Monitoring the Nation's Health

SAFER • HEALTHIER • PEOPLE™

National Health and Nutrition Examination Survey: Estimation Procedures, 2007–2010

Vital and Health Statistics

Series 2, Number 159

August 2013

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

Suggested citation

Mirel LB, Mohadjer LK, Dohrmann SM, et al. National Health and Nutrition Examination Survey: Estimation procedures, 2007–2010. National Center for Health Statistics. Vital Health Stat 2(159). 2013.

Library of Congress Cataloging-in-Publication Data

National Health and Nutrition Examination Survey : Estimation procedures, 2007–2010.
p. ; cm.— (Vital and health statistics, series 2, data evaluation and methods research ; number 159) (DHHS publication ; no. (PHS) 2013–1359)
Estimation procedures, 2007–2010
Includes bibliographical references.
ISBN 0–8406–0659–1
I. National Center for Health Statistics (U.S.), issuing body. II. Title: Estimation procedures, 2007–2010. III. Series: Vital and health statistics. Series 2, Data evaluation and methods research ; no. 159. IV. Series: DHHS publication ; no. (PHS) 2013–1359.
[DNLM: 1. National Health and Nutrition Examination Survey (U.S.) 2. Health Surveys—United States. 3. Body Weight—United States. 4. Nutrition Surveys—United States. 5. Sampling Studies—United States. 6. Statistics as Topic—United States. W2 A N148vb no.159 2013]
RA407.3
614.4'273—dc23

2013022786

For sale by the U.S. Government Printing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20402–9328
Printed on acid-free paper.

Vital and Health Statistics

Series 2, Number 159

National Health and Nutrition Examination Survey: Estimation Procedures, 2007–2010

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
August 2013
DHHS Publication No. 2013–1359

Contents

Acknowledgments

The authors gratefully acknowledge the assistance of Margaret Carroll, Michele Chiappa, Rosemarie Hirsch, Deanna Kruszon-Moran, and Ryne Paulose-Ram in the preparation of this report.

Abstract

Background

Data collection for the National Health and Nutrition Examination Survey (NHANES), comprises three levels: an initial household screening interview (or "screener"), an in-home personal interview, and a physical examination. The primary objective of the screener is to determine whether any household members are eligible for the interview and examination. Eligibility is determined by preset selection probabilities for the desired demographic subdomains. After an eligible sample person is selected, the in-home interview collects person-level demographic, health, and nutrition information, as well as information about the household. The examination includes physical measurements such as blood pressure, a dental examination, and the collection of blood and urine specimens for laboratory testing.

Objectives

This report provides background for the NHANES program and summarizes the sample design specifications for the 2007–2010 survey cycle. Estimation procedures are then presented, including the methods used to calculate survey weights for the full sample and for examination subsamples, as well as guidelines for combining 2-year weights for the analysis of multiyear data. Finally, the appropriate variance estimation methods are described. The sample selection methods, survey content, data collection procedures, and methods for assessing nonsampling errors are documented elsewhere.

Keywords: sampling • weighting • variance estimation

National Health and Nutrition Examination Survey: Estimation Procedures, 2007–2010

by Lisa B. Mirel, M.S., National Center for Health Statistics; Leyla K. Mohadjer, Ph.D., Sylvia M. Dohrmann, M.S., and Jason Clark, M.S., Westat; and Vicki L. Burt, Sc.M., R.N., Clifford L. Johnson, M.S.P.H., and Lester R. Curtin, Ph.D., National Center for Health Statistics

Introduction

The National Health and Nutrition Examination Survey (NHANES) is one of a series of health-related programs conducted by the Centers for Disease Control and Prevention's (CDC) National Center for Health Statistics (NCHS) to provide information on the health and nutritional status of the U.S. population. This information is used to estimate the prevalence of various diseases and conditions and to provide information for use in planning health policy.

NHANES provides information on the noninstitutionalized civilian resident population of the United States. It excludes all persons in supervised care or custody in institutional settings, all active-duty military personnel, active-duty family members living overseas, and any other U.S. citizens residing outside the 50 states and the District of Columbia. See the glossary in Appendix I for further details on institutional and noninstitutional group quarters.

Data collection for NHANES comprises three levels: an initial household screening interview (or ''screener''), an in-home personal interview, and a standardized physical examination in a specially equipped mobile examination center (MEC), including selected objective measures of health status. The primary objective of the screener is to determine whether any household members are eligible for the interview and examination. The in-home interview collects person-level demographic, health, and nutrition information, as well as information about the household. The examination includes physical measurements such as blood pressure, a dental examination, and the collection of blood and urine specimens for laboratory testing. Conducting the examinations in the MEC helps to standardize their administration.

The practical considerations surrounding the transportation of the MECs across the country limit the number of locations NHANES can visit each year. As a result, parameter estimates for single-year data are relatively unstable due to large variance estimates. To improve the statistical reliability and stability of estimates with larger variances, it is recommended that analysts use combinations of 2-year cycles. Combining data from 2-year cycles is particularly appropriate for rare events, for estimates pertaining to detailed demographic subdomains, and for measures that may have considerable geographic variation. The differences in the sample sizes and designs for all cycles of NHANES should be considered when comparisons are made across various HANES surveys. More details on the sample design for NHANES 2007–2010, as well as the sample design parameters for all six NHANES surveys, are outlined in

"National Health and Nutrition Examination Survey: Sample Design, 2007–2010" (1).

Aggregate-level national annual estimates may be made for NHANES 1999–2010 through the NCHS Research Data Center. Although the annual samples are nationally representative, annual estimates should be produced only for the nation as a whole, for race and Hispanic origin subdomains, or for very broad sex-age subdomains within race and Hispanic origin subdomains because of limited sample sizes and larger variances of point estimates (where annual samples are not publicly available).

The next section briefly summarizes the sample design specifications for the 2007–2010 survey, and the remainder of the report presents the estimation procedures. For more detail on the 2007–2010 sample design, see the Sample Design Report (1). The "Weighting the Sample Data" section discusses the creation of weights for the entire sample and subsamples, and the "Variance Estimation" section describes the appropriate variance estimation methods. Appendix I presents a glossary of terms, and Appendix II contains tables of supporting material. Documentation of the survey content, data collection procedures, and methods for assessing nonsampling errors is provided elsewhere (see http://www.cdc.gov/nchs/nhanes.htm).

Sample Design Summary

The NHANES sample represents the total noninstitutionalized civilian population residing in the 50 states and the District of Columbia. As with previous NHANES samples, a four-stage sample design was used in NHANES 2007–2010. The first stage consisted of selecting the primary sampling units (PSUs) from a sampling frame of all U.S. counties. The PSUs in the first stage were mostly individual counties; in a few cases, adjacent counties were combined to keep PSUs above a certain minimum size. NHANES PSUs were selected with probabilities proportionate to a measure of size (PPS).

The second stage of selection for the NHANES 2007–2010 sample included a sample of area segments, comprising census blocks or combinations of blocks. However, because these samples were based on U.S. Census 2000 data, the measure of size (MOS) used for sampling was updated, if necessary, for PSUs that had experienced large growth since 2000.

The sample was designed to produce approximately equal sample sizes per PSU. "Certainty PSUs"—those selected with certainty (with a probability of one)—may have more than or fewer than 24 segments, to ensure appropriate representation in the sample. "Noncertainty PSUs" have 24 segments. Additionally, some large certainty PSUs were treated as multiple study locations with varying numbers of segments in each location, again to ensure appropriate representation of the PSU. The segments were also selected with PPS. The MOS of the segments, when combined with the subsampling rates used within the segments, provided approximately equal numbers of sampled participants per segment.

The third stage of sample selection consisted of dwelling units (DUs), including noninstitutional group quarters. In a given PSU, following the selection of segments, a listing of all DUs in the sampled segments was prepared, and a subsample of these was designated for screening to identify potential sampled participants. The screening rate was designed to produce the desired number of sampled participants for the most difficult race-Hispanic origin-income-sex-age domain (i.e., the domain sampled at the highest rate).

