true” design variance estimates. MVUs have been created for all 2-year survey cycles from NHANES 1999–2000 through 2017–2018. They can also be used for analyzing any combined 4- to 20-year data set.

**Measure of size (MOS)**—A value assigned to every sampling unit in a sample selection, usually a count of units associated with the elements to be selected. For NHANES, MOS is a weighted average of estimates of population counts for the race–Hispanic-origin–income groups of interest.

**Participant**—A person selected into the NHANES sample during screening (sample person) who agrees to participate in the survey. In NHANES, people agreeing to complete the in-home interview are considered “interview participants.” People agreeing to complete both the in-home interview and an examination in a mobile examination center (MEC) are considered “MEC participants.”

**Primary sampling unit (PSU)**—The first-stage selection unit in a multistage area probability sample. In NHANES, PSUs are counties or groups of counties in the United States. Some PSUs are so large that they are selected into the survey with a probability of one. These are referred to as PSUs selected with certainty (“certainty PSUs” or often referred to as self-representing [SR] PSUs); all other PSUs are selected without certainty (“noncertainty PSUs” or non-self-representing PSUs).

**Public-use data file**—An electronic data set containing respondent records from a survey with a subset of variables collected in the survey that have been reviewed by analysts from the National Center for Health Statistics (NCHS) to ensure that participant identities are protected. NCHS distributes these files to encourage public use of the survey data.

**Race and Hispanic origin**—Unless otherwise specified, race and Hispanic origin in this report is used the same as for NHANES sample selection and sample weight construction. It refers to Hispanic people; non-Hispanic Black people; non-Hispanic, non-Black Asian people; and a fourth group that includes non-Hispanic White and all other races and ethnicities. This is different from the race and Hispanic-origin categories in the public-use data files as reported by participants in the interview (non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, Mexican American, other Hispanic, and other race including multiracial).

**Replicates**—Subsamples selected repeatedly from a sample used in some variance estimation approaches. With these approaches, the statistic of interest is calculated for each subsample, and the variability among the replicate statistics is used to estimate the variance of the full sample statistic. The jackknife and Fay’s adjusted balanced repeated replication (Fay’s BRR) methods are two common procedures for the derivation of replicates from a full sample. Fay’s BRR method was used to create replicate weights for most of the



NHANES 2015–2018 and prepandemic multiyear samples. Replicate weights are available through the NCHS Research Data Center (RDC).

**Response rate**—The number of survey respondents divided by the number of people selected into the sample.

**Restricted-use data file**—An electronic data set of survey respondent records, containing some information that may, if released to the public, risk identifying individual survey respondents. These data are available only through the NCHS RDC. These include data sets with (a) data items collected from a single year, (b) data geographically linked to other contextual data files (often supplied by the data user), and (c) data items determined to be too sensitive or too detailed to be released to the public due to confidentiality restrictions.

**Sample weight**—The estimated number of people in the target population that an NHANES respondent represents. For example, if a man in the sample represents 12,000 men in his race–Hispanic-origin–sex–age domain, then his sample weight is 12,000. The NHANES sample weights were adjusted for different sampling rates (of the race–Hispanic-origin–income–sex–age groups), different response rates, and different coverage rates among people in the sample, so accurate national estimates could be made from the sample. The product of these adjustments is sometimes called the “final” sample weight.

**Sampled person**—A person selected for the NHANES sample based on oversampling targets for certain demographic subgroups. Sampled persons may become participants (agree to take part in the survey) or nonrespondents.

**Sampling domain**—NHANES 2015–2018 and NHANES 2019–2022 sample designs each included 87 sampling domains, which were defined by race and Hispanic origin, income, age, and sex. Every person in the NHANES target population can be classified into exactly one of the 87 sampling domains. See Domain.

**Sampling error**—The portion of the difference between the sample estimate and the true population value due to only observing a sample rather than the entire population.

**Sampling rate**—The rate at which a unit is selected from a sampling frame. For NHANES, the rates required for sampling people in the race–Hispanic-origin–income–sex–age domains were designed to achieve the designated number of MEC examinations in each of those domains. The sampling rates are the driving force in all stages of sampling.

**Screener**—An interview (usually short) containing a set of questions asked of a household member to determine whether the household contains anyone who could be selected into the sample for the survey. For NHANES, the screener consisted of compiling a household roster and collecting the income level of the household, and the race and Hispanic origin, age, and sex of all household members. For NHANES, only people aged 18 and over can answer the screener.

**Screening**—The process of conducting, or attempting to conduct, the screener interview in selected DUs (households). Occupied DUs are screened using the screener. Other units can also be screened; the process for these units is verification that they are either vacant or not DUs. See Screener.

**Secondary sampling unit**—The second-stage selection unit in a multistage area probability sample. For NHANES, these are typically referred to as “segments.” See Segment.

**Segment**—A group of housing units located near each other, all of which were considered for selection into the sample. For NHANES, segments consist of a census block, or group of blocks, and their selection makes up the second stage of sampling. Within each segment, a sample of DUs was selected.

**Stratification; Strata**—The dividing of a population of sampling units into mutually exclusive categories (strata). Typically, stratification is used to increase the precision of survey estimates for subpopulations important to the survey’s objectives. To select the PSUs fielded during 2015–2018, PSUs were stratified based on health status, metropolitan statistical area status, and various population demographics.

**Study location**—The set of segments within a PSU that were fielded together, with all MEC examinations conducted at the same physical location. The distinction between a PSU and a study location is necessary because some large certainty PSUs were divided into multiple study locations and fielded at different times.

**Target population**—The population to be described by estimates from the survey. In NHANES, the target population is the resident civilian noninstitutionalized population of the United States, which excludes all people in supervised care or custody in institutional settings, all active-duty military personnel, active-duty family members living overseas, and any other people living outside the 50 states and the District of Columbia.

**Undercoverage**—The result of failing to include all the target population within the sampling frame.

**Variance**—A measure of the dispersion of a set of numbers. In this report, the variance is specifically the sample variance, which is a measure of the variation of a statistic, such as a proportion or a mean, calculated as a function of the sampling design and the population parameter being estimated. Many common statistical software packages compute “population variances” by default; these may underestimate the sampling variance because they do not incorporate any effects of taking a sample instead of collecting data from every person in the full population. Estimating the variance in NHANES requires special statistical software, as discussed in this report.

*Variance stratum*—The cluster of variance units used when forming a replicate for variance estimation. For NHANES, the PSU sampling strata usually correspond to the variance strata.

*Variance unit*—A collection of secondary sampling units aggregated into groups and excluded when forming a replicate for variance estimation. For NHANES, an entire PSU usually corresponds to a variance unit.

*Weight*—See Sample weight.

# 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

---

OFFICIAL BUSINESS  
PENALTY FOR PRIVATE USE, \$300

FIRST CLASS MAIL  
POSTAGE & FEES PAID  
CDC/NCHS  
PERMIT NO. G-284



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