SAFER • HEALTHIER • PEOPLE™

Monitoring the Nation's Health

The National Health and Nutrition Examination Survey: Sample Design, 1999–2006

Vital and Health Statistics
Series 2, Number 155
May 2012

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

Suggested citation

Curtin LR, Mohadjer L, Dohrmann S, et al. The National Health and Nutrition Examination Survey: Sample design, 1999–2006. National Center for Health Statistics. Vital Health Stat 2(155). 2012.

Library of Congress Cataloging-in-Publication Data

National health and nutrition examination survey : sample design, 1999–2006.
p. ; cm.— (Vital and health statistics. Series 2, Data evaluation and methods research ; no. 155) (DHHS publication ; no. (PHS) 2012–1355)
NHANES
''April 2012.''
Includes bibliographical references.
ISBN 0–8406–0649–4
1. National Center for Health Statistics (U.S.) II. Title: NHANES. III. Series: Vital and health statistics. Series 2, Data evaluation and methods research ; no. 155. IV. Series: DHHS publication ; no. (PHS) 2012–1355.
[DNLM: 1. National Health and Nutrition Examination Survey (U.S.) 2. Health Surveys—United States—Statistics. 3. Nutrition Surveys—United States—Statistics. W2 A N148vb no.155 2012]
614.4'273—dc23 2012003361

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 155

The National Health and Nutrition Examination Survey: Sample Design, 1999–2006

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

Hyattsville, Maryland
May 2012
DHHS Publication No. (PHS) 2012–1355

Contents

Abstract

Background

Data collection for the National Health and Nutrition Examination Survey (NHANES) comprises three levels: a household screener, an 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 the preset selection probabilities for the desired demographic subdomains. After selection as an eligible sample person, the interview collects person-level demographic, health, and nutrition information as well as information about the household. The examination includes physical measurements, tests such as eye and dental examinations, and the collection of blood and urine specimens for laboratory testing.

Objectives

This report will first describe the broad design specifications for the 1999–2006 survey including survey objectives, domain and precision specifications, operational requirements, sample design, and estimations procedures. Details of the sample design are divided into two sections. The first section (NHANES 1999–2001 Sample Design) broadly describes the sample design and various design changes during the first three years of the continuous NHANES (1999–2001). The second section (NHANES 2002–2006 Sample Design) describes the final sample design developed and applied for 2002–2006. Weighting and variance estimation procedures are presented in the same manner; however, to correspond to the public data release cycles, the weighting and variance sections are separated into those used for 1999–2002, and those used for 2003–2006. Much of this report is based on survey operations documents and sample design reports prepared by Westat. Documentation of the survey content, procedures, and methods to assess nonsampling errors are reported elsewhere.

Keywords: sampling • weighting • variance estimation

The National Health and Nutrition Examination Survey: Sample Design, 1999–2006

by Lester R. Curtin, Ph.D., National Center for Health Statistics; Leyla K. Mohadjer, Ph.D., Westat; Sylvia M. Dohrmann, M.S., Westat; Jill M. Montaquila, Ph.D., Westat; Deanna Kruszon-Moran, M.S., National Center for Health Statistics; Lisa B. Mirel, M.S., National Center for Health Statistics; Margaret D. Carroll, M.A., M.S.P.H., National Center for Health Statistics; Rosemarie Hirsch, M.D., M.P.H., National Center for Health Statistics; Susan Schober, Ph.D., National Center for Health Statistics; Clifford L. Johnson, M.S.P.H., National Center for Health Statistics

Background

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 National Center for Health Statistics (NCHS). A unique feature of this survey is the collection of physical examination data for a nationally representative sample of the resident civilian noninstitutionalized U.S. population. The survey consists of questionnaires administered in the home followed by a standardized physical examination in specially equipped mobile examination centers (MECs).

NHANES I, the first cycle of NHANES, was conducted from 1971 through 1975. The second cycle (NHANES II) was conducted from 1976 through 1980. Although NHANES I and NHANES II each examined more than 20,000 individuals, the representation of the minority population in the sample was not large enough to permit adequate estimates of the health status of Mexican-American, Cuban-American, or Puerto Rican persons, or even the three groups combined. The objective of the Hispanic Health and Nutrition Examination Survey (HHANES), conducted from 1982 through 1984, was to produce estimates of health and nutritional status for the three major Hispanic subgroups that were comparable with the estimates available for the general population. NHANES III was the seventh in a series of surveys conducted by NCHS since 1960 using health examination procedures, and the third to include a nutrition component. It was fielded from 1988 through 1994. Details on the target populations, sample designs, and data collection procedures for these previous health examination surveys have been described in previous reports (6–13).

NHANES was again fielded in 1999, and in the tradition of the past national surveys it continues to be a keystone in providing critical and unique information on the health and nutritional status of the U.S. population. This information is essential for estimating the prevalence of various diagnosed and undiagnosed diseases and conditions, examining the differences in disease prevalence, and developing health policy. NHANES provides information to more than a dozen individual agencies and reflects coordination within the U.S. Department of Health and Human Services (HHS) in the collection of direct physical measurement data.

The differences in the sample sizes and sample designs for the five cycles of NHANES and for HHANES should be considered when comparisons are

made across various NHANES surveys. For example, note that until NHANES III, the HANES surveys did not include persons aged 75 and over and that NHANES I and NHANES II did not oversample Hispanic persons. The sample design parameters for the five NHANES surveys are compared in Table A.

Introduction

Data collection for NHANES comprises three levels: a household screener, an 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 the preset selection probabilities for the desired demographic subdomains. After selection as an eligible sample person, the interview collects person-level demographic, health, and nutrition information as well as information about the household. The examination includes physical measurements, tests such as eye and dental examinations, and the collection of blood and urine specimens for laboratory testing.

Beginning in 1999, the design and operations for NHANES have taken a new direction. The major difference from previous cycles is that the current NHANES is being implemented as a continuous, annual survey of the noninstitutionalized civilian resident population of the United States. NHANES 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. Each calendar year of the NHANES comprises a nationally representative sample of this portion of the U.S. population. This design facilitates potential linkage to other health and nutrition surveys that provide yearly estimates and allows aggregate-level national estimates from NHANES each year or from combinations of years.

Because NHANES can visit only a small number of locations each year, variance estimates for single-year data are relatively unstable. In addition, releasing only 1 year of data increases the possibility of disclosure of a sample person's identity. These two factors, plus the need to provide timely national estimates, resulted in the decision to publicly release data in 2-year cycles. Annual estimates may only be made through limited access to the data in the NCHS Research Data Center (RDC). Although the annual samples are nationally representative, annual estimates should be produced only for the nation as a whole, for the recommended race and Hispanic origin subdomain, or for very broad sex-age subdomains within race and Hispanic origin. It is also recommended that in order to improve the statistical reliability and stability of estimates with larger variances, analysts use combinations of 2-year cycles. Combining data from 2-year cycles is particularly appropriate for rare events, for preparing estimates for very detailed demographic

Table A. Selected sample design parameters for the National Health and Nutrition Examination Surveys

Characteristic	NHANES I	NHANES II	Hispanic HANES	NHANES III	NHANES 1999–2006
Age of noninstitutionalized civilian target population	1–74 years	6 months–74 years	6 months–74 years	2 months and over	0 months and over
Geographic areas	United States (excluding Alaska and Hawaii)	United States (including Alaska and Hawaii)	Southwest for Mexican Americans; NY, NJ, CT for Puerto Rican persons; Dade County, FL, for Cuban persons	United States (including Alaska and Hawaii)	United States (including Alaska and Hawaii)
Average number of sample persons per eligible household	1	1	2–3	2–3	2
Number of survey locations	100	64	17 in Southwest; 9 in NY, NJ, CT; 4 in Dade County, FL	89	117
Domains for oversampling	Low income: children aged 1–5 years; women aged 20–44 years; persons aged 65 years and over	Low income: children aged 6 months–5 years; persons aged 60–74 years	Dade County, FL: 6 months–19 years and 45–74 years; Southwest and NY, NJ, and CT: persons aged 6 months–19 years and 45–74 years	52 subdomains were predesignated consisting of sex-age groups for black, Mexican American, and all others; target sample sizes were established for the subdomains	76 subdomains were predesignated consisting of sex-age groups for black and Mexican-American persons and income-sex-age groups for all other persons; target sample sizes were established for the subdomains[1]
Total sample size	28,043	27,801	15,931	39,695	50,939
Examined sample size	20,749	20,322	11,672	30,818	39,352
Years covered	1971–1974	1976–1980	1982–1984	1988–1994	1999–2006

[1]In 1999, only 53 subdomains were predesignated for black, Mexican American, and others (no income used).

subdomains, and for measures that may have considerable geographic variation.

The study locations selected for NHANES 1999–2004 were linked to the National Health Interview Survey (NHIS) at the county level (1). It was later determined that only the locations selected for 1999–2001 would be fielded, and that a second sample would be selected for NHANES 2002–2007 from a national frame utilizing the 2000 decennial census data. At that time, the planned data release cycle was to combine 3 years of data as 1999–2001, 2002–2004, and 2005–2007. During the fielding of the 2002–2006 sample, the NHANES data release cycles were set as 2-year cycles due to the need for more timely data release. As a result, the 2-year data cycle 2001–2002 uses the 2001 locations based on the linked NHIS design and the 2002 locations based on the independent design for 2002–2007. The locations designated for 2007 were discarded to permit 2-year data releases through 2006, and a new sample for 2007–2010 was selected, again from a national frame, to produce 2-year estimates for 2007–2008 and 2009–2010. The 2007–2010 sample design will be described in a separate report.

This report will first describe the broad design specifications for the 1999–2006 survey including survey objectives, domain and precision specifications, operational requirements, sample design, and estimations procedures. Details of the sample design will be divided into two sections. The first section (NHANES 1999–2001 Sample Design) will broadly describe the sample design and various design changes during the first three years of the continuous NHANES (1999–2001). The second section (NHANES 2002–2006 Sample Design) will describe the final sample design developed and applied for 2002–2006. Weighting and variance estimation procedures will be presented in the same manner; however, to correspond to the public data release cycles, the weighting and variance sections are separated into those used for 1999–2002, and those used for 2003–2006. Much of this report is based on survey operations documents and sample design reports prepared by Westat. Documentation of the survey content, procedures, and methods to assess nonsampling errors are reported elsewhere (see http://www.cdc.gov/nchs/nhanes.htm for more information).

Design Specifications

Survey Objectives

A necessary first step in designing a survey is to define the survey's primary analytical objectives. As in the previous NHANES, a primary purpose of NHANES 1999–2006 is to produce a broad range of descriptive health and nutrition statistics for sex, race and Hispanic origin, and age subdomains of the population. These data can then be used to measure and monitor the health and nutritional status of the noninstitutionalized population. Because NHANES was designed to produce cross-sectional data and because respondents may be recontacted over time for future interviews or examinations, a set of cross-sectional and longitudinal objectives was developed. The analytic goals of NHANES are to:

- estimate the number and percentage of persons in the U.S. population and in designated subgroups with selected diseases and risk factors;
- monitor trends in the prevalence, awareness, treatment, and control of selected diseases;
- monitor trends in risk behaviors and environmental exposures;
- analyze risk factors for selected diseases;
- study the relationship between diet, nutrition, and health;
- explore emerging public health issues and new technologies;
- establish a national probability sample of genetic material for future genetic research;
- provide baseline health characteristics for longitudinal analysis when data are linked to mortality outcomes through the National Death Index; and
- provide health information that can be linked, subject to confidentiality concerns, to contextual and other data sets such as the Centers for Medicare & Medicaid Services.

Domain and Precision Considerations

The set of domains for which specified reliability was desired in NHANES 1999–2006 consisted of sex and age groups for black persons, Mexican-American persons, and the remainder of the U.S. population. Table B provides the set of sampling domains in NHANES 1999–2006. To increase the precision of estimates for certain subdomains, oversampling was carried out for adolescents (aged 12–19), older Americans (aged 60 and over), Mexican-American persons, and the black population. In addition, in NHANES 2000–2006 pregnant women and all others at or below 130% of the poverty level were also oversampled. Even though data are released in 2-year cycles, the accumulation of at least 4 years of data is required to obtain an acceptable level of reliability for the domains given in Table B. Thus, to create estimates for smaller 2-year samples (or any annual estimates), collapsing of some of the above domains is necessary to produce adequate sample sizes for analysis purposes.