The fourth stage of sample selection consisted of persons within occupied DUs, or "households." All eligible members within a household were listed, and a subsample of individuals was selected based on domains defined by race and Hispanic origin, income, sex, and age. The subsampling rates and designation of potential sampled participants within screened households were arranged to provide approximately self-weighting samples for each subdomain while maximizing the average number of sampled participants per sample household.

The set of domains for which specified reliability was desired in NHANES 2007–2010 consisted of sex-age groups for non-Hispanic black persons and Hispanic persons, and income-sex-age groups for the remainder of the U.S. population. The Table on page 3 provides the set of sampling domains in NHANES 2007–2010. To increase the precision of estimates for certain subdomains, oversampling was carried out for non-Hispanic white persons and others aged 80 and over and for Hispanic persons, non-Hispanic black persons, and all other persons at or below 130% of the federal poverty level. Although data are released in 2-year cycles, at least 4 years of data must be accumulated to obtain an acceptable level of reliability for the domains given in the table. Thus, to create estimates for smaller, 2-year samples (or any annual estimates), some of the domains must be collapsed to produce adequate sample sizes for analysis.

Oversampling of population subgroups was changed in the sample design for the 2007–2010 survey period, compared with the 1999–2006 NHANES. Adolescents were no longer oversampled, and the supplemental sample of pregnant women was eliminated. The age domains 12–15 and 16–19 years were combined, and the resulting domain sample size was reduced. For the non-Hispanic black and Hispanic domains, the 40–59 age group was split in half into the 10-year age domains 40–49 and 50–59, while the total sample size for 40–59 stayed the same. This led to an increase in the number of participants aged 50–59 and a decrease in those aged 12–19 when compared with domains in previous cycles.

The most significant change is that all Hispanic persons were targeted for oversampling, rather than just Mexican-American persons. In addition to allowing estimates for the total group of Hispanic persons, sampling fractions were set so that the sample size for

Table. Sampling subdomains classified by race and Hispanic origin, income, sex, and age: National Health and Nutrition Examination Survey, 2007–2010

Non-Hispanic black	Hispanic	Non-Hispanic white and other: Non-low-income	Non-Hispanic white and other: Low-income[1]
		Males and females, age (years)	
Under 1	Under 1	Under 1	Under 1
1–2	1–2	1–2	1–2
3–5	3–5	3–5	3–5
		Males, age (years)	
6–11	6–11	6–11	6–11
12–19	12–19	12–19	12–19
20–39	20–39	20–29	20–29
. . .	. . .	30–39	30–39
40–49	40–49	40–49	40–49
50–59	50–59	50–59	50–59
60 and over	60 and over	60–69	60–69
. . .	. . .	70–79	70–79
. . .	. . .	80 and over	80 and over
		Females, age (years)	
6–11	6–11	6–11	6–11
12–19	12–19	12–19	12–19
20–39	20–39	20–29	20–29
. . .	. . .	30–39	30–39
40–49	40–49	40–49	40–49
50–59	50–59	50–59	50–59
60 and over	60 and over	60–69	60–69
. . .	. . .	70–79	70–79
. . .	. . .	80 and over	80 and over

. . . Category not applicable.
[1]Persons at or below 130% of the federal poverty level.

Mexican-American persons would not be very different from previous years and would be sufficient to produce reliable estimates for this group. The methodology for the oversampling of Hispanic persons does not provide sufficient sample sizes for calculating estimates for other Hispanic subgroups.

The overall selection probability for a person in race-Hispanic origin-income-sex-age subdomain k is

$$P_h\, P_{hj} \frac{\max_k\{r_k\}}{P_h\, P_{hj}} \frac{r_k}{\max_k\{r_k\}} = r_k \qquad [1]$$

where

h = the PSU

hj = segment within the hth PSU

k = race-Hispanic origin-income-sex-age subdomain

P_h = probability of selecting the hth PSU

P_{hj} = probability of selecting segment j within the hth PSU

r_k = sampling rate of persons in the kth race-Hispanic origin-income-sex-age subdomain

$\max_k\{r_k\}$ = maximum sampling rate across all k race-Hispanic origin-income-sex-age subdomains.

Weighting the Sample Data

The goal of NHANES is to produce data representative of the civilian noninstitutionalized U.S. population. The weighting of sample data permits analysts to produce estimates of the statistics that would have been obtained if the entire sampling frame had been surveyed. Sample weights can be considered as measures of the number of persons represented by the particular sampled participant. Weighting takes into account several features of the survey: the differential probabilities of selection for the sampling domains, nonresponse to survey instruments, and differences between the final sample and the total population.

The NHANES samples were weighted to meet the following objectives:

1. Compensate for differential probabilities of selection among subgroups defined by race and Hispanic origin, income, sex, and age.
2. Reduce biases arising from the fact that nonrespondents may be different from respondents.
3. Fix weighted sample data to match an independent U.S. Census Bureau estimate of the target population totals.
4. Compensate, to the extent possible, for inadequacies in the sampling frame (e.g., resulting from omission of some housing units in the listing of area segments and omission of persons with no fixed address).
5. Reduce variances in the estimation procedure by using auxiliary information that is known with a high degree of accuracy (Table I in Appendix II).

The sample weighting was carried out in three steps. The first step involved the computation of weights to compensate for unequal probabilities of selection (objective 1). The second step adjusted for nonresponse (objective 2). In the third step, the sample weights were poststratified to Census Bureau estimates of the U.S. population, to simultaneously accomplish objectives 3–5. These steps were performed for respondents at each stage of the survey: the screener, personal interview, and examination.

The weights described in the ''Calculating Base Weights'' section were the starting point for the screener weight calculation. Those weights were adjusted for nonresponse to the screener, and then poststratified. The resulting weights were the starting point for the calculation of the interview weights, which were then adjusted for nonresponse to the interview, inspected for extreme weights, and again poststratified. Finally, the poststratified interview weights were the starting point for calculation of the examination weights. The weights were adjusted for nonresponse to the examination, inspected for extreme weights, and then poststratified. Table II in Appendix II lists the appropriate use of each set of weights calculated for NHANES 2007–2010, including those calculated for examination and laboratory subsamples.

Note that extreme variability in the weights results in reduced reliability (increased sampling error) of some survey estimates. The NHANES sample was designed to minimize variability in the weights, subject to operational and analytic constraints. Additionally, measures such as weight-trimming were implemented to further reduce variability. The impact of weight variability is minimal when estimates are for the demographic subdomains used in the design; however, when estimates are for domains that are aggregated across design domains (for example an estimate for the total population), then the impact of weight variability is greater.

Calculating Base Weights

The first-stage (or base) weight for each sampled participant, $w_{i(\text{base})}$, was calculated as the reciprocal of the sampled participant's probability of selection. These sampling rates are provided in Table III (in Appendix II), and their derivation is described in the Sample Design Report (1).

The base weight for a sampled participant is simply the reciprocal of the sampling rate for the sampling domain of the sampled participant (denoted r_k). For NHANES 2007–2010, the base weight was adjusted further to account for the following:

- Proportion of DUs released, $f_{i(\text{release})}$
- Proportion of deselected DUs, $f_{i(\text{desel})}$
- Increase in sample size due to larger segment sizes (2008–2010), $f_{i(\text{sizeinc})}$
- Number of years in the sample, $f_{i(\text{year})}$.

The final base weight was calculated as

$$w_{i(\text{base})} = \frac{1}{r_k}\left(f_{i(\text{release})} f_{i(\text{desel})} f_{i(\text{sizeinc})} f_{i(\text{year})}\right) \quad [2]$$

where subscript i indicates the sampled participant. The following sections briefly describe each component of this calculation.