Two main requirements were established for NHANES III when considering the utility of a sample for analysis purposes. These two conditions were considered in the sample design of NHANES 1999–2006 as well. They included the following:

1. An estimated prevalence statistic on the order of 10% in a sex-age domain should have a relative standard error of 30% or less; and
2. Estimated (absolute) differences between domains of at least 10% should be detectable with a Type I error rate (α) of ≤ 0.05 and a Type II error rate (β) of ≤ 0.10.

To satisfy the first condition a sample size of about 150 examined persons was necessary. This assumed a design effect of 1.5 resulting from the variability in sampling rates across

Table B. Analytical subdomains classified by race and Hispanic origin, income, sex, and age: National Health and Nutrition Examination Survey, 1999–2006

		Others[1]	
Black	Mexican American	Non-low income	Low income (2000–2006 only)
Sex and age			
Males and females under 1 year	Males and females under 1 year	Males and females under 1 year	Males and females under 1 year
Males and females 1–2 years.	Males and females 1–2 years	Males and females 1–2 years	Males and females 1–2 years
Males and females 3–5 years.	Males and females 3–5 years	Males and females 3–5 years	Males and females 3–5 years
Males 6–11 years	Males 6–11 years	Males 6–11 years	Males 6–11 years
Males 12–15 years	Males 12–15 years	Males 12–15 years	Males 12–15 years
Males 16–19 years	Males 16–19 years	Males 16–19 years	Males 16–19 years
Males 20–39 years	Males 20–39 years	Males 20–29 years	Males 20–29 years
. . .	. . .	Males 30–39 years	Males 30–39 years
Males 40–59 years	Males 40–59 years	Males 40–49 years	Males 40–49 years
. . .	. . .	Males 50–59 years	Males 50–59 years
Males 60 years and over	Males 60 years and over	Males 60–69 years	Males 60–69 years
. . .	. . .	Males 70–79 years	Males 70–79 years
. . .	. . .	Males 80 years and over	Males 80 years and over
Females 6–11 years	Females 6–11 years	Females 6–11 years	Females 6–11 years
Females 12–15 years.	Females 12–15 years	Females 12–15 years	Females 12–15 years
Females 16–19 years.	Females 16–19 years	Females 16–19 years	Females 16–19 years
Females 20–39 years.	Females 20–39 years	Females 20–29 years	Females 20–29 years
. . .	. . .	Females 30–39 years	Females 30–39 years
Females 40–59 years.	Females 40–59 years	Females 40–49 years	Females 40–49 years
. . .	. . .	Females 50–59 years	Females 50–59 years
Females 60 years and over	Females 60 years and over	Females 60–69 years	Females 60–69 years
. . .	. . .	Females 70–79 years	Females 70–79 years
. . .	. . .	Females 80 years and over	Females 80 years and over

. . . Category not applicable.

[1]NHANES 1999 included only 23 subdomains for others. Separate subdomains for low-income others were not included until 2000.

density strata necessary to accommodate oversampling. The sample necessary to satisfy the second condition was about 420 examined persons; therefore, the second condition was the more stringent one.

These were the general ideas used in the sample design that provide guidance on potential analytic considerations. For example, for a very small demographic group such general considerations may indicate the need to combine 4 years of NHANES data for a specific variable and analysis; however, the sample design effects for each measured NHANES variable and for specific demographic subdomains can be quite different from the assumed general design effect of 1.5. The issues of precision and statistical power should be addressed for each specific analysis.

Operational Requirements

A unique feature of NHANES is the complete physical examination for each respondent in the sample. To standardize the administration, the examinations are carried out in MECs. Three separate MECs are in service at any given time. Following a carefully designed schedule, two MECS are in operation at study locations, while the third is either traveling or being prepared for operation at a new location.

In order to maintain a cost-efficient workload within each location while considering the time and the cost involved in moving a MEC between survey locations, the maximum number of study locations NHANES may visit in each annual sample is 15. Taking this into account, the number of sample participants selected in each study location should be between 300 and 600, with an average of approximately 450, to yield approximately 333 examined persons in each of the 15 locations visited that year.

Previous experience with NHANES I–III and HHANES indicated that response rates increased when a larger sample of persons was selected within households. One factor thought to be responsible for the increased response rates in multiple-sample participant households was that each person was given remuneration for his or her time and participation. Thus, another important factor considered in the final design was to maximize the response rates and reduce screening costs by selecting as large an average number of sample participants per household as possible. Another factor affecting response rates was the amount of travel necessary for respondents to visit a MEC. The primary sampling units (PSUs) for NHANES are typically defined as individual counties—rather than combinations of counties as in other area surveys—to increase the likelihood of achieving high response rates.

Sample Design

In 1996, a decision was made by NCHS to field NHANES as a continuous survey and to fully automate data collection in the field. The plan was to pilot test the content and operations for new survey in 1997–1998, and to field the survey starting in 1999 as a 6-year survey (with public data release and published national estimates based on two 3-year cycles) similar to NHANES III.

In response to a recommendation from the National Academy of Sciences, HHS submitted a plan to the Office of Management and Budget proposing to integrate and coordinate data collection systems within HHS (1,5). Due to a number of sample design factors, linkage of NHIS and NHANES at the sample participant level was determined to be not feasible. The initial sample design decision was that the NHANES 1999–2004 cycle of the continuous survey would be linked to NHIS at the county level. Instead of an independent sampling frame based on all counties in the United States, only the counties selected in the 1995–2004 NHIS sample design (2) were considered as the sample frame for NHANES.

In early 2001, after 2 years of field work, it was determined that the NHIS PSUs were not a cost-efficient design for the NHANES sample due to the need to oversample race and Hispanic origin domains. An independent national design was reinstituted for future years of the NHANES continuous survey. This change required many modifications of the original survey plans, including the selection of PSUs and the creation of sample weights and variance calculations. As a result only the NHIS PSUs selected for 1999–2001 were fielded, and an independent sample was selected for NHANES 2002–2007 from a national frame utilizing the 2000 decennial census data.

Later in 2001, once data collection and cleanup were finalized for 1999–2000, a decision was made to release all NHANES survey data in 2-year cycles to be timelier, while still protecting confidentiality and maintaining reasonable sample sizes for analytic purposes. Public-use data files were eventually released for 1999–2000 and for 2001–2002. The 2001–2002 data release combined the NHIS linked design for 2001 with the independent design for 2002. Based on the independent design, data release cycles for 2003–2004 and 2005–2006 were planned and implemented. Rather than combine the 2007 PSU with another independent selection for 2008, the set of PSUs designated for 2007 were discarded, and a new sample for 2007–2010 was selected, again from a national frame. This allowed the 2007–2008 data release to be based on the same sample design. In summary, the sample design for the ''continuous'' NHANES has evolved into the current approach of allowing content changes and releasing data every 2 years with the survey sample design staying the same for 4-year intervals (two 2-year cycles, with independent sample designs for the intervals 2003–2006, 2007–2010, and 2011–2014).

Given this history, details of the sample design for 1999–2006 will be presented in two sections. The first section (NHANES 1999–2001 Sample Design) will broadly describe the sample design and various design changes during the first 3 years of the continuous NHANES (1999–2001). The second section (NHANES 2002–2006 Sample Design) will describe the final sample design developed and implemented during 2002–2006. The impact of the sample design changes over the 8-year period 1999–2006 for estimation in 2-year data release cycles (sample weighting and variance estimation) will be described in the ''Weighting the Sample Data'' and the ''Variance Estimation'' sections of this report.

NHANES 1999–2001 Sample Design

Because the PSU selection and some aspects of the within-PSU selection for NHANES 1999–2001 differ greatly from the later years, those procedures are briefly reviewed in the sections below. The detailed selection procedures used for 2002–2006, including the PSU selection procedures for the NHANES 2002–2006 sample, are described in the ''NHANES 2002–2006 Sample Design'' section.

PSU sample selection for 1999–2001

The NHANES 1999–2001 PSUs were selected from a frame that included all counties in two panels of NHIS PSUs. Initially, a 6-year sample of PSUs (for NHANES 1999–2004) was selected from this frame; however, these PSUs were used only for NHANES 1999–2001. The sample was redesigned and a new sample of PSUs was selected from a national frame for NHANES 2002–2007. NHIS is an ongoing survey of the noninstitutionalized civilian population that collects data on general health-related issues such as health status, health care utilization and access, chronic diseases, and acute conditions. It also collects demographic data, as well as data on health insurance coverage, income, and program participation. The sample design for NHIS has undergone periodic revisions since its inception in 1957; however, the basic design is a multistage stratified area sample, with metropolitan statistical areas (MSAs) or groups of counties as the PSUs. The NHIS PSUs are selected approximately every 10 years by the U.S. Census Bureau based upon the most current decennial census information and in conjunction with other household surveys conducted by the Census Bureau.

To facilitate linkage to other health surveys, the PSUs selected in the 1995–2004 NHIS design were distributed into four panels, each panel considered to be a nationally representative sample. Two NHIS panels were allocated to linkage for the Medical Expenditure Panel Survey, conducted by the Agency for Health Care Research and Quality. The remaining two panels of NHIS PSUs were made available for linkage to NHANES.

The sampling domains created for NHANES 1999–2006 are shown in Table B. NHANES 1999 sampling domains are the same as the NHANES 2000–2001 domains, except that there are no breakouts by low-income status for all others. The NHANES sample was designed to yield a self-weighting sample for each sampling domain while producing an efficient workload for each PSU. PSUs were selected with probabilities proportionate to a measure of size (MOS). The selection probability of a PSU determines the maximum rate at which persons residing in that particular PSU can be selected. For

NHANES 1999–2001, the PSU MOS was defined to be

$$M_h' = \frac{M_h}{\pi_h} = \frac{l}{\pi_h}\sum_k A_k C_{hk}',$$

where

$$A_k = \sum_{i,l} T_{ik} r_{ikl} \frac{C^*_{..kl}}{C^*_{..k.}}$$

and

M_h' = PSU MOS for 1999–2001 accounting for the NHIS linkage;

M_h = basic PSU MOS;

Σ = summation;

A_k = coefficient of the race and Hispanic origin group k;

h = the NHANES PSU;

i = the Mexican-American density stratum;

k = the race and Hispanic origin subdomain;

l = the sex-age subdomain;

π_h = the selection probability for the NHIS PSU containing NHANES PSU h;

r_{ikl} = the sampling rate of persons in density stratum i for the (k,l)-th race and Hispanic origin-sex-age subdomain;

C_{hk}' = the population estimate for the year 1996 for race-ethnicity subdomain k in PSU h (Estimates for Mexican-American persons were not available; therefore, estimates from the 1990 census of the proportion of Hispanic persons who were Mexican were applied to estimates of the Hispanic population to obtain population estimates for Mexican-American persons.);

T_{ik} = the proportion of the U.S. population in the i-th density stratum for the k-th race and Hispanic origin subdomain;

$C^*_{..kl}$ = the most recent projection for the year 2000 total population count for race and Hispanic origin-sex-age subdomain (k,l) [Projections for Mexican-American persons were not available; therefore, estimates from the March 1994 Current Population Survey (CPS) of the proportion of Hispanic persons who were Mexican were applied to projections of the Hispanic population to obtain population projections for Mexican-American persons.]; and,

$C^*_{..k.}$ = the most recent projection for the year 2000 total population count for race and Hispanic origin subdomain k.

When the MOS is defined in terms of the sampling rates as well as population counts, PSUs with larger populations for oversampled subdomains have a greater probability of being selected. This reduces the amount of screening required compared with a design in which the MOS is a function of population counts alone.

The probability of selection of an NHANES PSU, conditional on the NHIS PSU having been selected, is

$$P_h = k_1 \frac{M_h'}{\sum_h M_h'} = k_1 \frac{M_h/\pi_h}{\sum_h (M_h/\pi_h)}$$

where k_1 is the number of PSUs selected.

Within-PSU sample selection for 1999–2001

Within each PSU, a sample of segments was selected. In NHANES 1999–2001, two types of segments were used: area segments, which were typically census blocks or groups of blocks; and new construction segments, which were sets of building permits for new residential construction. Two approaches to segment selection were used for NHANES 1999–2001. In all NHANES 1999 study locations and the first four study locations of NHANES 2000, dwelling units (DUs) constructed after 1990 were sampled by forming new construction segments containing data gathered from census permit files and field visits to the permit offices. As a result of concerns about the continued use of permit sampling, a double sampling (or two-phase) approach was developed to update the MOS of area segments prior to segment selection. This two-phase approach to segment selection was used in the last 11 PSUs of NHANES 2000, and in all the PSUs in 2001.