Adjustment for number of sampled DUs released to the field

The first component, the release factor $f_{i(\text{release})}$, was introduced to reflect the procedures used to obtain a relatively fixed sample size within each study location in NHANES. The sampled participant base weight was adjusted according to the proportion of the total sample released to the field. $f_{i(\text{release})}$ was calculated as

$$f_{i(\text{release})} = \frac{1}{D_i} \quad [3]$$

where D_i represents the proportion of sampled DUs released for screening in the location where sampled participant i was selected. If response rates were close to the predicted values and the MOSs used during sampling were current, the subsample factor would be approximately 1.5 for study locations fielded in 2007 to mid-2008 and 1.8 for locations fielded afterward, given that the overall DU sample size was increased. That is, approximately two-thirds of the sampled cases were expected to be released for the earlier locations, and approximately five-ninths for later study locations. See the section "Adjustment for increase in segment size (2008–2010)" for more details.

Deselection of released DUs

The sample yield monitoring and evaluation methods used in NHANES III, and subsequently in NHANES 2007–2010, occasionally suggested that the expected number of sampled participants from released DUs would exceed the target sample size for the study location. In these instances, DUs from the set of DUs released but not yet screened, to keep the sample size near the target. To account for the deselection, an adjustment factor was applied to the base weight of sampled participants who were identified in the remaining units. In 2007–2010, the expected number of sampled participants exceeded the manageable sample size in two study locations. The deselection factors for those locations were 1.82 and 2.93. The factor, denoted by $f_{i(\text{desel})}$, was calculated as

$$f_{i(\text{desel})} = \frac{1}{(1 - D_i)} \quad [4]$$

The denominator $(1 - D_i)$ represents the proportion of released DUs deselected from the sample. The deselection factor for all remaining study locations was set equal to 1.

Adjustment for increase in segment size (2008–2010)

For study locations in 2008–2010, the segment sizes and probability of DU selection were increased so that it would be easier to reach the target number of identified sampled participants. To make this adjustment, the overall selection probability for a sampled participant changed to

$$P_h P_{hj} \frac{(1.2)\max_k\{r_k\}}{P_h P_{hj}} \frac{r_k}{\max_k\{r_k\}} = 1.2\, r_k \quad [5]$$

Thus, a factor for these study locations, denoted by $f_{i(\text{sizeinc})}$, was calculated as

$$f_{i(\text{sizeinc})} = \frac{1}{1.2} \quad [6]$$

Note that one study location was small enough that the increase in the probability of DU selection caused every DU in the county to be selected for the sample. For this location only, the valuwere deselected or randomly removed e of $f_{i(\text{sizeinc})}$ was slightly smaller.

Adjustment for number of years in the sample

Because the selected sample of study locations on which the sampling rates were based was fielded over 4 years, the base weights calculated from the original sampling rates also correspond to a 4-year sample. In weighting subsets of those 4 years, the following factor must be applied:

$$f_{i(\text{year})} = \frac{\text{AWF}_i}{\text{Number of years in sample}} \quad [7]$$

where AWF_i represents the factor that, when applied to the weights, converts the 4-year weights to annual weights. For the 2007–2010 sample, AWF_i is 4. The divisor of $f_{i(\text{year})}$ is simply the number of years in the sample to be weighted. For example, the divisor for the records in the 2007–2008 sample is 2.

Nonresponse Adjustment

If every selected household had agreed to complete the screener, and every selected person had agreed to complete the interview and the examination, weighted estimates using the base weights described in the ''Calculating Base Weights'' section would be approximately unbiased estimates of characteristics for the civilian noninstitutionalized U.S. population. But in reality, some of the sampled participants who were screened refused to be interviewed (interview nonresponse), and some of the interviewed sampled participants refused the examination (examination nonresponse). Thus, nonresponse bias may result.

Bias in the survey estimates occurs when the characteristics of nonrespondents are very different from those of respondents. The best approach to minimizing nonresponse bias is to plan and implement field procedures that maintain high cooperation rates. For NHANES, the payment of cash incentives and repeated callbacks for refusal conversion are very effective in reducing nonresponse, and thus nonresponse bias. However, some nonresponse occurs even with the best strategies; therefore, adjustments are always necessary to minimize potential nonresponse bias.

A multistage procedure for nonresponse adjustment was carried out to adjust for nonresponse to the screener, interview, and examination. The nonresponse adjustment procedure consists of computing adjustment factors and applying these to the survey weights separately by nonresponse cell. Nonresponse adjustment reduces bias if response rates and survey characteristics vary from cell to cell and if respondents and nonrespondents sharing the same characteristics are in the same cell. The nonresponse adjustment factors are the reciprocals of the weighted response rates within the selected cells.

A negative effect of nonresponse adjustment is that it increases the variability of the weights, which in turn increases the variance. When the nonresponse cells contain a sufficient number of 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 the small number of respondents in that cell. To avoid having nonresponse adjustments based on very small sample sizes, or having large nonresponse adjustment factors, cells are usually collapsed to form larger cells. The following criteria were used in NHANES to determine whether to collapse cells: (a) a minimum of 30 respondents in each cell and (b) a maximum adjustment factor of 1.35.

Nonresponse adjustments were carried out separately for screener nonresponse, interview nonresponse, and examination nonresponse. In general, nonresponse adjustment cells were generated using variables with known values for both respondents and nonrespondents. A few variables with low item nonresponse rates were considered when creating nonresponse adjustment cells. For the screener nonresponse adjustment, cells were defined by segments within each location. For the interview and examination nonresponse adjustments, the Chi-squared Automatic Interaction Detector (CHAID) was used to identify variables most highly related to response propensity. See Tables I and IV in Appendix II for the variables used to form the nonresponse adjustment cells.

The nonresponse adjustment factors, $f_{i(\text{NR})}$, were calculated as

$$f_{i(\text{NR})} = \frac{\sum_{i=1}^{n_{a_s}} w_{i(\text{base})}}{\sum_{i=1}^{n_{a_r}} w_{i(\text{base})}} \quad [8]$$

where $w_{i(\text{base})}$ is the base weight for the ith sampled participant in the ath cell, n_{a_s} is the total sample size in the ath nonresponse adjustment cell, and n_{a_r} is the number of respondents in the ath cell. The summation was carried out separately for each cell. Thus, the nonresponse-adjusted weights, $w_{i(\text{NR})}$, were calculated as

$$w_{i(\text{NR})} = w_{i(\text{base})} f_{i(\text{NR})} \quad [9]$$

Trimming

Because nonresponse adjustments can contribute to extreme weights, trimming of the weights was considered. Extreme weights may also occur when units are sampled to yield fixed sample sizes within a PSU, as was the case in NHANES. Even a few unexpectedly large weights can seriously inflate the variance of survey estimates. Thus, weight-trimming procedures may be used to reduce the impact of any such large sampled participant weights on the estimates produced from the sample. Because trimming introduces a bias in the estimates, it is hoped that the resulting reduction in variances will also decrease the mean squared error.

The inspection method was used for trimming weights in NHANES. This method involves inspecting the distribution of weights in the sample and applies to samples (or subsets of samples) that were originally designed to be self-weighting.

The subdomains for trimming are the race-Hispanic origin-income-sex-age sampling domains. Because some groups are oversampled, the weights within domains may be quite variable. For this reason, trimming thresholds were dependent on the amount of oversampling used in these domains.

Once the weights to be trimmed had been identified, the weights of the nontrimmed cases were also adjusted so that the weights for each sampling domain summed to the corresponding weighted sum prior to trimming. This is referred to as ''preserving weighted totals.'' Failure to preserve weighted totals may lead to serious understatements in estimated totals; thus, preserving weighted totals is an important characteristic for a trimming procedure.