To reduce the amount of screening required, NHANES segments were stratified according to the proportion of the population that was Mexican American. These strata are referred to as "density strata." Higher sampling rates were used to sample Mexican-American persons within the density strata having higher proportions of this population.

Segments were also selected with probability proportionate to an MOS. The segment MOS has the same form as the PSU MOS. The MOS is based on the sampling rates as well as the population counts in order to give segments with larger populations for oversampled subdomains a greater probability of being selected. Because segments consist of census blocks or combinations of blocks, the segment MOS is calculated as a sum of block sizes. Let M_{hib} denote the MOS of block b, in density stratum i, in PSU h, where

$$M_{hib} = \sum_k A_{ik} C_{hibk} \frac{C_{hk}'}{C_{h.k.}},$$

$$A_{ik} = \sum_l r_{ikl} \frac{C^*_{..kl}}{C^*_{..k.}}, \quad \text{and}$$

r_{ikl} = the sampling rate of persons in the (k,l)-th race and Hispanic origin-sex-age-income (income was not a subdomain in 1999) subdomain in density stratum i;

C_{hibk} = the 1990 decennial census population count for race and Hispanic origin k in block b in density stratum i in PSU h;

$C_{h.k.}$ = the 1990 decennial census population count for race and Hispanic origin k in PSU h; and

C_{hk}' = the population estimate for the year 1996 for race and Hispanic origin k in PSU h.

The MOS for the segments j were the sums of the MOS of the block(s) comprising each segment. These MOS are denoted by M_{hij}.

The conditional probability of selection of segment j in stratum i within PSU h is then

$$P_{hij} = k_h \frac{M_{hij}}{\sum_{j \in h} M_{hij}},$$

where k_h is the number of segments selected in PSU h (typically 24, but there is a small amount of variation in the number of segments in each PSU).

With two-phase segment selection, the basic segment MOS, M_{hij}, was calculated as described above. However, the final MOS used in the first phase of segment selection was obtained by inflating M_{hij} to account for expected growth in the segment due to new construction. Ideally, M_{hij} would be adjusted to account for expected declines as well as expected growth, but intercensal data on population declines are not available at this level.

That is, the segment MOS for the first-phase selection is

$$M^*_{hij(1)} = g_{hj,E} M_{hij},$$

where $g_{hj,E}$ is the expected growth rate (computed as the ratio of the 1990 population plus the estimated population in new construction to the 1990 population) for the census-designated place containing segment j.

Based on the DU counts obtained for the first-phase sample, an actual growth rate, $g_{hj,A}$, was estimated for the segment by computing the ratio of the actual number of DUs counted to the expected number of DUs in the segment (based on 1990 census data). The MOS used for the second-phase segment selection was

$$M^*_{hij(2)} = \frac{g_{hj,A}}{g_{h,E}}.$$

Within NHANES segments, the sampling procedures for 1999–2001 were similar to those used in 2002–2006, as described in the ''NHANES 2002–2006 Sample Design'' section of this document. DUs were selected with equal probability, at a rate equal to the maximum within-segment sampling rate required to attain the subdomain sampling rates. (Note that the sampling rates used in NHANES 1999–2001 were different than those described in the ''NHANES 2002–2006 Sample Design'' section.) Persons were selected within DUs using the ratio of the subdomain sampling rate to the maximum subdomain sampling rate. The overall selection probabilities yield approximately equal sample sizes for each PSU.

NHANES 2002–2006 Sample Design

The survey locations in the NHANES 2002–2006 sample were selected as part of a larger sample intended for 2002–2007 based on a design much different from that used for the locations visited in 1999–2001. As mentioned earlier, once it was determined that the data would be released every 2 years (that is, 2003 and 2004 together, followed by 2005 and 2006), the locations intended to be fielded in 2007 were never visited. However, when referencing sample design issues such as the determination of sampling rates (see ''Sampling rates for 2002–2006'') and PSU selection (see ''Stratification and selection of PSUs for 2002–2006''), the intended 2002–2007 6-year sample is often referenced.

Summary for 2002–2006

The NHANES sample represents the total noninstitutionalized civilian population residing in the 50 states and District of Columbia. As with previous NHANES, a four-stage sample design was used in NHANES 2002–2006. The first stage of the sample design consisted of selecting the PSUs from a frame of all U.S. counties, using the 2000 U.S. Census Bureau data. The PSUs in the first stage were mostly 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).

For most of the sample, the second sampling stage consisted of area segments comprising census blocks or combinations of blocks. A regular sample of area segments was selected for the NHANES 2002–2006 samples. However, because these samples were based on the 2000 census data, the MOS used for sampling was updated if necessary for PSUs experiencing large growth since 2000.

Within PSUs, an average of 24 segments was sampled. The sample was designed to produce approximately equal sample sizes per PSU; most PSUs have exactly 24 segments. PSUs selected with certainty (with a probability of one) may have more or fewer than 24 segments to ensure appropriate representation in the sample. Additionally, some large certainty PSUs were treated as multiple-study locations with varying numbers of segments in each location to, again, 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 sample participants per segment.

The segment sample can be viewed as having two components. In most of the United States segments were selected with PPS at a uniform rate. The sampling rate used in this component was sufficient to satisfy the sample size requirements for all the sex-age domains except some of those for Mexican-American persons. The second component was restricted to 2000 census block groups with moderate or large proportions of Mexican-American persons according to the 2000 decennial census. A higher sampling rate was designated in such areas. The block groups were stratified according to the proportion of the population in the minority subdomains.

The third stage of sample selection consisted of DUs, including noninstitutional group quarters such as dormitories. 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 were designated for screening in order to identify potential sample participants. The subsampling rates were set up to produce a national and approximately equal probability sample of households in most of the United States, with higher rates for the geographic strata with high minority concentrations. The screening rate in each stratum was designed to produce the desired number of sample

participants for the most difficult race and Hispanic origin-sex-age-income domain (i.e., the domain sampled at the highest rate) in the minority stratum.

Persons within the occupied DUs, or households, were the fourth stage of sample selection. All eligible members within a household were listed, and a subsample of individuals was selected based on sex, age, race and Hispanic origin, income, and pregnancy status. The subsampling rates and designation of potential sample participants within screened households were arranged to provide approximately self-weighting samples for each subdomain within minority density strata and simultaneously to maximize the average number of sample participants per sample household.

A summary of the expected annual sample sizes at the design stage is shown below:

Number of study locations	15
Number of segments	360
Number of DUs to be screened annually	12,500
Number of households to be screened annually	11,000
Number of sampled persons each year	6,525
Number of examined persons	5,000

The remainder of this section will focus on the calculation of the sampling rates (see ''Sample rates for 2002–2006''), the selection of PSUs (see ''Stratification and selection of PSUs for 2002–2006''), and within-PSU selection procedures (see ''Selection of segments''). Again, since the sampling rates and PSU sample were created based on a 6-year sample design, much of those sections will reference the 2002–2007 sample from which the 2002–2006 study locations were selected.

Sampling rates for 2002–2006

The sampling rates used for the 2002–2006 sample were developed for a sample intended to span the years 2002–2007. The rates required for sampling race and Hispanic origin-sex-age-income domains are the driving force in all stages of sampling for NHANES. The ''Calculation of sampling rates and screening to achieve a self-weighting sample'' section describes the calculation of those sampling rates and indicates the amount of screening required to achieve a self-weighting sample. The ''Oversampling in Mexican-American domains'' section addresses the increased efficiencies that can be gained by oversampling Mexican-American domains in high-Mexican areas. The final overall sampling rates used for PSU selection are given in the ''Final sampling rates'' section. In the ''Departures from a self-weighting sample'' section, the conditions under which the sample may deviate from a self-weighting design are presented.

Calculation of sampling rates and screening to achieve a self-weighting sample

NHANES is a multistage national area probability survey with fixed sample size targets for sampling domains defined by race and Hispanic origin, sex, age, and low-income status. Thus, the first step in determining the measures of size to be used for sampling at each stage is to calculate the sampling rate for each domain. The sampling rate for a domain depends on the target examination sample size, the expected examination response rate, and the estimated size of the population. These sampling rates determine the amount of screening that will be required.

To calculate sampling rates, expectations for response rates must be set. In most cases, the domain-specific NHANES 1999–2000 response rates were assumed to be unchanged for NHANES 2002–2007; for a few domains for which the NHANES 1999–2000 response rate was 100%, a response rate of 95% was assumed. The response rates used in the calculations ranged from 58% to 98%, with the lowest response rates assigned to the most challenging sampling domains, such as older persons.

Several data sources were used to obtain national estimates of the 2004 noninstitutionalized civilian population by race and ethnicity for the NHANES 2002–2007 sample. At the time of PSU selection, population projections were not available for certain U.S. subpopulations, such as Mexican-American persons or noninstitutional civilian residents, so multiple data sources were needed to create these estimates. Age-sex distributions by race were from the U.S. Census Bureau's projections of the total resident U.S. population in 2004 for middle series migration data. The proportion of Hispanic persons who were Mexican American (noninstitutionalized civilians) came from the March 2000 CPS. These data were used to adjust the counts of Hispanic persons in order to estimate the number of Mexican-American persons in the 2004 population. The proportion of the total resident population that is civilian and noninstitutionalized was calculated using April 1, 2000, census projections for the monthly postcensal resident population and the monthly postcensal noninstitutionalized civilian population. National poverty estimates for others by sex, age, and ratio of income to poverty threshold for 1999 were found in Table 2 of ''Poverty in the United States,'' issued by the U.S. Census Bureau.

Oversampling in Mexican-American domains

The information in Table I in Appendix II was used to determine the overall sample size required to meet each domain target in NHANES 2002–2006. Among black persons, the domain requiring the most screening contained males aged 16–19; it was expected that 60,639 households must be screened to meet the 6-year target for this domain. The domain requiring the most screening among others contained low-income males aged 80 and over; it was expected that 59,738 households must be screened to meet the 6-year target for this domain. The domain requiring the most screening among Mexican-American persons contained males aged 60 and over; it was expected that 115,176 households would need to be screened to meet the 6-year target for this domain.

In NHANES, the screening cost per household is only a small fraction of the cost of the interview and examination, and up to a certain point it has only a

minor effect on total cost; however, when screening must be performed for 100,000 households or more, the screening cost begins to account for a substantial part of the total cost. In addition, the workload associated with screening so many households has a major impact on recruitment, supervision, and control of the entire operation. To attain the target sample sizes for these domains, areas with high proportions of Mexican-American persons were oversampled because it was expected that over the course of 6 years a self-weighting national sample of 60,639 households would attain the target sample sizes for the black and others domains.

Census blocks (the basis for the formation of second-stage sampling units, or segments) were stratified according to the population percentage of Mexican-American persons in the block group. Table C gives the definitions of the Mexican-American density strata.

To determine the oversampling rates to be used, the distribution of the total population and of the Mexican-American population across the density strata were examined. The concentration of the Mexican-American population in each density stratum determined the operational efficiency of the stratification scheme; the variation in concentrations and sampling rates across strata determined the statistical efficiency.

Table D provides the expected design effects due to oversampling in the high-density strata for each of the Mexican-American domains for which this oversampling was intended in NHANES 2002–2007. The expected design effect is highest for Mexican-American males aged 60 and over, the domain that required the most oversampling.

Table C. Definitions of the Mexican-American density strata for National Health and Nutrition Examination Survey, 2002–2007

Density stratum	Percent Mexican American in stratum
1	0–10
2	10–30
3	30–60
4	60 or more

Table D. Expected design effects for the analysis of annual samples due to oversampling in the high-density strata, for Mexican-American sex-age sampling domains in which this oversampling was used

Sex-age sampling domain (Mexican-American persons only)		Expected design effect for the analysis of annual samples
Male and female	Under 1 year	1.20
Male and female	1–2 years	1.00
Male	12–15 years	1.03
Male	16–19 years	1.03
Male	60 years and over	1.20
Female	12–15 years	1.03
Female	16–19 years	1.04
Female	60 years and over	1.07

The following is a derivation of the screening rate for the full 6-year basic sample:

- Base screening sample size for 6 years = 60,639 households;
- Fifty percent add-on for reserve = 90,959 households;
- Projected total households in the United States in 2000 = 107,492,953 households; and,
- Sampling rate for base sample = 90,959/107,492,953 ≈ 1/1,182.

The amount of additional screening required as a result of oversampling Mexican-American persons is the sum of the amounts of additional screening required in each density stratum in order to attain the highest sampling rate for the density stratum. In this case, a total of 5,321 additional households were estimated to be needed for this purpose.

Final sampling rates

Table II in Appendix II shows the sampling rates used for the selection of PSUs in NHANES 2002–2006 for each of the sampling domains in the four density strata. Despite no explicit target sample sizes for the infant domains (sample participants under age 1 year), persons in these domains were sampled at the maximum rate for their race and ethnicity within the density stratum; that is, the infant domains were ''take all'' domains. The sampling rates given in Table I in Appendix II were designed to provide a 50% reserve sample as well as a provision for expected nonresponse in each subdomain.

The sampling rates were calculated using the approach described in the ''Calculation of sampling rates and screening to achieve a self-weighting sample'' and ''Oversampling in Mexican-American domains'' sections. In each density stratum, the maximum sampling rate determined the screening sample; that is, in each density stratum, the sample of households to be screened was to be selected at the highest sampling rates that appeared for that density stratum. All screened persons in the subdomain having that maximum rate were to be retained in the sample. The screened persons in other subdomains were to be subsampled to bring the sampling rates for those subdomains down to the desired levels. The subsampling rates were designed to minimize the variability in sampling rates among strata while still achieving the desired precision.

Departures from a self-weighting sample

Calculating the sampling rates required several assumptions relating to population size and response rates. To the extent that these assumptions were not met, the actual screening required to reach the target sample sizes differed from the expected screening.

As stated in the ''Calculation of sampling rates and screening to achieve a self-weighting sample'' section, several data sources were used to develop the 2004 population projections used in the sampling rate calculations. These sources included the U.S. Census Bureau's July 2004 projections of the resident population by age, sex, and race and Hispanic origin, the 2000 postcensal

projections of the resident and noninstitutionalized civilian population, and the March 2000 CPS estimates of the proportion of Hispanic persons who were also Mexican American. National poverty estimates for white and others originated with census data from 1999. The population projections and resulting expected screening requirement numbers depended on the assumption that these 1999, 2000, and 2004 proportions continued to hold in the years of data collection.

Finally, as noted in the ''Calculation of sampling rates and screening to achieve a self-weighting sample'' section, the expected examination response rates were set equal to achieved examination response rates by domain for earlier years of NHANES. Screening requirements also varied from expectations depending on how much these earlier response rates differed from the actual experience in 2002–2006.

Stratification and selection of PSUs for 2002–2006

As mentioned previously, the operational requirements for NHANES are such that the amount of travel necessary for a sample participant to visit a MEC should be minimized to increase the likelihood of achieving high response rates. As a result, individual counties were chosen as the PSUs for NHANES. However, some counties have such small populations that their probabilities of selection would be lower than what is required to attain the sampling rates for some of the domains. If selected for the sample, they would introduce considerable variability into the weights. Consequently, these small counties were combined with one or more adjacent counties to form more efficient sampling units. For the same reason, independent cities in Virginia were combined with nearby counties.

The PSUs selected for the 2003–2006 sample were developed for a sample intended to span from 2002 to 2007. The frame for the full NHANES 2002–2007 included all counties in the entire country. From the approximately 3,100 counties and county equivalents in the United States, 2,882 PSUs were formed (most of which consisted of individual counties), a sample of 90 study locations was selected, and 15 per year were randomly allocated to each of the years.

Calculation of PSU MOS

The NHANES sample was designed to yield a self-weighting sample for each sampling domain while producing an efficient workload in each study location. PSUs were selected with probabilities proportionate to an MOS. The selection probability of a PSU determines the maximum rate at which persons residing in that particular PSU can be selected.

The expression used to define the PSU MOS is similar to that used in NHANES III. The MOS of PSU h, denoted by M_h, is a weighted average of estimated populations by race and ethnicity and was calculated as follows:

$$M_h = \sum_k A_k C_{hk},$$

$$A_k = \sum_{i,l} T_{ik} r_{ikl} \frac{C^*_{..kl}}{C^*_{..k.}},$$

where

i = the within-PSU Mexican-American density stratum;

k = the race and Hispanic origin and income subdomain;

l = the sex-age subdomain;

C_{hk} = the 2000 census population estimate for race and Hispanic origin and income subdomain k in PSU h (see below);

T_{ik} = the proportion of the U.S. population in the i-th density stratum for the k-th race and Hispanic origin and income subdomain;

r_{ikl} = the sampling rate of persons in density stratum i for the (k,l)-th race and Hispanic origin and income-sex-age subdomain;

$C^*_{..kl}$ = the most recent projection of the 2004 total population count for race and Hispanic origin and income-sex-age subdomain (k,l); and

$C^*_{..k.}$ = the most recent projection of year 2004 total population count for race and Hispanic origin and income subdomain k.

Because single counties rather than larger areas made up of groups of counties are optimal as NHANES PSUs, the M_h was first calculated with h representing a single county.

Population estimates for the total, black, and Hispanic populations were obtained from the ''Census 2000 Redistricting Data.'' Because county-level data for Mexican-American population data were not available, a ratio of Mexican-American to Hispanic persons was calculated using 1990 decennial census data for the county. Under the assumption that this ratio has not changed significantly, it was applied to the county's 2000 Hispanic population count to estimate the 2000 Mexican-American population, as follows:

$$C_{hM} = \frac{C^{1990}_{hM}}{C^{1990}_{hH}} C_{hH},$$

where C^{1990}_{hM} and C^{1990}_{hH} are the 1990 Mexican-American and Hispanic populations in county h, C_{hM} is the 2000 Hispanic population in county h, and C_{hM} is the estimated Mexican-American population for county h. The population for the ''other'' race and Hispanic origin domain was calculated by subtracting the 2000 Mexican-American population estimate and the black population from the total 2000 population for the county.

In order to arrive at estimates of the civilian, noninstitutional population for each racial and ethnic group used in the PSU MOS, these estimates were further adjusted by the percentage of the total population in the PSU that was institutionalized according to the 2000 decennial census because more detailed estimates were not available. (Note that this is not included notationally above.)

The factors A_{ik} shown in Table E are the weights used to assign the relative contribution from each racial and ethnic group in the computation of the MOS. Population concentration in each density stratum, T_{ik} must be estimated for Mexican-American persons because the sampling rates vary across density strata for some domains.

Table E. Values of T_{ik} and A_{ik} used in the calculation of the primary sampling unit measure of size

Race and Hispanic origin	Density	T_{ik}	A_{ik}
Mexican American	1	0.30	0.000614
	2	0.23	0.000614
	3	0.20	0.000614
	4	0.27	0.000614
Black	. . .	. . .	0.000405
Other, low income	. . .	. . .	0.000227
Other, non-low income	. . .	. . .	0.000118

. . . Category not applicable.

These concentrations were estimated using 1990 census data at the block group level and are also shown in Table E. Estimates of the factors T_{ik} are not necessary for the other race and Hispanic origin and income domains. For these groups it can be shown that $A_{ik} = A_k$ because $r_{ikl} = r_{kl}$ and $\Sigma_i T_{ik} = 1$.

Minimum measures of size

The selection probability of a PSU determines the maximum rate at which persons residing in that particular PSU can be selected for NHANES while retaining the self-weighting nature of the sample. If the MOS of a PSU is too small, the required sampling rates for some subdomains cannot be achieved. Consequently, special weighting procedures would be required for such PSUs, and the resulting variability in weights would increase sampling errors. To ensure that all required sampling rates could be achieved, counties with very small MOS were combined with other adjacent counties.

The condition that determines the minimum MOS of a PSU is

$P_h \geq \hat{r}_k$ for all h and k,

where

P_h = the probability of selecting PSU h; and,

$\hat{r}_k$ = the maximum sampling rate among the sampling domains for race and Hispanic origin k.

For certainty PSUs, this condition always holds, because $\hat{r}_k \leq 1$ and $P_h = 1$. For noncertainty PSUs, the probability of selecting PSU h is

$$P_h = c_{NC} \frac{M_h}{\sum_{h \epsilon NC} M_h}$$

where

NC = the set of noncertainty counties;

c_{NC} = the number of noncertainty PSUs to be selected; and,

M_h = the MOS for PSU h.

Thus, the condition that determines the minimum MOS is equivalent to

$$M_h \geq \hat{r}_k \frac{\sum_{h \epsilon NC} M_h}{c_{NC}}.$$

For each county, it was necessary to check whether the MOS of the county met the ''minimum MOS'' condition. Because the right-hand side is a constant, the first step in this check was to compute this product. The number of noncertainty locations, c_{NC}, was 72.

The second term on the right-hand side of the expression was found to be

$$\frac{\sum_{h \epsilon NC} M_h}{c_{NC}} = 578.39$$

Many counties have very low Mexican-American or black populations. For these counties, extreme measures were taken to ensure that the sampling rates for Mexican-American or black domains could be met. For this reason, the value of $\hat{r}_k$ used in the calculation varied depending on the Mexican-American or black population proportions in a given PSU. The values of $\hat{r}_k$ used in the calculation of the minimum PSU MOS are noted below:

- If at least 3% of the PSU population distribution was Mexican American, then $\hat{r}_k = \hat{r}_{\text{Mexican American}}$
- If less than 3% of the PSU population distribution was Mexican American, but at least 3% black, then $\hat{r}_k = \hat{r}_{\text{Black}}$
- If less than 3% of the PSU population distribution was Mexican American and less than 3% black, then $\hat{r}_k = \hat{r}_{\text{White/other}}$

The same rule was used in the selection of PSUs in 1999–2001. Cutoffs of at least 3% Mexican-American or at least 3% of black persons have traditionally been used in NHANES for this purpose and were again used here.

Based on this minimum MOS criterion, 256 counties were found to have MOS that were too small. These counties were combined with neighboring noncertainty counties. The neighboring counties had to be adjacent, and the maximum distance between any two points in the combined-county area had to be less than 125 miles. Also, unless there were no alternatives that met the aforementioned criteria, counties combined were from the same state.

Twenty counties that failed to be combined either had no surrounding counties that fit this qualification, or adjacent counties already belonged to a combination. These counties were manually combined by slightly loosening the mileage restrictions where necessary. Three boroughs in Alaska were not combined either by computer or manually because the distances between them and areas under consideration for combining were too great; in some cases, the areas were separated by water, making travel difficult between the areas.

After the necessary county combinations were made, the PSU MOS, M_h was recalculated with h representing the combined counties as a single PSU.

Selection of certainty PSUs

Some counties had an MOS large enough that they were selected with certainty, and a few of these were selected multiple times. These certainty PSUs were removed from the county frame prior to noncertainty PSU selection.

A PSU was identified as a certainty if its weighted MOS exceeded 75% of the initial sampling interval; that is, PSU h was included in the sample with certainty if

$$M_h > 0.75 \frac{\sum_{h=1}^{H} M_h}{90},$$

where H is the number of PSUs on the entire sampling frame.

Some certainty PSUs were so large that they warranted more than one study location; otherwise, weighting factors would have to be applied to ensure appropriate representation, and these weighting factors would reduce the efficiency of estimates. The number of study locations allocated to each certainty PSU was obtained by comparing the weighted MOS M_h for the PSU with the initial PSU sampling interval, $(1/90)\ \Sigma_{h=1}^{H} M_h$.

A total of 18 study locations in the full NHANES 2002–2007 90-location sample were assigned to certainty PSUs. These locations were in 11 counties; 3 counties contained multiple-study locations.

Stratification and selection of noncertainty PSUs

The stratification scheme for the full NHANES 2002–2007 PSUs was developed with the primary goal of efficiency for the 6-year sample and with the secondary goals of efficiency for 3-year and annual samples. Recall that at the time the 2002–2007 was selected, data were intended to be released in 3-year cycles. The stratification scheme has the added benefit of producing efficient 2-year samples because it was designed to ensure that the PSUs comprising the annual and multiyear samples were distributed evenly in terms of geography and certain population characteristics.

For the full 6-year sample, 12 major strata were defined based on geography and the MSA status of the PSUs. Seventy-two minor strata were defined based on the demographics of the PSUs. Each major stratum included six minor strata, and one PSU was selected from each of these final strata.