The trimming factors, $f_{i(\mathrm{TR})}$, were calculated as

$$f_{i(\mathrm{TR})} = \frac{\sum_{i=1}^{n_b} t_i}{\sum_{i=1}^{n_b} w_{i(\mathrm{base})} f_{i(NR)}} \qquad [10]$$

where n_b is the sample size of the bth race-Hispanic origin-income-sex-age sampling domain, and t_i is equal to $w_{i(\mathrm{base})} f_{i(\mathrm{NR})}$, provided that this product does not exceed the threshold and is set to be equal to the threshold otherwise. The trimmed weights, $w_{i(\mathrm{TR})}$, were calculated as

$$w_{i(\mathrm{TR})} = w_{i(\mathrm{NR})} f_{i(\mathrm{TR})} \qquad [11]$$

Poststratification

The final step in the weighting procedure was poststratification to known population totals, to compensate for undercoverage or overcoverage of certain demographic groups and for any residual differential nonresponse among these groups. Poststratification of sample weights to independent population estimates has several purposes. In most household surveys, certain demographic groups in the U.S. population (e.g., young black males) experience fairly high rates of undercoverage in survey efforts. Poststratification to census estimates partially compensates for such undercoverage and for any differential nonresponse, and can help to reduce the resulting bias in the survey estimates. Poststratification can also help reduce the variability of sample estimates and achieve consistency with accepted U.S. figures for various subpopulations.

Poststratification involves applying a ratio adjustment to the survey weights. Broad classes—called poststratification cells, or poststrata—are constructed using auxiliary data, and a single ratio-adjustment factor is applied to all units in a given poststratification cell. The numerator of the ratio is a ‘‘control total’’ obtained from a secondary source; the denominator is a weighted total obtained using the survey weights. Therefore, at the poststratum level, estimates obtained using the poststratified survey weights will correspond to the control totals used. Because poststratification is a ratio adjustment, this process will improve the efficiency of estimates provided that the variables used in constructing poststratification cells are associated with the analysis variables of interest. Such gains in efficiency are most evident in the case of linear estimates such as means or totals; for ratio estimates, the ratio adjustments cancel each other out at the poststratum level, and the overall gains in efficiency due to poststratification tend to be small.

A major effect of poststratification is that it implicitly imputes for nonresponse of survey characteristics for the missed persons. The assumption is that these missed persons not covered by the survey have the same distribution of characteristics as interviewed persons within the poststratification cells. This is obviously an oversimplification; the missed persons are likely to be different. However, in the absence of any detailed information on the characteristics of the missed persons, poststratification appears to be the only reasonable technique available for reducing bias due to undercoverage and nonresponse.

All control totals were obtained using undercount-adjusted weights from the yearly March supplement to the U.S. Census Bureau’s Current Population Survey (CPS). For the 2-year samples, the control totals were calculated as the midpoint between the CPS estimates of the poststrata population totals for the two sample years. These CPS weights have undergone poststratification to the Census Bureau’s best estimates of the total noninstitutionalized civilian population of the United States, including those not counted in surveys or in the most recent decennial census. Poststratification, therefore, brings the weighted totals up to the level of the presumed total noninstitutionalized civilian population in the United States.

The poststratification factors, $f_{i(\mathrm{PS})}$, were calculated as

$$f_{i(\mathrm{PS})} = \frac{N_c}{\sum_{i=1}^{n_c} w_{i(\mathrm{TR})}} \qquad [12]$$

where N_c is the control total and n_c is the sample size of the poststratification cell. Thus, the poststratified weights, $w_{i(\mathrm{PS})}$, were calculated as

$$w_{i(\mathrm{PS})} = w_{i(\mathrm{NR})} f_{i(\mathrm{PS})} \qquad [13]$$

Computing Final Weights

The final weight for each sampled participant at each stage was calculated as the product of the base weight and the nonresponse adjustment, trimming, and poststratification factors; that is,

$$w_i = w_{i(\mathrm{base})} f_{i(\mathrm{NR})} f_{i(\mathrm{TR})} f_{i(\mathrm{PS})} \qquad [14]$$

More specifically, the final screening weight was calculated as

$$w_{i(\mathrm{S})} = w_{i(\mathrm{base})} f_{i(\mathrm{NR,S})} f_{i(\mathrm{TR,S})} f_{i(\mathrm{PS,S})} \qquad [15]$$

the final interview weight was calculated as

$$w_{i(\mathrm{I})} = w_{i(\mathrm{base})} f_{i(\mathrm{NR,S})} f_{i(\mathrm{TR,S})} f_{i(\mathrm{PS,S})} f_{i(\mathrm{NR,I})} f_{i(\mathrm{TR,I})} f_{i(\mathrm{PS,I})} \qquad [16]$$

and the final examination weight was calculated as

$$w_{i(\mathrm{E})} = w_{i(\mathrm{base})} f_{i(\mathrm{NR,S})} f_{i(\mathrm{TR,S})} f_{i(\mathrm{PS,S})} f_{i(\mathrm{NR,I})} f_{i(\mathrm{TR,I})} f_{i(\mathrm{PS,I})} f_{i(\mathrm{NR,E})} f_{i(\mathrm{TR,E})} f_{i(\mathrm{PS,E})} \qquad [17]$$

Only the interview and examination weights were released to the public.

Any sampled participant who did not respond to the interview was assigned an interview weight of zero. These sampled participants were considered ineligible for the examination and were also assigned an examination weight of zero. Their records were not released to the public. Sampled participants who did complete the interview and were eligible for the examination but did not respond were assigned examination weights of zero and their records were included in the public release.

The interview weight should be used for analyses of data from the household interview only. The examination weights should be used exclusively for analyses of data from the examination, or in conjunction with the household interview data. This includes data from the MEC interview and examination, or laboratory data on the entire examination sample (Table II).

Subsample Weights

Some NHANES respondents were also asked to participate in survey components that were statistically defined (or random) subsamples of the NHANES examined sample. Data collected from these participants include a variety of laboratory, nutrition and dietary, environmental, audiometry, and mental health components. Each subsample was selected to be a nationally representative sample.

For example, some but not all participants were selected to give a fasting blood sample on the morning of their examination. The subsamples selected for these components were chosen at random with a specified sampling fraction (e.g., one-half of the total examined group), according to the protocol for that component. Each component subsample has its own designated weight, which accounts for the additional probability of selection into the subsample component, as well as any additional nonresponse to the component. For some components, subsample weights were calculated to incorporate additional information relevant to data collection (such as day of the week for the dietary recall data).

The special survey weights were included in the released data file. Because these weights differ from the examination weights for the whole sample, the subsample-specific weights must be used for statistical estimation of measures collected only in that subsample. (See Table II for a list of these weights and information regarding their appropriate use.) For example, no sample weights are provided for the overlap of the fasting subsample with an environmental subsample. Subsample weights from the same survey cycle are not designed to be combined within the data release cycle. In fact, many subsamples are mutually exclusive. To combine two or more subsamples, random overlap would have to occur between the subsamples, and appropriate weights would need to be recalculated. See the respective survey protocol and documentation (available from: http://www.cdc.gov/nchs/nhanes.htm) for more specific information on each subsample.

Combining 2-year Weights to Analyze Other Multiyear Samples

Only 2-year weights were calculated for NHANES 2007–2008 and 2009–2010. To combine these cycles to produce 4-year estimates, the 4-year weights are calculated as one-half times the 2-year weights. Six-year weights may similarly be created for 2005–2010 estimates by multiplying the 2-year weights by one-third. Future years of data can continue to be added, using these same methods for each sample. However, as a result of sample design changes initiated in 2007, it is important to consider the following when calculating estimates based on data from the 6-year period 2005–2010. First, because of the small sample numbers and design inefficiency for non-Mexican-American Hispanic persons (and thus for total Hispanic persons) in 2005–2006, estimates should not be created for total Hispanic persons for the 2005–2010 data period. Second, for the non-Hispanic white, non-Hispanic black, and Mexican-American sample domains, rescaling the sample weights to create 6-year weights (as described in the NHANES analytic guidelines) will be sufficient for producing combined 6-year estimates for most analytic variables.