The 6-year sample had a one-PSU-per-minor stratum design; each annual sample had a one-PSU-per-major-stratum design. That is, all multiyear samples contained only one PSU per sampled minor stratum rather than multiple PSUs from the same stratum. Annual samples contained only one PSU from each major stratum.

In forming the major strata, the variables used were census region (Northeast, Midwest, South, and West) and MSA status. This resulted in the major stratum containing the Southern non-MSA PSUs being larger than the other strata. This particular stratum had a total MOS of approximately 4,990, while the MOS for the other strata ranged from 3,300 to 4,200. The Figure in Appendix II contains a depiction of the major strata; Table F contains descriptions of the PSUs within each major stratum.

The minor strata were constructed in such a way that they were equal in size to the extent possible (in terms of total MOS). The variables used to form the boundaries of the minor strata were minority status and the percentage of the population below the poverty level. The percentages of black persons and Mexican-American persons in each PSU were obtained from the 2000 census. The percentage of the population below poverty was based on the 1990 census, because more current data were unavailable.

Allocation of PSUs to time period

To have nationally representative annual samples (a design requirement of NHANES), study locations had to be assigned to years in a random fashion. The certainty PSUs were sorted according to their MOS, and the noncertainty PSUs were sorted by order of selection (for the NHANES 2002–2007 sample, this was within each major and minor stratum). Within each major stratum, minor strata were paired. Each pair was randomly assigned to the study years 3 years apart. The assignment of the pairs to the particular sets of study years and the assignment of the study years within the pair were random within the first major stratum, and all other major strata followed the same pattern.

The large certainty PSUs were assigned in a manner that appropriately reflected their relative size. For example, because one PSU was large enough to be selected with certainty in a 1-year sample and contained six study locations within the 90-location sample for 2002–2007, one location was assigned to each year. In 1999, this PSU was divided into three study locations along tract boundaries: the northeastern, southern, and northwestern areas of the county. These locations have been and will continue to be fielded in that order. A few PSUs were selected with certainty in a 3-year sample but not in a 2- or 1-year sample. These PSUs have two locations assigned; one of the study locations was randomly assigned to one of the first 3 years and the other to one of the second 3 years in a balanced fashion (3 years apart). Finally, several PSUs were selected with certainty in a 6-year sample but not in a 3-, 2-, or 1-year sample. These PSUs were paired; one PSU in each pair was randomly assigned such that the two PSUs in each pair were scheduled 3 years apart.

Targeted number of sampled persons in each PSU

The initial target number of examined persons per location was 333 based on the assumption of a total of 5,000 examined persons per year in 15 study locations. Once the sample of locations was selected, the examination targets were adjusted. The final target number of examined persons for certainty locations was obtained by adjusting this initial target by the relative size of the location. For certainty locations, this was calculated as the MOS allocated to the survey location divided by the initial sampling interval used for selecting noncertainty PSUs, $(1/72)\ \Sigma_{h \epsilon NC} M_h$. The relative size of certainty locations ranged from 0.80 to 1.20.

For all other noncertainty locations, the initial examination target was

Table F. Descriptions of the major strata formed for selection of the 2002–2007 primary sampling units

Major stratum	Description[1]
1.	Northeastern PSUs with less than 7.8% of the population in poverty
2.	Northeastern PSUs with more than 7.8% of the population in poverty
3.	Nonmetropolitan Midwestern and Western PSUs
4.	Metropolitan Midwestern PSUs with population less than 9.8% black
5.	Metropolitan Midwestern PSUs with population more than 9.8% black
6.	Nonmetropolitan Southern PSUs
7.	Metropolitan Southern PSUs with population less than 1.4% Mexican American and: - Less than 21% black; or, - Between 21% and 27.6% black and less than 10.8% of the population in poverty
8.	Metropolitan Southern PSUs with population less than 1.4% Mexican American and: - More than 27.6% black; or, - Between 21% and 27.6% black and more than 10.8% of the population in poverty
9.	Metropolitan Southern PSUs with population between 1.4% and 3.8% Mexican American
10.	Metropolitan Southern PSUs with population more than 3.8% Mexican American
11.	Metropolitan Western PSUs with population less than 14% Mexican American
12.	Metropolitan Western PSUs with population more than 14% Mexican American

[1]Threshold values are approximate.

NOTES: PSU is primary sampling unit. PSUs selected for 2007 were not used.

adjusted by the relative contribution of the location's stratum to the total noncertainty MOS. Finally, the target number of identified sample participants for a given study location was derived from the desired number of examined persons by inflating that number to account for the predicted combined screener, interview, and examination response rate for the study location.

For NHANES 2003–2004, combined screener, interview, and examination response rates were predicted using a logistic regression based on the exam response experience in NHANES III and 1999–2001. The same procedure was used for NHANES 1999–2001, using the NHANES III data. The predicted exam rates were used to determine the expected number of screened sample participants needed to reach each location's exam target. The data were used to fit a logistic regression model at the person-level with exam participation as the response and various sample participant characteristics such as race and Hispanic origin, age, and gender as well as PSU characteristics such as population size, region, poverty level, and education as independent variables. Probabilities of exam participation were calculated by race and Hispanic origin of the sample participant and selected PSU characteristics. The number of expected sample participants screened for each PSU was calculated by dividing the appropriate cell probabilities into the expected number of sample participants examined for that cell. After summing the expected number of sample participants to the location level, the response rates were calculated as the total expected number of examined sample participants divided by the total expected number of identified sample participants.

Beginning in 2005, NHANES response rates (combined screener, interview, and examination) for each location have been predicted annually using a linear regression based on the actual response rates and location-level characteristics of prior study locations. Prediction based on previous experience has proven more accurate than simply applying a single response rate across all study locations.

Each year the model was refit with the most recent data available at the time of the prediction. A relatively large number of geographic, demographic, and economic variables from the U.S. Census Bureau were assembled and brought into a linear regression model as potential independent variables with the study location response rate as the dependent variable. A stepwise regression was used. The final model for each year was decided based on a combination of the regression correlation coefficient and a statistic that adjusts for the total number of variables included in the model. The model was applied to the values of the selected variables for the current year's study locations to predict their response rates.

After these predicted response rates were reviewed with senior project staff, some of the rates were adjusted based on past experience. Once the response rate predictions were finalized, they were applied to the target number of exams to predict the number of identified sample participants that would be required in each study location to achieve the targets. This became the initial target number of identified sample participants.

Selection of segments

The second stage of the design involved sampling segments within each PSU (study location). In order to utilize the most recent available data on new construction, segments were selected as a continuous process about 5 to 6 months prior to the start of the field period for the study location.

The usual practice in area samples is to list all DUs in sampled segments and apply a prespecified sampling rate to the listed DUs. This approach gives all DUs the desired probabilities of selection. For example, if the sampling rate is 50%, then one-half of the DUs listed in the segments will be included in the sample. If the number of DUs has tripled due to new construction (that is, housing units built since the most recent decennial census), the same sampling rate will produce three times as many interviews and examinations as the number originally expected. Such

dramatic changes in the segment size are expected when the data collection period is several years after the most recent decennial census for which data files are available.

If the segment contains much new construction, the segment MOS may be inaccurate. As a result, either a larger-than-expected sample must be drawn from that segment or a weighting factor must be applied to all sample participants selected from that segment. Because highly variable sample sizes are not operationally feasible for NHANES, subsampling within PSUs would be necessary to attain equal sample sizes across PSUs; however, this would require the application of a weighting factor, which would reduce the efficiency of the sample.

To update a sampling frame when the sample is to be selected with respect to an MOS but a reliable estimate of the MOS is not available, double sampling (or two-phase sampling) can be used. Beginning with the PSUs sampled for the 2005 NHANES year, after field staff determined the number of DUs in the first-phase sample of segments, an updated MOS that reflected the ratio of the actual number of DUs to the expected number of DUs was calculated. The final sample of segments was selected by subsampling from the first-phase segments using the updated MOS (14).

Stratification within PSUs

The procedures for selecting the segment sample involve both explicit and implicit modes of stratification. The PSU strata described in the ''Stratification and selection of noncertainty PSUs'' section and the minority-density geographical strata described in the ''Oversampling in Mexican-American domains'' section comprise explicit stratification. To keep combined blocks within a single-block group, the stratification is based on characteristics of the block group in which segments are located. Within the geographical strata, there is implicit stratification created by sorting the area segments by tract number, block group within tract, and segment number within block group, and selecting a systematic sample with PPS.

Segment MOS

The segment MOS calculation is similar to that for the PSU MOS. Prior research on intraclass correlations and unit costs indicated that an average of 14 examined sample participants per segment was reasonably close to an optimum for most statistics in NHANES. As indicated earlier, operational requirements make it necessary to have a fairly constant number of examined sample participants per study location, usually about 333. This implies having 24 segments per PSU.

Because segments consist of census blocks or groups of blocks, the segment MOS is a sum of MOS calculated at the block level. Further, in study locations that have experienced significant growth since the 2000 decennial census, a two-phase sampling procedure was followed with the segment MOS calculated separately for each phase.

For the first phase, let $M_{hib(1)}$ denote the MOS of block b in density stratum i in PSU h, where

$$M_{hib(1)} = \sum_{k^*} A_{ik^*} C_{hibk^*},$$

$$A_{ik^*} = \sum_{l} r_{ik^*l} \frac{C^*_{..k^*l}}{C^*_{..k^*.}},$$

where

h = PSU;

i = density stratum;

b = block;

k^* = the race and Hispanic origin subdomain (black, Hispanic, and others income levels combined);

C_{hibk^*} = the 2000 population of race and Hispanic origin k^* in block b in PSU h;

r_{ik^*l} = the sampling rate of persons in density stratum i for the (k^*,l)-th race and Hispanic origin and sex-age subdomain;

$C^*_{..k^*l}$ = the most recent projection of the year 2004 total population count for race and Hispanic origin and sex-age subdomain (k^*,l); and

$C^*_{..k^*.}$ = the most recent projection of the year 2004 total population count for race and Hispanic origin subdomain k^*.

The factor A_{ik^*} is similar to the A_k used in the PSU sampling MOS calculation described in ''Calculation of PSU MOS;'' however, because the r_{ik^*l} vary across density strata for Mexican-American persons, the values for Mexican-American persons must be calculated by density stratum. Further, because income level is not available at the block level, the value for the non-low-income other persons was used for all other persons. Table G contains the A_{ik^*} values used in calculating the segment MOS for NHANES 2002–2006.

The MOS for the first-phase segments $j^{[1]}$ are the sums of the MOS

Table G. Values of A_{ik^*} used in calculating measures of sizes, by race and Hispanic origin and density stratum for segment selection in National Health and Nutrition Examination Survey, 2002–2006

Density stratum (i)	Numerator of A_{ik^*} for race and Hispanic origin k^* (denominator is 1,182)		
	Black	Mexican American	Other
1	0.478	0.616	0.14
2	0.478	0.694	0.14
3	0.478	0.755	0.14
4	0.478	0.851	0.14

of the block(s) comprising each segment. These MOS are denoted by $M_{hij^{[1]}(1)}$.

The MOS used for the second-phase segment selection is the segment growth rate. Based on the DU counts obtained for the first-phase sample segments, the growth rate is estimated by computing the ratio of the actual number of DUs (counted by NHANES staff in the field) to the expected number of DUs in the segment (based on the 2000 decennial census data).

Let $U'_{hij^{[1]}}$ denote the number of DUs found by NHANES staff in the field for the first-phase segment $j^{[1]}$ in PSU h. The growth ($g_{hij^{[1]}}$) of the first-phase segment $j^{[1]}$ is estimated as

$$g_{hij^{[1]}} = \frac{U'_{hij^{[1]}}}{U^{[0]}_{hij^{[1]}}},$$

where $U^{[0]}_{hij^{[1]}}$ is the number of DUs in segment $j^{[1]}$ according to the 2000 decennial census.

Thus, the second-phase MOS for segment $j^{[1]}$ selected in the first phase of sampling in PSU h is equal to

$$M_{hij^{[1]}(2)} = M_{hij^{[1]}(1)} g_{hij^{[1]}}.$$

Number of segments and their probability of selection

As discussed in the "Targeted number of sampled persons in each PSU" section, the person sample sizes for some survey locations selected with certainty were adjusted to account for their size relative to the other selected locations to minimize the effects of intraclass correlation. The number of segments selected in the certainty locations was also adjusted from 24 to account for the relative size of the location. As a result, some survey locations selected with certainty contained as few as 19 segments or as many as 29 segments in the second phase of segment sampling, denoted $n_{h(2)}$. To achieve proper within-segment sampling rates in study locations in which the segment sample was selected in two phases, the first-phase sample must be larger than the ultimate sample. Beginning in 2005, approximately twice as many segments were selected than ultimately needed (50 segments).

For each survey location, the conditional probability of selection of first-phase segment $j^{[1]}$ is

$$P_{hj^{[1]}(1)} = \min\left[\frac{n_{h(1)} M_{hj^{[1]}(1)}}{\sum_{j^{[1]}=1}^{N_h(1)} M_{hj^{[1]}(1)}}, 1\right],$$

where

$N_{h(1)}$ = total number of segments in first-phase segment frame in the h-th PSU;

$n_{h(1)}$ = total number of first-phase segments to be selected in the h-th PSU; and

$M_{hj^{[1]}(1)}$ = first-phase MOS of segment $j^{[1]}$ in the h-th PSU.

Given the first-phase segments, the conditional selection probability of second-phase segment $j^{[2]}$ is

$$P_{hj^{[2]}(2)} = \min\left[\frac{n_{h(2)} M_{hj^{[2]}(2)}}{\sum_{j=1}^{n_h(1)} M_{hj^{[2]}(2)}}, 1\right],$$

where

$n_{h(2)}$ = total number of second-phase segments to be selected in the h-th PSU; and

$M_{hj^{[2]}(1)}$ = second-phase MOS of segment $j^{[2]}$ in the h-th PSU.

The actual probability of selection of a segment depends on the MOS of the segment and the probability of selection of the location from which the segment is selected. The overall probability of selection of a second-phase segment $j^{[2]}$ is

$$P_{hj^{[2]}(2)} P_{hj^{[1]}(1)} P_h = n_{h(2)} n_{h(1)} \left(\frac{M_{hij^{[2]}(2)}}{\sum_{j^{[2]}=1}^{n_h(1)} M_{hij^{[2]}(2)}}\right)\left(\frac{M_{hij^{[1]}(1)}}{\sum_{j^{[1]}=1}^{N_h(1)} M_{hij^{[1]}(1)}}\right) P_h \quad [1]$$

Note that in survey locations that do not require the two-phase procedure, $n_{h(2)} = n_{h(1)} = n_h$. Also, in survey locations where no second phase is needed, $M_{hj^{[2]}(2)} = 1$. Substituting into the second-phase probability of selection above results in

$$P_{hj^{[2]}(2)} = \min\left(\frac{n_h(1)}{\sum_{j=1}^{n_h}(1)}, 1\right) = \min\left(\frac{n_h}{n_h}, 1\right) = 1,$$

so the segment probability of selection within one-phase locations is simply the first-phase probability of selection.

Minimum segment MOS

One goal of the sample design is to create equal probabilities of selection for each domain within a study location. This enables the selection of a nearly self-weighting sample and facilitates the selection of persons. To create equal probabilities, the within-segment sampling rate for a domain in study locations selected without certainty should be

$$r_{hijkl} = \frac{r_{ikl}}{P_h P_{hj^{[1]}(1)} P_{hj^{[2]}(2)}}. \quad [2]$$

For locations selected with certainty, $P_h = 1$, so the segment sampling rate should be $r_{hijkl} = r_{ikl} \,/\, P_{hj^{[1]}(1)} P_{hj^{[2]}(2)}$.

The within-segment sampling rates must be less than or equal to 1. The most severe constraint is for domains with the highest value of r_{ikl}. These maximum sampling rates are referred to as $\hat{r}_i$; that is, $\hat{r}_i = \max_{k,l}\{r_{ikl}\}$, thus

$$\max_{k,l}\{r_{hijkl}\} = \frac{\hat{r}_i}{P_h P_{hj^{[1]}(1)} P_{hj^{[2]}(2)}} \le 1. \quad [3]$$

Replacing the denominator in expression [3] with its equivalent as given in expression [1], the condition given in expression [3] becomes

$$\frac{\hat{r}_i}{n_{h(2)} n_{h(1)} \left(\frac{M_{hij^{[2]}(2)}}{\sum_{j^{[2]}=1}^{n_h(1)} M_{hij^{[2]}(2)}}\right)\left(\frac{M_{hij^{[1]}(1)}}{\sum_{j^{[1]}=1}^{N_h(1)} M_{hij^{[1]}(1)}}\right) P_h} \le 1$$

which is equivalent to

$$M_{hij^{[1]}(1)} \ge \left(\frac{\hat{r}_i}{P_h \frac{n_{h(2)}}{\sum_{j^{[1]}=1}^{N_h(1)} M_{hij^{[1]}(1)}}}\right)\left(\frac{\sum_{j^{[2]}=1}^{n_h(1)} M_{hij^{[2]}(2)}}{n_{h(1)} M_{hij^{[2]}(2)}}\right). \quad [4]$$

The first-phase minimum measure of size is a product of two factors. The first factor was calculated for the survey location based on known information. The second factor was based on the second-phase measures of size, which was not known at the time of selection of the first-phase segments.

For survey locations that do not require the two-phase process, the

second factor reduces to 1:

$$\frac{\sum_{j^{[2]}=1}^{n_{h(1)}} M_{hij^{[2]}(2)}}{n_{h(1)}\, M_{hij^{[2]}(2)}} = \frac{\sum_{j^{[2]}=1}^{n_{h(1)}}}{n_h 1} = \frac{n_h}{n_h} = 1,$$

and equation [4] reduces to

$$M_{hij^{[1]}(1)} \geq \frac{\hat{r}_i}{P_h \dfrac{n_h}{\sum_{j^{[1]}=1}^{N_{h(1)}} M_{hij^{[1]}(1)}}}, \qquad [5]$$

the minimum MOS for segments.

For survey locations where the full two-phase process was implemented, $M_{hij^{[2]}(2)}$ was not known when the first-phase segment was selected. In this case the second factor must be considered:

$$\frac{\sum_{j=1}^{n_{h(1)}} M_{hij^{[2]}(2)}}{n_{h(1)}\, M_{hij^{[2]}(2)}} = \frac{ave(M_{hij^{[2]}(2)})}{M_{hij^{[2]}(2)}}.$$

This factor would inflate the minimum MOS to account for expected growth in the segment due to new construction. Because the actual values were not known, an inflation factor constant across all segments was used. Based on empirical research, this inflation factor was set to 1.25.

In implementing the sample selection, the minimum MOS was made 50% greater than needed to attain the maximum sampling rates $\hat{r}_i$. As a result, the MOS was increased to permit the selection of a reserve 50% sample.

Within each PSU, the blocks reported on the block-level census files in each minority density stratum were sorted by tract, block group, and block number. Blocks with MOS below the minimum were combined with succeeding blocks until the desired measure was achieved. To the extent possible, the combinations were kept to the same block group. When the combinations came to the end of a block group without reaching the minimum, earlier blocks within the same block group were added. When necessary, blocks were combined across block groups within the same tract to form segments; however, collapsing of blocks across tracts was not permitted. Consequently, the combinations consisted of blocks in close geographic proximity, and, in most cases, they were adjacent blocks. As a result of the method of combination, some large blocks that could have been segments by themselves were combined with small blocks.

At the second phase of segment selection the constraint in expression [4] is equivalent to

$$M_{hij^{[2]}(2)} \geq \frac{\hat{r}_i}{P_h P_{hj^{[1]}(1)} \dfrac{n_{h(2)}}{\sum_{j^{[2]}=1}^{n_{h(1)}} M_{hij^{[2]}(2)}}}. \qquad [6]$$

The right side of expression [6] is the minimum MOS for the second-phase segment selection. Any first-phase segments, $j^{[1]}$, with MOS less than the minimum second-phase MOS were combined with adjacent segments to form the second-phase segments, $j^{[2]}$, prior to selection.

After second-phase selection, any $j^{[2]}$ segments that had been formed as a combination of first-phase segments to achieve the second-phase minimum MOS were disaggregated into their first-phase components for operational reasons. The within-PSU probability of selection was equal for the constituent segments. After completing the segment selection, the selected segments were denoted by j (with the superscript dropped) to simplify the notation.

Controlling sample size per PSU

Screening and interviewing begins approximately 3 weeks before the first examinations in a location. This ensures that there are enough identified and interviewed sample participants to fill available examination sessions. Once the MEC team arrives at a location (after conducting exams in a previous location only days before), examinations for interviewed sample participants begin. Examinations continue for approximately 5 weeks. After the last examination day, the field staff has limited time to travel to the next study location.

This strict time schedule for examining the sample participants in each study location necessitates the advance establishment of a fixed screening and examination workload in each location (see the "Minimum measures of size" section). As with any survey, it is not possible to predict the exact number of screened households that will supply the desired number of sample participants and examinations. This is further aggravated by variations in response rates from location to location.

A fixed number of sample participants is expected in the locations selected without certainty as a result of the constant sampling rate defined for each domain-density stratum across all study locations, r_{ikl}. Within the study location, the sampling rate used in density stratum i for domain (k,l) is

$$P_{hj} \frac{\hat{r}_i}{P_h P_{hj}} \frac{r_{ikl}}{\hat{r}_i} = \frac{r_{ikl}}{P_h}.$$

Therefore, the total number of sample participants in a noncertainty location is expected to be

$$\sum_{i,k,l} \left(\frac{r_{ikl}}{P_h}\right) C_{hikl} = \sum_{i,k,l} \left\{ \frac{r_{ikl}}{c_{NC} \left[\dfrac{\Sigma_k C_{hik} \Sigma_l r_{ikl} \dfrac{C_{..kl}}{C_{..k.}}}{\Sigma_h \Sigma_k C_{hik} \Sigma_l r_{ikl} \dfrac{C_{..kl}}{C_{..k.}}} \right]} \right\} C_{hikl},$$

which can be written as

$$= \left[\frac{1}{c_{NC} \Sigma_{h,i,k,l} C_{hik} r_{ikl} \dfrac{C_{..kl}}{C_{..k.}}} \right] \left[\frac{\Sigma_{i,k,l} r_{ikl} C_{hikl}}{\Sigma_{i,k,l} r_{ikl} \left[C_{hik} \dfrac{C_{..kl}}{C_{..k.}} \right]} \right],$$

where C_{hikl} is the actual population of race and Hispanic origin and sex-age domain (k,l) in location h and density stratum i, and c_{NC} is the number of locations selected without certainty. Note that the first term on the right-hand side is a constant. To the extent that the population distribution is approximately the same as in 2000, $C_{hikl} \approx C_{hik}\, C_{..kl} / C_{..k.}$, and the second term is approximately equal to 1. Therefore, the number of sample participants is approximately constant across these locations.

Because the number of segments per location is constant at 24 for all but the certainty PSUs, the variation in quotas per location is also reflected in segment sample sizes. In addition, the changes in the population distribution since the most recent census are likely to be greater among segments than

among locations. The average segment size is thus expected to vary more than the average location size, but even this variation is generally within a moderate range. The approximate equality that exists in sample participant sample sizes per location and segment does not occur in the screening sample. The amount of screening in a location is partially based on what proportion of the location population lives in high-density strata. The amount of screening per segment will vary considerably among the density strata. Consequently, it is necessary to use a procedure that can produce samples that are either somewhat larger or somewhat smaller than those arising from the application of the self-weighting sampling rates. See the ''Selection of sample participants'' section for more information on this procedure.

Selection of DUs and persons

The third stage of sample selection consisted of DUs including certain types of group quarters. All DUs in the sample segments were listed, and a subsample of DUs was designated for screening to identify potential sample persons for interviews and examinations. The subsampling rates are designed to produce a national approximately equal probability sample of DUs in most of the United States, and higher rates for the geographical strata with high minority concentrations. Within each geographical stratum, there was an approximately equal probability sample of DUs across all PSUs. While the discussion in this section is phrased in terms of the 2002–2006 locations, the same procedures also apply to the 1999–2001 locations.

Within-segment sampling rates

Within segments, DUs were selected with equal probability at a rate equal to the maximum within-segment sampling rate required to attain the subdomain sampling rates. That is, the sampling rate used to select DUs within segment j in PSU h is

$$\frac{\max_{i,k,l}\{r_{ikl}\}}{P_h P_{hj}},$$

where i is the density stratum of segment j.