With this reweighting, the target population is the U.S. noninstitutionalized population at the midpoint of the combined interval, and the sum of combined weights should be reasonably close to an independent estimate of that midpoint population. Although combining years of data is useful for rare events, users are cautioned that there is an inherent assumption of no trend in the estimate over the time period, or an interpretation that the estimate is the average over the time period.

Variance Estimation

Sampling errors should be calculated for all survey estimates, to aid in determining the statistical reliability of those estimates. For complex sample surveys, exact mathematical formulas for variance estimates are usually not available. Variance approximation procedures are required in order to provide reasonable, approximately unbiased, and design-consistent estimates of variance. Although each 2-year sample is nationally representative, it was selected from only 30 PSUs, and the sample sizes for some specific race-Hispanic origin-income-sex-age subdomains may be small.

The small number of PSUs also poses challenges for variance estimation. With a small number of PSUs, direct design-based variance estimates may be unstable for some measures. In addition, because variance computations must incorporate the NHANES design, standard statistical software routines (i.e., software packages that assume a simple random sample) should not be used for computing variances for NHANES. This section introduces design-based methods of variance estimation for complex sample survey data and describes the creation of variables necessary for variance estimation on the public- and restricted-use data files for the 2007–2010 samples.

Two variance approximation procedures that account for the complex sample design and allow the computation of design effects are replication methods and Taylor Series Linearization. Replication methods provide a general means for estimating variances for the types of complex sample designs and weighting procedures usually encountered in practice. The basic idea behind the replication approach is to select subsamples repeatedly from the whole sample, to calculate the statistic of interest for each of these subsamples (or ‘‘replicates’’), and then to use the variability among these replicate statistics to estimate the variance of the full-sample statistic. The jackknife and balanced repeated replication (BRR) methods are two common procedures for deriving replicates from a full sample. The jackknife procedure retains most of the sample in each replicate, whereas the BRR approach retains a portion of the sample in each replicate.

BRR was used for NHANES III. Initially, the delete-one jackknife method (a replication method) was used to estimate variances based on data from the NHANES 1999–2000 survey. However, jackknife method replicate weights were provided only for the 1999–2000 data release. If replication methods are to be used for any other survey years, replicate weights must be computed by the analyst (2).

For the linearization approach, nonlinear estimates are approximated by linear estimates for the purpose of variance estimation. The linear approximation is derived by taking the first-order Taylor series approximation for the estimator. Standard variance estimation methods for linear statistics are then used to estimate the variance of the linearized estimator. Currently, NCHS recommends using Taylor Series Linearization methods for variance estimation in analyses of all NHANES data. SUDAAN, Stata, R, and SAS survey procedures can be used to obtain variance estimated by this method.

Variance Estimation for Publicly Released Data

PSUs are selected from strata defined by geography and proportions of minority populations. In any 2-year sample there are two PSUs from each strata; these are used as variance strata to estimate sampling error in the Taylor Series Linearization approach. Within each variance stratum, two variance units are generally defined as the PSU. However, because 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. As a result, some of these variance strata may have one PSU split into multiple variance units, while other variance strata may have three PSUs for variance estimation, depending on the number and size of the certainty locations that year.

The small number of PSUs in a 2-year NHANES sample, geographic data and other characteristics of the area on the data files, and local publicity campaigns while the survey is in the field all pose a risk for data disclosure. As a result, masked variance units (MVUs) are provided for use with the public-use data files to reduce the chance of an intruder being able to match PSUs in the sample to PSUs in the population, while minimizing the bias in the variance caused by altering the PSU structure. MVUs can be used as if they were pseudo-PSUs to estimate sampling errors (similar to past NHANES).

The MVUs or pseudo-PSUs on the data file are not the ''true'' design PSUs but are a collection of secondary sampling units aggregated into groups for the purpose of variance estimation. They 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 2009–2010. They can also be used for analyzing any combined 4-, 6-, or 8-year data set after the appropriate adjustment described in ''Combining 2-year Weights to Analyze Other Multiyear Samples.''

Many surveys swap data values between cases to limit disclosure. Rather than swapping individual values, however, the procedure used in NHANES 2007–2010, and described in Park et al. (3), swapped entire segments (secondary sampling units) between PSUs. That is, for two similar segments in different PSUs, the PSU and variance stratum identifiers for all sampled cases were swapped. Because any PSUs with swapped segments are no longer completely associated with a single real PSU, the chance of correctly matching a given individual within the PSU is limited. The point estimates of the overall population means do not change under this PSU masking, but the variance estimates may change slightly.

To identify which segments to swap in NHANES 2007–2010, estimates were first calculated for all of the segments in all of the study locations, for comparative purposes. These estimates provide general descriptions of the segments, such as percentage of Hispanic sampled participants, prevalence of home ownership, and obesity prevalence, that should be similar for swapped segments. Then, study locations that were the most at risk for data disclosure (locations with smaller populations or in rural areas) were identified.

Within each of these at-risk locations, each segment was paired with all segments from the other study locations (including other at-risk locations) and a distance measure was calculated to determine the effect of swapping the pair on variance. The distance measure was calculated as

$$D = \sum_{l=1}^{q} \left| \frac{v(\bar{x}_l|S^r) - v(\bar{x}_l|S)}{v(\bar{x}_l|S)} \right| \qquad [18]$$

where q is the number of variables used to calculate the estimates, l is an individual estimate, $\bar{x}_l$ is the mean of that estimate, $v(\bar{x}_l \mid S^r)$ is the variance of the estimate after swapping, and $v(\bar{x}_l \mid S)$ is the variance of the estimate before swapping.

Within each at-risk location, the segments were sorted by smallest distance measure achieved, and some segments were selected to be swapped. Generally, pairs with the smallest distances were swapped, but if any two pairs included the same segment, one pair was not used for swapping. In this way, a single segment was swapped only once. Consideration was also given to pairs of segments that both came from at-risk study locations; swapping of such pairs was minimized where possible.

Further research by Park (4) indicated that variance estimates generally tend to increase as more segments are swapped, although the variance for specific analysis variables could also be underestimated after swapping. For this reason, the amount of swapping (i.e., the number of study locations determined to be at risk and the number of segments swapped per location) is limited.

Variance Estimation in the Research Data Center

For the current sample design, NHANES data are released to the public in 2-year data cycles. In addition to public-use data files, there are special data sets that are available only through the NCHS Research Data Center (RDC) (see http://www.cdc.gov/rdc/index.htm). These include data sets with (a) data items that were collected for an odd number of calendar years (1, 3, or 5 years), (b) data linked geographically to some other contextual data files (often supplied by the data user), and (c) data items that are determined to be too sensitive or too detailed to be released to the public due to confidentiality restrictions.

Some of these data files have special sample weights that should be used when these nonpublic data sets are analyzed within the confines of the RDC environment. For example, single-calendar-year data files have a single-year examination weight. This single-year weight can be combined with the examination weight provided on the 2-year public-use file to create a 3-year examination weight. All single-year sample weights were calculated in the same manner as the public-use 2-year weights described in ''Weighting the Sample Design.'' If a special data file involves subsampling, then special subsample weights were created for that file that reflect the number of calendar years in the data file and the rate of subsampling. For all special data files, appropriate documentation is provided in the RDC to describe the necessary sample weights.

Special unmasked PSU and stratum codes (which differ from the MVU codes provided for the public-use files) are provided for variance estimation for data from those special files, using the true PSU and stratum codes. These unmasked design codes are necessary given the need for true geographic linkage with some data sets. Providing the unmasked PSU and stratum codes poses no disclosure risk because of the restrictions of the RDC, such as the prohibition of publication of PSU-level estimates. Further, any subnational estimate that is generated from an RDC analysis must be reviewed and approved by NCHS staff to protect the confidentiality of sampled participants.

More information on the RDC, and lists of special NHANES data files, are available from the NCHS website at: http://www.cdc.gov/rdc/. Information on proposals for use of stored specimens is available from: http://www.cdc.gov/nchs/nhanes/genetics/stored_specimens.htm.