Sample participants were selected within DUs using the ratio of the subdomain sampling rate to the maximum subdomain sampling rate. Thus, the overall selection probability for a person in race and Hispanic origin and sex-age-income subdomain (k, l) in density stratum i is

Pr[select PSU h] • Pr[select segment hj|select PSU h] • Pr[select a given DU in segment hj|select segment hj]

Pr[domain (ikl) flagged for selection in the given DU|the given DU in segment hj selected]

$$= P_h P_{hj} \frac{\max_{i,k,l}\{r_{ikl}\}}{P_h P_{hj}} \frac{r_{ikl}}{\max_{i,k,l}\{r_{ikl}\}} = r_{ikl} \ .$$

It is easily shown that these probabilities yield approximately equal sample sizes for each PSU.

Selection of sample participants

Once the DU sample was released to the field, each DU was screened to determine whether it is occupied, vacant, or for seasonal use only. Only occupied DUs, or households, were eligible. Once the sampled households were identified, a sample of persons to be interviewed and examined from each individual household was selected. All eligible members within a household were listed and a subsample of individuals was selected based on sex, age, race and Hispanic origin, income, and pregnancy status. Sample participants were selected at rates established to ensure that the target sample sizes by subdomain were achieved, and the average number of sample participants per household was maximized.

Considerable subsampling was needed to reduce the screening sample of households to the desired number of sample participants. If independent random or systematic selections had been made for the subdomains, in most cases, only one person in a household would have been selected and the average sample size per household would have been quite low, not much above one.

Experience with recent cycles of NHANES and with HHANES indicated that response rates improve when sample sizes within households are larger. Therefore, a method of subsampling was used to maximize the number of sample participants per household (conversely, this method minimizes the number of households containing sample participants). The effect of within-household clustering is not a large concern for NHANES because most analyses are done within subdomains and there is generally little within-household clustering at the subdomain level.

The method begins with the designated screening sample from which persons are to be subsampled. The persons are classified into Q subdomains with sampling rates r_1, r_2, ... r_Q. The subdomains are ordered by subsampling rate so that $r_q \le r_{q+1}$. Note that the screening rates are set so that $r_Q = 1$; that is, the screening rate is equal to the maximum subsampling rate.

The set of households designated for screening is partitioned into L unequally sized random subsets, such that the sizes of the subsets are proportionate to r_1, $r_{(2)}-r_{(1)}$, $r_{(3)}-r_{(2)1}$,..., $r_{(q+1)} - r_{(q)}$, $r_Q - r_{(Q\text{-}1)}$. It is clear that the sum of these proportions is equal to $r_Q = 1$, so that each screened household is assigned to exactly one of the sets.

The subsampling is then carried out as follows:

- In the first random subset of households (corresponding to $100*r_1$% of all screened households), all persons in the household are designated as sample participants.
- In the second random subset of households (corresponding to $100 * (r_{(2)} - r_{(1)})$%, of all screened households), all persons in the household are sample participants except those in the subdomain (1); therefore, those persons in the first subdomain were selected only in the first random subset, with probability r_1;

- In the third random subset of households, all persons in the household are sample participants except those in the subdomains (1) and (2). Thus, those in the second subdomain were selected only in the first two random subsets, with probability $r_1 + (r_{(2)} - r_{(1)}) = r_{(2)}$.
- This procedure is continued in this manner through the Q-th random subset, for which only persons in subdomain Q are sample participants.

This sampling procedure was implemented using a set of sampling flags that designate for each DU, the domains eligible for sampling. The interviewers were not required to carry out any subsampling operation. They were instead instructed by the system (based on the set of domain flags provided for each household) on which persons to include as sample participants. Note that because the sampling domain flags were prepared in advance of the screening, they were based on the expected distribution of the screened sample by race and Hispanic origin, sex, age, and income rather than the distribution actually achieved. Thus, this procedure was expected to produce small deviations in the sample from the desired number in each domain. Such deviations are inevitable when subsampling rates must be established before the screening is completed.

Instead of unrestricted randomization, a pseudorandom procedure was used that guaranteed that all sampled DUs within each sequence of 100 consecutive DUs were assigned different random numbers (because the random number assigned determined the set of domains to be selected). To start, a random number between 0.00 and 0.99 was assigned to the first DU, with a separate initial random number used in each study location. The number 0.41 was then used like a skip interval and was added successively to obtain the random number for the next case. The random number was then used in the manner described above to determine the sampling domain flags assigned to each case.

Initially, a screening sample was drawn for each study location using sampling rates 50% larger than those required to attain the target sample sizes in each domain. Each study location's screening sample was then divided into release groups. Each group was a systematic subsample of the screening sample, with the screening sample sequenced by segment number and a temporary geographically based sequence number prior to subsampling. Thus, each release group contained cases from all segments, except as limited by release group and segment size. Table H gives the expected distribution of the sample of DUs across release groups.

In most study locations the 50% release group (that is, group A) was released to the interviewers first. The yield from this group was monitored and used to project estimates of the total yield of sample participants expected from this group. Based on these figures, additional groups (or portions of groups) were released as needed. The sample was monitored on a daily basis to determine whether additional release groups were required. The cases in group Z had sampling flags indicating that all persons in the household should be sampled; it was designed to be used as a last resort, only when the sample yield in a study location was low after all other groups had been released. This release group was never utilized in NHANES 1999–2006.

Table H. Release group distribution for National Health and Nutrition Examination Survey, 2002–2006

Release group	Percent for 150% sample[1]
A	50
B	10
C	8
D	8
E	6
F	6
G	3
H	3
I	2
J	2
K	1
Z	1

[1]The amount the National Health and Nutrition Examination Survey needs (100% sample) plus a 50% reserve.

NOTE: The same distribution was used in 1999–2001.

Special samples

Supplemental sample of pregnant women

To improve the precision of the estimates for pregnant women, a supplemental sample of pregnant women was selected. Only women aged 15–39 were eligible for this supplemental sample. The original plan was to include in the sample all pregnant women (that is, all women reported by the screener respondent to be pregnant); however, taking all non-Hispanic white or other pregnant women in the high-minority density strata would have reduced the efficiency of the sample because of the impact on the differential probabilities of selection.

Thus, the decision was made to select pregnant women according to the maximum sampling rates in each density strata. Because the sampling rates do not vary by density strata for black persons and others, the same rate was used to sample all pregnant women in these groups. Mexican-American women were sampled at different rates depending on their density stratum. Therefore, the subsampling approach described above yielded a self-weighting sample of others and black pregnant women. The sampling rates for the pregnant women are shown in Table J.

Examination session subsamples

NHANES has two examination session subsamples: the morning subsample and the afternoon or evening subsample. Sample participants selected for the morning sessions were instructed

Table J. Sampling rates used to sample pregnant women in National Health and Nutrition Examination Survey, 2002–2006

Numerator of sampling rate[1] Race and Hispanic origin	Density stratum 1	Density stratum 2	Density stratum 3	Density stratum 4
Black	1.00	1.00	1.00	1.00
Mexican American	1.00	1.90	3.00	3.75
Other	1.00	1.00	1.00	1.00

[1]The denominator of the sampling rates is 1,182.

NOTE: Similar rates were used in 1999–2001.

to fast overnight; those selected for the afternoon or evening sessions were also instructed to fast, but for a shorter period of time. Data that are sensitive to fasting times should be analyzed separately for these two groups.

Because it is generally more convenient for household members to come to the MEC at the same time—which is believed to favorably affect response rates—the examination session subsample assignment was made at the household level. The assignment was based on the household identifier (ID). If the household ID was an even number, the household was assigned to the morning subsample; if the household ID was an odd number, the household was assigned to the afternoon or evening subsample. The examination session subsample was assigned immediately after DUs were selected.

Although the examination session subsamples were designed to be approximately half-samples, some deviations resulted. Additionally, sample participants did not always report to the assigned examination session. For example, some sample participants assigned to be examined in a morning session may have been unable to report to the MEC at that time; in such cases, the sample participants were permitted to schedule afternoon or evening examinations.

Examination and laboratory subsamples

The examination component of NHANES consisted of medical, dental, and physiological measurements, as well as numerous laboratory tests to assess various aspects of health. For some of these components, subsampling was required to reduce respondent burden and facilitate the scheduling and completion of examinations.

Sample participants were assigned to examination and laboratory subsamples by first using an algorithm to randomly divide the sample participants into 12 groups; combinations of these groups were predetermined to create the various subsamples. In some cases, sample participants were assigned to multiple subsamples. After subsample assignment, weighting factors were attached to each sample participant record as appropriate to reflect this stage of subsampling. Table III provides the specifications for the components requiring subsampling.

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 statistics they would have 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 sample observation. Weighting takes into account several features of the survey: the differential probabilities of selection for the individual domains, nonresponse to survey instruments, and differences between the final sample and the total population.

NHANES-sampled participants were weighted in order to accomplish the following objectives:

1. To compensate for differential probabilities of selection among subgroups (race and Hispanic origin and sex-age-income and pregnancy subdomains, and persons living in different geographic strata sampled at different rates).
2. To reduce biases arising from the fact that nonrespondents may be different from respondents.
3. To fix weighted sample data to match an independent estimate from the U.S. Census Bureau, of the target population totals.
4. To compensate, to the extent possible, for inadequacies in the sampling frame (resulting from omissions of some housing units in the listing of area segments, omissions of persons with no fixed address, etc.).
5. To reduce variances in the estimation procedure by using auxiliary information that is known with a high degree of accuracy.

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 above). The second step adjusted for nonresponse (objective 2). In the third step, the sample weights were poststratified to U.S. Census Bureau estimates of the U.S. population to simultaneously accomplish objectives three, four, and five. These steps were performed for respondents to each stage of the survey: the screener, personal interview, and MEC examination.

Because the sample design for the 1999–2001 PSU selection differed significantly from that for the 2002–2006 sample, weighting the 1999–2000 and 2001–2002 samples also differed somewhat from the latter years. For that reason, this section is divided into two parts: the first section describes those weighting procedures unique to the 1999–2002 samples; otherwise, the

second section describes the methods used for the 2003–2006 samples, which generally apply to the earlier years.

The national inflation weights described in ''Weighting the sample data for 2003–2006'' were the starting point for the screener weight calculation. Those weights were then adjusted for nonresponse (see ''Nonresponse adjustment and trimming of weights for 2003–2006'') to the screener and then post-stratified (see ''Post-stratification for 2003–2006''). The resulting weights were then the starting point (or base weights) for the calculation of the interview weights, which were then adjusted for nonresponse to the interview, and again post-stratified. Finally, those post-stratified interview weights were the base weights used in the calculation of the MEC examination weights. Those weights were adjusted for nonresponse to the MEC examination and then post-stratified. See the ''Computing final weights for 2003–2006'' section for the calculation of the final interview and MEC examination weights.

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 the variability in the weights, subject to operational and analytic constraints. Additionally, measures such as weight trimming were implemented to reduce the variability in the weights for NHANES. The impact of weight variability is minimal when estimates are for the demographic subdomains used in the design, but 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.

Weighting the Sample Data for 1999–2002

Weighting procedures for the first two 2-year cycles, 1999–2000 and 2001–2002 (as well as special weights for 1999–2002), were very similar to those described in ''Weighting the sample data for 2003–2006,'' with a few notable differences. First, an additional adjustment to the base weights was included to account for the number of newly constructed DUs completed between DU sample selection and data collection. Second, because the NHANES 1999–2001 PSUs were selected as part of a larger sample, the adjustment described in the ''Weighting the sample data for 2003–2006'' section required some modifications. Third, special 4-year weights (for 1999–2002) were required for these years to accommodate multiyear analyses using these data. These deviations from the general weighting methodology are described in turn below.

Adjustment for new construction

Building permit data were used to select segments that contained housing units built between 1990 and the 6-month period prior to data collection for all study locations in 1999 and for the first four locations in NHANES 2000 (16 study locations in all). Census data were used to select segments that contained housing units built before 1990; building permit data were used to select segments that contained housing units built after 1990. Because segments were selected about 6 months prior to the data collection in each study location, a new construction factor was calculated to account for housing units built after new construction segments were selected and prior to the beginning of data collection.

For those 16 study locations, a new construction factor, denoted by $f_{i(nc)}$ was calculated as

$$f_{i(nc)} = \frac{\sum(TOTMOS_{nc1} + TOTMOS_{nc2})}{\sum TOTMOS_{nc1}},$$

where $TOTMOS_{nc1}$ is the total MOS for new construction segments in the original frame and $TOTMOS_{nc2}$ is the total MOS for new construction segments in the supplemental frame (that is, the frame of permits issued after the selection of new construction segments). For all other study locations in the 1999–2002 sample, the new construction factor was set equal to 1. This factor was included in the base weight calculation described in the ''Calculating basic national inflation weights'' section.