References

1. Curtin LR, Mohadjer LK, Dohrmann SM, et al. National Health and Nutrition Examination Survey: Sample design, 2007–2010. National Center for Health Statistics. Vital Health Stat 2(160). 2013. Available from: http://www.cdc.gov/nchs/data/series/sr_02/sr02_160.pdf.
2. Korn EL, Graubard BI. Analysis of health surveys. New York, NY: Wiley;29. 1999.
3. Park I, Dohrmann S, Montaquila J, Mohadjer LK. Curtin LR. Reducing the risk of data disclosure through area masking: Limiting biases in variance estimation. Proceedings of the Survey Research Methods Section, American Statistical Association;1761–7. 2006.
4. Park I. Primary sampling unit (PSU) masking and variance estimation in complex surveys. Surv Methodol 34(2):183–94. 2008.

Appendix I. Glossary

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 ''housing unit.'' A house, apartment, mobile home or trailer, group of rooms, or single room occupied as separate living quarters (see *Group quarters*) or, if vacant, intended for occupancy as separate living quarters. Separate living quarters are those in which the occupants live separately from other individuals in the building and which 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 (i.e., excluding institutional group quarters), or that could become eligible (e.g., vacant at the time of sampling, but could be occupied once screening begins).

Group quarters—A place where people live or stay that is normally owned or managed by an entity or organization providing housing or services for the residents. These services may include custodial or medical care, as well as other types of assistance, and residency is commonly restricted to those receiving these services. People living in group quarters usually are not related to each other. Group quarters include such places as college residence halls, residential treatment centers, skilled nursing facilities, group homes, military barracks, correctional facilities, workers' dormitories, and facilities for people experiencing homelessness. These are generally grouped into two categories: institutional group quarters and noninstitutional group quarters.

Institutional group quarters—Group quarters providing formally authorized, supervised care or custody in an institutional setting, such as correctional facilities, nursing and skilled nursing facilities, inpatient hospice facilities, mental health or psychiatric facilities, and group homes and residential treatment centers for juveniles. Institutional group quarters are not included in the NHANES sample.

Noninstitutional group quarters—Group quarters that do not provide formally authorized, supervised care or custody in an institutional setting. These include college and university housing, group homes intended for adults, residential treatment facilities for adults, workers' group living quarters, Job Corps centers, and religious group quarters. Noninstitutional group quarters are included in the NHANES sample.

Household—The person or group of persons living in an occupied DU.

Low-income—Beginning in 2000, NHANES split the sampling domains for white and other persons based on their income status into low-income or non-low-income. Low-income persons were defined as those at or below 130% of the federal poverty level. The poverty threshold used in this determination was based on the most recent poverty guidelines published by the U.S. Department of Health and Human Services; these thresholds are updated annually by the U.S. Census Bureau.

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 using the units as sampled, some pseudounits are created by swapping segments between PSUs. 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 2009–2010. They can also be used for analyzing any combined 4-, 6-, or 8-year data set.

Maximum sampling rate ($\max_k\{r_k\}$)—The largest probability of selection assigned to a demographic group within a survey design. This value within certain strata and demographic groups was used to determine the sample size and other sampling parameters in NHANES.

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, the MOS is actually a weighted average of estimates of population counts for the race-Hispanic origin-income groups of interest.

National Center for Health Statistics (NCHS)—As the nation's principal health statistics agency, NCHS designs, develops, and maintains a number of systems that produce data related to demographic and health concerns. These include data on registered births and deaths collected through the National Vital Statistics System, and data collected by the National Health Interview Survey, NHANES, the National Health Care Surveys, and the National Survey of Family Growth, among others. NCHS is part of the Centers for Disease Control and Prevention, an operating division of the U.S. Department of Health and Human Services.

Noninstitutionalized civilian population—Includes all people living in households, excluding those in institutional group quarters and those on active duty with the military. This is the target population for NHANES.

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''); all other PSUs are selected without certainty (''noncertainty PSUs'').

Probability proportionate to size (PPS) sampling—In this method, the probability of selecting any unit varies with the size of the unit, giving larger units a greater probability of selection and smaller units a lower probability. NHANES uses PPS sampling in the selection of PSUs and secondary sampling units.

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 within NCHS to ensure that respondent identities are protected. NCHS disseminates these files to encourage public use of the survey data.

Race and Hispanic origin—This phrase is used in this report as it was used in NHANES sample selection. It refers to Hispanic persons, non-Hispanic black persons, and a third group consisting of all other persons.

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 balanced repeated replication (BRR) methods are two common procedures for the derivation of replicates from a full sample. The BRR method was used in the creation of replicate weights for most of the NHANES 2007–2010 multiyear samples.

Respondent—A person selected into a sample who agrees to participate in all aspects of a survey. In NHANES, persons agreeing to complete the in-home interview are considered ''interview respondents.'' Persons agreeing to complete both the in-home interview and an examination in a mobile examination center (MEC) are considered ''MEC respondents.''

Response rate—The number of survey respondents divided by the number of persons selected into the sample. The response rates in this report are MEC response rates—calculated as the number of people receiving examinations in the MEC divided by the total number of people sampled.

Restricted-use data file—An electronic data set of survey respondent records, containing some information that may, if released to the public, risk disclosing individual survey respondents. These data are available only through the NCHS Research Data Center. These include data sets with (a) data items collected for an odd number of calendar years (1, 3, or 5 years); (b) data geographically linked to other contextual data files (often supplied by the data user); (c) data items determined to be too sensitive or too detailed to be released to the public due to confidentiality restrictions; and (d) surplus sera projects where past biological samples have been stored and subsequently used based on a formal proposal submitted as a special study (on either the full sample or a special subsample).

Sample weight—The estimated number of persons 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-income-age subdomain, 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 persons in the sample, so that accurate national estimates could be made from the sample. The product of all of these adjustments is sometimes called the ''final'' sample weight.

Sampling domain—NHANES 2007–2010 included 72 sampling domains. See *Domain*.

Sampling rate—The rate at which a unit is selected from a sampling frame. For NHANES, the rates required for sampling persons 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 eligible for the survey. In 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. In NHANES, only persons aged 18 and over can answer the screener.

Screening—The process of conducting, or attempting to conduct, the screener interview in selected DUs. Occupied DUs (households) are screened using the screener. Other units can also be screened; the process for these units is simply 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 groups of blocks, and their selection makes up the second stage of sampling. Within each segment, a sample of DUs was selected.

Self-weighting sample—A sample for which each elementary unit in the population has the same, nonzero chance of selection into the sample; that is, they are selected with the same constant probability. Higher-stage sampling units may be selected with differing probabilities, but such differences in selection probabilities at various stages cancel each other out. NHANES is a self-weighting sample of persons within each sampling domain.

Stratification; Strata—The partitioning 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 in 2007–2010, PSUs were stratified based on region, 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 persons in supervised care or custody in institutional settings, all active-duty military personnel, active-duty family members living overseas, and any other persons residing 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 having taken 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*.

Appendix II. Supporting Tables

Table I. Variables used in forming nonresponse adjustment cells for weighting interview samples: National Health and Nutrition Examination Survey, 2007–2010

Variables considered for nonresponse, by age group	Order and categories of variables cross-classified to form nonresponse adjustment cells	
	2007–2008	2009–2010
0–5 years		
Race and Hispanic origin of sampled person	1. Non-Hispanic black, Hispanic, other	. . .
Urbanicity	2. Areas with population over 3 million, other	. . .
Race and ethnicity of household reference person	. . .	4. Non-Hispanic black, Hispanic, other
Sex of household reference person	3. Male, female	5. Male, female
Household size	4. 1–2, 3–4, 5–6, 7 or more	1. 1–4, 5–8, 9 or more
Census region	5. Northeast, Midwest, South, West	2. Northeast, Midwest, South, West
Number of sampled persons in household	. . .	3. 1–3, 4, 5 or more
6–19 years		
Sex of household reference person	1. Male, female	. . .
Number of sampled persons in household	2. 1–5, 6 or more	. . .
Household size	. . .	1. 1–4, 5–8, 9 or more
Census region	3. Northeast, Midwest, South, West	3. Northeast, Midwest, South, West
Urbanicity	4. Urban, suburban, rural	. . .
Presence of a child under 18 in household	. . .	2. Yes, no
Race and Hispanic origin of sampled person	. . .	4. Non-Hispanic black, Hispanic, other
Sex of sampled person	. . .	5. Male, female
20–59 years		
Household size	1. 1–2, 3–4, 5–6, 7 or more	1. 1–2, 3–4, 5–6, 7 or more
Urbanicity	2. Urban, suburban, rural	4. Areas with population over 3 million, other
Census region	3. Northeast, Midwest, South, West	2. Northeast, Midwest, South, West
Race and Hispanic origin of sampled person	4. Non-Hispanic black, Hispanic, other	5. Non-Hispanic black, Hispanic, white, other
Presence of a child under 18 in household	5. Yes, no	. . .
Presence of a child under 6 in household	. . .	3. Yes, no
60 years and over		
Urbanicity	1. Areas with population over 3 million, other urban areas, suburban, rural	2. Urban, suburban, rural
Census region	2. Northeast, Midwest, South, West	4. Northeast, Midwest, South, West
Household size	3. 1–2, 3–4, 5–6, 7 or more	. . .
Presence of a child under 18 in household	. . .	1. Yes, no
Number of sampled persons in household	. . .	3. 1, 2 or more

. . . Category not applicable.

Table II. Most common survey sample weights and their appropriate use: National Health and Nutrition Examination Survey, 2007–2010

Weight	Application
Interview	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).
Examination	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 (e.g., fasting or environmental).
Dietary day 1 sample	Use when analyzing data from the day 1 24-hour dietary recall or the Flexible Consumer Behavior Survey telephone follow-up module for examined persons who completed one or both of these interviews.
Dietary day 2 sample	Use when analyzing data from the day 1 and day 2 24-hour dietary recalls for examined persons who completed these interviews.
Fasting subsample	Use when analyzing the fasting glucose, insulin levels, triglycerides, and LDL cholesterol (lipids) only for examined persons assigned to and meeting the criteria for the fasting subsample.
Oral glucose tolerance test (OGTT) subsample	Use when analyzing only OGTT glucose levels or OGTT glucose levels with other data such as insulin or fasting levels for examined persons assigned to and meeting the criteria for the OGTT.
Volatile organic compound subsample	Use when analyzing data from volatile organic compound one-half laboratory subsample for examined persons assigned to and meeting the criteria for this subsample.
Laboratory subsample A	Use when analyzing data from the one-third laboratory environmental subsample A for examined persons assigned to and meeting the criteria for this subsample.
Laboratory subsample B	Use when analyzing data from the one-third laboratory environmental subsample B for examined persons assigned to and meeting the criteria for this subsample.
Laboratory subsample C	Use when analyzing data from the one-third laboratory environmental subsample C for examined persons assigned to and meeting the criteria for this subsample.

NOTE: LDL is low-density lipoprotein.

Table III. Final sampling rates and base weights: National Health and Nutrition Examination Survey, 2007–2010

Race and Hispanic origin-income-sex-age sampling domain[2]	2007–2008		2009–2010[1]	
	Numerator of sampling rate[3]	Base weight	Numerator of sampling rate[3]	Base weight
Non-Hispanic black				
Male and female:				
Under 1	0.74	2,197.71	0.74	2,197.71
1–2	0.67	2,423.66	0.67	2,423.66
3–5	0.50	3,265.82	0.50	3,265.82
Male:				
6–11	0.52	3,128.20	0.52	3,128.20
12–19	0.39	4,173.43	0.39	4,173.43
20–39	0.27	5,888.48	0.27	5,888.48
40–49	0.30	5,449.23	0.30	5,449.23
50–59	0.36	4,547.44	0.36	4,547.44
60 and over	0.91	1,786.37	0.91	1,786.37
Female:				
6–11	0.54	3,007.61	0.54	3,007.61
12–19	0.36	4,550.57	0.36	4,550.57
20–39	0.22	7,215.89	0.22	7,215.89
40–49	0.24	6,860.13	0.24	6,860.13
50–59	0.28	5,729.55	0.28	5,729.55
60 and over	0.61	2,641.55	0.61	2,641.55
Hispanic				
Male and female:				
Under 1	1.00	1,617.80	1.00	1,617.80
1–2	0.61	2,667.76	0.61	2,667.76
3–5	0.41	3,936.02	0.41	3,936.02
Male:				
6–11	0.46	3,521.77	0.46	3,521.77
12–19	0.35	4,631.85	0.35	4,631.85
20–39	0.22	7,425.66	0.22	7,425.66
40–49	0.29	5,501.76	0.29	5,501.76
50–59	0.48	3,404.92	0.48	3,404.92
60 and over	1.00	1,617.80	1.00	1,617.80
Female:				
6–11	0.47	3,457.27	0.47	3,457.27
12–19	0.36	4,513.21	0.36	4,513.21
20–39	0.24	6,845.99	0.24	6,845.99
40–49	0.29	5,515.27	0.29	5,515.27
50–59	0.44	3,702.10	0.44	3,702.10
60 and over	0.95	1,706.07	0.95	1,706.07
Non-Hispanic white or other, low-income				
Male and female:				
Under 1	1.00	1,617.80	1.00	1,617.80
1–2	0.65	2,473.22	0.68	2,384.89
3–5	0.46	3,534.90	0.49	3,291.11
Male:				
6–11	0.28	5,850.60	0.34	4,786.85
12–19	0.19	8,298.50	0.29	5,601.49
20–29	0.19	8,564.47	0.27	6,034.06
30–39	0.33	4,896.73	0.38	4,216.63
40–49	0.26	6,310.83	0.31	5,148.31
50–59	0.31	5,161.38	0.36	4,444.52
60–69	0.37	4,377.29	0.41	3,991.06
70–79	0.68	2,374.67	0.70	2,300.46
80 and over	0.73	2,228.42	0.73	2,228.42

See footnotes at end of table.

Table III. Final sampling rates and base weights: National Health and Nutrition Examination Survey, 2007–2010—Con.

Race and Hispanic origin-income-sex-age sampling domain[2]	2007–2008		2009–2010[1]	
	Numerator of sampling rate[3]	Base weight	Numerator of sampling rate[3]	Base weight
Non-Hispanic white or other, low-income—Con.				
Female:				
6–11	0.27	6,033.45	0.33	4,936.46
12–19	0.19	8,592.04	0.28	5,799.63
20–29	0.14	11,981.17	0.24	6,632.44
30–39	0.20	7,894.48	0.30	5,438.42
40–49	0.23	7,032.79	0.31	5,190.87
50–59	0.25	6,530.81	0.30	5,327.77
60–69	0.27	5,902.28	0.33	4,945.16
70–79	0.33	4,961.80	0.38	4,272.66
80 and over	0.38	4,310.49	0.40	4,049.25
Non-Hispanic white or other, non-low-income				
Male and female:				
Under 1	0.37	4,346.57	0.37	4,346.57
1–2	0.20	8,108.34	0.20	8,108.34
3–5	0.14	11,763.13	0.14	11,763.13
Male:				
6–11	0.14	11,263.77	0.14	11,263.77
12–19	0.09	17,699.38	0.09	17,699.38
20–29	0.09	17,977.79	0.09	17,977.79
30–39	0.10	15,962.95	0.10	15,962.95
40–49	0.08	20,314.70	0.08	20,314.70
50–59	0.09	18,720.29	0.09	18,720.29
60–69	0.13	12,555.30	0.13	12,555.30
70–79	0.21	7,534.69	0.21	7,534.69
80 and over	0.40	4,056.66	0.40	4,056.66
Female:				
6–11	0.14	11,690.30	0.14	11,690.30
12–19	0.09	17,930.75	0.09	17,930.75
20–29	0.09	18,359.86	0.09	18,359.86
30–39	0.09	17,654.24	0.09	17,654.24
40–49	0.07	21,643.66	0.07	21,643.66
50–59	0.07	22,132.77	0.07	22,132.77
60–69	0.11	14,603.62	0.11	14,603.62
70–79	0.20	8,063.03	0.20	8,063.03
80 and over	0.30	5,447.32	0.30	5,447.32

[1]Sample includes an additional target of 150 examinations of low-income non-Hispanic white persons or others older than 1 year, altering their sampling rates and base weights. All other rates and base weights were unchanged.

[2]Age in years.

[3]Corresponds to a 150% sample; sampling rates may be calculated by dividing the numerator by 1,618.

Table IV. Variables used in forming nonresponse adjustment cells for weighting MEC examination samples: National Health and Nutrition Examination Survey, 2007–2010

Variables considered for nonresponse, by age group	Order and categories of variables cross-classified to form nonresponse adjustment cells	
	2007–2008	2009–2010
0–5 years		
Sex of sampled person	1. Male, female	. . .
Urbanicity	2. Urban, suburban, rural	. . .
Household size	3. 1–2, 3–4, 5–6, 7 or more	2. 1–2, 3–6, 7 or more
Census region	4. Northeast, Midwest, South, West	. . .
Reported health status of sampled person	. . .	1. Excellent/very good/good/not known, fair/poor
Household composition	. . .	3. One sampled person under age 16; one aged 16 or over; more than one, all under age 16; more than 1, all aged 16 or over; more than one, mixed ages
6–11 years		
Household size	1. 1–2, 3–4, 5–6, 7 or more	. . .
Race and Hispanic origin of sampled person	. . .	2. Non-Hispanic black, Hispanic, other
Sex of household reference person	. . .	3. Male, female
Home ownership	2. Yes, no	. . .
Urbanicity	3. Urban, suburban, rural	. . .
Census region	4. Northeast, Midwest, South, West	. . .
Number of sampled persons in household	. . .	4. 1–8, 9 or more
12–19 years		
Education of sampled person	1. Less than high school or not known, high school, more than high school	1. Less than high school or not known, high school, more than high school
Urbanicity	2. Urban, suburban, rural	. . .
Census region	3. Northeast, Midwest, South, West	. . .
Sex of sampled person	4. Male, female	. . .
20–59 years		
Urbanicity	1. Urban, suburban, rural	. . .
Census region	2. Northeast, Midwest, South, West	. . .
Race and Hispanic origin of sampled person	3. Non-Hispanic black, Hispanic, other	5. Non-Hispanic black, Hispanic, other
Sex of sampled person	4. Male, female	4. Male, female
Household size	5. 1–2, 3–4, 5–6, 7 or more	1. 1–2, 3–6, 7 or more
Self-reported health status	. . .	2. Fair/more than fair/not known, poor
Limited physically, emotionally, or mentally	. . .	3. Yes, no
60 years and over		
Self-reported health status	1. Excellent/very good, good/fair, poor/unknown	. . .
Urbanicity	2. Urban, suburban, rural	3. Areas with population over 3 million, other
Sex of household reference person	3. Male, female	. . .
Education of sampled person	4. Less than high school or not known, high school, more than high school	. . .
Census region	5. Northeast, Midwest, South, West	. . .
Limited physically, emotionally, or mentally	. . .	1. Yes, no
Hispanic origin of sampled person	. . .	2. Hispanic, not Hispanic
Number of sampled persons in household	. . .	4. 1, 2–3, 4 or more

. . . Category not applicable.

NOTE: MEC is mobile examination center.

Vital and Health Statistics Series Descriptions

ACTIVE SERIES

Series 1. **Programs and Collection Procedures**—This type of report describes the data collection programs of the National Center for Health Statistics. Series 1 includes descriptions of the methods used to collect and process the data, definitions, and other material necessary for understanding the data.

Series 2. **Data Evaluation and Methods Research**—This type of report concerns statistical methods and includes analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. Also included are experimental tests of new survey methods, comparisons of U.S. methodologies with those of other countries, and as of 2009, studies of cognition and survey measurement, and final reports of major committees concerning vital and health statistics measurement and methods.

Series 3. **Analytical and Epidemiological Studies**—This type of report presents analytical or interpretive studies based on vital and health statistics. As of 2009, Series 3 also includes studies based on surveys that are not part of continuing data systems of the National Center for Health Statistics and international vital and health statistics reports.

Series 10. **Data From the National Health Interview Survey**—This type of report contains statistics on illness; unintentional injuries; disability; use of hospital, medical, and other health services; and a wide range of special current health topics covering many aspects of health behaviors, health status, and health care utilization. Series 10 is based on data collected in this continuing national household interview survey.

Series 11. **Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey**—In this type of report, data from direct examination, testing, and measurement on representative samples of the civilian noninstitutionalized population provide the basis for (1) medically defined total prevalence of specific diseases or conditions in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics, and (2) analyses of trends and relationships among various measurements and between survey periods.

Series 13. **Data From the National Health Care Survey**—This type of report contains statistics on health resources and the public's use of health care resources including ambulatory, hospital, and long-term care services based on data collected directly from health care providers and provider records.

Series 20. **Data on Mortality**—This type of report contains statistics on mortality that are not included in regular, annual, or monthly reports. Special analyses by cause of death, age, other demographic variables, and geographic and trend analyses are included.

Series 21. **Data on Natality, Marriage, and Divorce**—This type of report contains statistics on natality, marriage, and divorce that are not included in regular, annual, or monthly reports. Special analyses by health and demographic variables and geographic and trend analyses are included.

Series 23. **Data From the National Survey of Family Growth**—These reports contain statistics on factors that affect birth rates, including contraception and infertility; factors affecting the formation and dissolution of families, including cohabitation, marriage, divorce, and remarriage; and behavior related to the risk of HIV and other sexually transmitted diseases. These statistics are based on national surveys of women and men of childbearing age.

DISCONTINUED SERIES

Series 4. **Documents and Committee Reports**—These are final reports of major committees concerned with vital and health statistics and documents. The last Series 4 report was published in 2002. As of 2009, this type of report is included in Series 2 or another appropriate series, depending on the report topic.

Series 5. **International Vital and Health Statistics Reports**—This type of report compares U.S. vital and health statistics with those of other countries or presents other international data of relevance to the health statistics system of the United States. The last Series 5 report was published in 2003. As of 2009, this type of report is included in Series 3 or another series, depending on the report topic.

Series 6. **Cognition and Survey Measurement**—This type of report uses methods of cognitive science to design, evaluate, and test survey instruments. The last Series 6 report was published in 1999. As of 2009, this type of report is included in Series 2.

Series 12. **Data From the Institutionalized Population Surveys**—The last Series 12 report was published in 1974. Reports from these surveys are included in Series 13.

Series 14. **Data on Health Resources: Manpower and Facilities**—The last Series 14 report was published in 1989. Reports on health resources are included in Series 13.

Series 15. **Data From Special Surveys**—This type of report contains statistics on health and health-related topics collected in special surveys that are not part of the continuing data systems of the National Center for Health Statistics. The last Series 15 report was published in 2002. As of 2009, reports based on these surveys are 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, and so compilations of Advance Data reports are no longer needed.

Series 22. **Data From the National Mortality and Natality Surveys**—The last Series 22 report was published in 1973. Reports from these sample surveys, based on vital records, are published in Series 20 or 21.

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, and so compilations of reports 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 5419
Hyattsville, MD 20782

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

U.S. DEPARTMENT OF
HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road, Room 5419
Hyattsville, MD 20782

DHHS Publication No. 2013–1359, Series 2, No. 159
CS240460