Adjustment for subsampling of 1999–2001 PSUs

Originally, the 1999–2001 sample was part of a larger 6-year design developed with the target of 20 study locations per year or 120 overall. Simultaneously, a subsample with 14 study locations per year was also developed in anticipation of a smaller NHANES. As the design evolved, a sample with 15 study locations each year was adopted. In 1999, 15 study locations were scheduled over the course of the calendar year, but in order to conduct some pilot testing (and due to budgetary considerations) the first three scheduled study locations were used as field test sites, so that sample was reduced to 12 locations. All other years contained 15 locations.

Sample weights begin with each sampled person's probability of selection. For persons selected in the 1999–2001 sample, because this probability represents the probability of the person being selected over a 6-year sample, the chance of the person being selected in only one of the 6 years is that probability divided by six. In addition, the reduction of three study locations in 1999, or 87 instead of 90 over the projected 6-year interval, was reflected by further adjustment. As a result, weights were adjusted by a factor of 7.25 (= 6 × 15/12 × 87/90) in 1999 and 5.8 (= 6 × 87/90) in 2000 and 2001. This factor is referred to as the annual weighting factor (AWF_i) in the ''Weighting the Sample Data for 2003–2006'' section.

Calculation of 4-year weights

Sample weights for NHANES 1999–2000 were based on population estimates developed by the U.S. Census Bureau before the 2000 decennial census counts became available. The 2-year sample weights for NHANES 2001–2002, and all other subsequent 2-year cycles, are based on population estimates that incorporate the year 2000 census counts. Because different population bases were used, the 2-year

weights for 1999–2000 and 2001–2002 are not directly comparable. So that analysts could combine the 1999–2000 with 2001–2002 survey years in analyses, 4-year sample weights were created to account for the two different reference populations. Because sample weights for the subsequent years after 2002 were based on population estimates after the 2000 census, special multiyear sample weights were not needed for these samples. For more information about combining multiple 2-year data sets for analyses, see the ''Combining 2-year weights to analyze other multiyear samples'' section.

Weighting the Sample Data for 2003–2006

Calculating basic national inflation weights

The first-stage (or base) weight for each sample participant, denoted by $w_{i(BASE)}$, was calculated as the reciprocal of the sample participant's probability of selection. These sampling rates are provided in Table II and their derivation is described in ''Variance estimation for publicly released data.''

The base weight for a sample participant is simply the reciprocal of the sampling rate for the sampling domain of the sample participant (denoted r_i here, where the subscript i indicates the sample participant). For NHANES 2003–2006, the base weight was adjusted further to account for the proportion of DUs released, the proportion of deselected DUs, and the number of years in the sample being weighted. The final base weight was calculated as

$$w_{i(Base)} = \frac{1}{r_i} f_{i(release)} f_{i(deselect)} f_{i(year)} \ .$$

The following section briefly describes each component of this calculation.

Adjustments for the number of release groups fielded

The first component, the release factor ($f_{i(release)}$), was introduced to reflect the procedures used to obtain a relatively fixed sample size within each study location in NHANES. See the ''Selection of DUs and persons'' section for a description of this procedure. The sample participant base weight was adjusted according to the proportion of the total sample released to the field. The release factor, denoted by $f_{i(release)}$, was calculated as

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

D_i represents the proportion of sampled DUs released for screening in the location sample participant i was selected. If response rates are as predicted and the MOS used during sampling were current, the subsample factor would be approximately 1.5. That is, approximately two-thirds of the sampled cases were expected to be released.

In rare instances, less than one-half of the sample selected was released, yielding a factor greater than 2.0. This situation resulted when the demographics of the segments were not as expected. Factors greater than 2.0 are generally much higher than those for other locations and, as a result, are likely to dominate the weights for the sample. When such a large factor resulted from the sample release in a study location, its value was trimmed to 2.0 to bring the weights down to a more reasonable level.

Deselection of released DUs

The sample yield monitoring and evaluation methods used in NHANES III and subsequently in NHANES 1999–2006, occasionally suggested that the expected number of sample participants from released DUs would exceed the target sample size for the study location. In these instances, DUs were deselected or randomly removed from the set of DUs released, but not yet screened, in order to keep the sample size near the target. To account for the deselection, an adjustment factor was applied to the base weight of sample participants identified in the remaining units. In 1999–2006, the expected number of sample participants exceeded the manageable sample size in five study locations. The deselection factors for those locations were 1.58, 1.82, 2.05, 2.07, and 3.08. The factor, denoted by $f_{i(deselect)}$, was calculated as

$$f_{i(deselect)} = \frac{1}{(1-D_i)} \ .$$

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 one.

Adjustment for the number of years in the sample

Because the original selected sample was intended to be fielded over 6 years, the base weights calculated from the original sampling rates also correspond to a 6-year sample. In weighting subsets of those 6 years, the following factor must be applied:

$$f_{i(year)} = \frac{AWF_i}{\text{Number of years in sample}} \ .$$

AWF_i represents the factor which, when applied to the weights, takes the 6-year weights and converts them to annual weights. The divisor of $f_{i(year)}$ is simply the number of years in the sample to be weighted. For example, the divisor for the records in the 2003–2004 sample is two.

Nonresponse adjustment and trimming of weights for 2003–2006

Nonresponse adjustment

If every selected household had agreed to complete the screener and every selected person had agreed to complete the interview and the medical examination, weighted estimates using the base weights described in the ''Weighting the Sample Data'' section would be approximately unbiased estimates of characteristics for the civilian noninstitutionalized U.S. population. But in reality, some of the sample participants who were screened refused to be interviewed (interview nonresponse) and some of the interviewed sample participants refused the medical 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 unit nonresponse in NHANES for each stage of nonresponse. The nonresponse adjustment procedure consists of computing adjustment factors and applying these factors to the survey weights separately by nonresponse cell. Nonresponse adjustment reduces bias if response rates and survey characteristics vary from cell to cell and 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 sampling 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:

- Minimum of 30 respondents in each cell.
- 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 was used to identify variables most highly related to response propensity. See Table IV for the variables used to form the nonresponse adjustment cells.

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

$$f_{i(NR)} = \frac{\sum_{i=1}^{n_{a_s}} w_{i(Base)}}{\sum_{i=1}^{n_{a_r}} w_{i(Base)}} ,$$

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

$$w_{i(NR)} = w_{i(Base)} f_{i(\mathrm{NR})} .$$

Trimming

Nonresponse adjustments can contribute to extreme weights; therefore, 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 sampling weights can seriously inflate the variance of survey estimates. Thus, weight trimming procedures may be used to reduce the impact of any such large sample participant weights on the estimates produced from the sample.

Because trimming introduces bias in the estimates, the hope was that the resulting reduction in variances would 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 and Hispanic origin and sex-age-income and pregnancy sampling domains. Because Mexican-American persons, pregnant women, and others with low income (beginning in 2000) were oversampled in NHANES, the weights in their domains may be 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 and reason for selection (income or pregnancy) 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, this is an important characteristic to have in a trimming procedure.

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

$$f_{i(TR)} = \frac{\sum_{i=1}^{n_b} t_i}{\sum_{i=1}^{n_b} w_{i(Base)} f_{i(NR)}} ,$$

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

$$w_{i(TR)} = w_{i(NR)} f_{i(TR)} .$$

Post-stratification

The final step in the weighting procedure was post-stratification to known population totals to compensate for undercoverage or overcoverage of certain demographic groups and for any residual differential nonresponse among these groups. Post-stratification of sample weights to independent population estimates is used for several purposes. In most household surveys, certain demographic groups in the U.S. population (for example, young black males) experience fairly high rates of

undercoverage in survey efforts. Post-stratification 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. Post-stratification can also help to reduce the variability of sample estimates and achieve consistency with accepted U.S. figures for various subpopulations.

Post-stratification involves applying a ratio adjustment to the survey weights. Broad classes—called post-stratification cells or post-strata—are constructed using auxiliary data, and a single ratio adjustment factor is applied to all units in a given post-stratification cell. The numerator of the ratio is a ''control total'' obtained from a secondary source (these are available from http://www.cdc.gov/nchs/nhanes.htm); the denominator is a weighted total obtained using the survey weights. Therefore, at the post-stratum level, estimates obtained using the post-stratified survey weights will correspond to the control totals used. Because post-stratification is a ratio adjustment, this process will improve the efficiency of estimates provided the variables used in constructing post-stratification 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 post-stratum level, and the overall gains in efficiency due to post-stratification tend to be small.

A major effect of post-stratification is that it implicitly imputes for unit 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 post-stratification cells. This is obviously an oversimplification as the missed persons are likely to be different; however, in the absence of any detailed information on the characteristics of the missed persons, post-stratification appears to be the only reasonable technique available for reducing the bias due to undercoverage and nonresponse.

All control totals were obtained using undercount-adjusted weights from the March Supplement to the CPS from the year closest to the midpoint of the sample years being weighted. These CPS weights have undergone post-stratification to the U.S. Census Bureau's best estimates of the total noninstitutionalized civilian population of the United States including those not counted in surveys or in the most recent decennial census. The post-stratification, therefore, brings the weighted totals up to the level of the presumed total noninstitutionalized civilian population in the United States.

The post-stratification factors, $f_{i(PS)}$, were calculated as

$$f_{i(PS)} = \frac{N_c}{\sum_{i=1}^{n_c} w_{i(TR)}},$$

where N_c is the control total for post-stratification cell c containing the i-th sample participant, and n_c is the sample size of the post-stratification cell c containing the i-th sample participant. Thus, the post-stratified weights, $w_{i(PS)}$, were calculated as

$$w_{i(PS)} = w_{i(NR)} f_{i(PS)}.$$

Computing final weights for 2003–2006

The final weight for each sample participant was calculated as the product of the base weight and the nonresponse adjustment, trimming, and post-stratification factors; that is,

$$w_i = w_{i(Base)} f_{i(NR)} f_{i(TR)} f_{i(PS)}.$$

More specifically, the final screening weight was calculated as

$$w_{i(S)} = w_{i(Base)} f_{i(NR,S)} f_{i(TR,S)} f_{i(PS,S)},$$

the final interview weight was calculated as

$$w_{i(I)} = w_{i(Base)} f_{i(NR,S)} f_{i(TR,S)} f_{i(PS,S)} f_{i(NR,I)} f_{i(TR,I)} f_{i(PS,I)},$$

and the final examination weight was calculated as

$$w_{i(E)} = w_{i(Base)} f_{i(NR,S)} f_{i(TR,S)} f_{i(PS,S)} f_{i(NR,I)} f_{i(TR,I)} f_{i(PS,I)} f_{i(NR,E)} f_{i(TR,E)} f_{i(PS,E)}.$$

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

Any sample participant who did not respond to the sample participant interview was assigned an interview weight of ''0.'' These sample participants were considered ineligible for the examination and assigned a MEC examination weight of ''0.'' These records were not released to the public. Sample participants who did complete the interview and were eligible for the examination but did not respond, were assigned MEC examination weights of ''0'' and their records are included in the public release.

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

Subsample weights

As discussed in the ''Special samples'' section, some NHANES respondents were also asked to participate in survey components that were statistically defined (or random) subsamples of the NHANES MEC-examined sample. Data collected from these participants include a variety of lab, nutrition or dietary, environmental, audiometry, and mental health components. Each of these subsamples was selected in a manner such that each subsample is a nationally representative sample.

For example, some but not all participants were selected to give a fasting blood sample on the morning of their MEC exam. The subsamples selected for these components were chosen at random with a specified sampling fraction (for example, one-half 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).

When data collected via one of these subsamples were released, special survey weights were constructed for that subsample and included in the data file. These weights differ from the full examination weight, and the subsample-specific weights must be used for statistical estimation of measures collected only in that 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, there would have to be random overlap between the subsamples and appropriate weights would need to be recalculated. For example, no sample weights are provided for the overlap in the fasting subsample with an environmental subsample. See the respective survey protocol or documentation for more specific information on each subsample.

Combining 2-year weights to analyze other multiyear samples

Only 2-year weights were calculated for NHANES 2003–2004 and 2005–2006. To combine these cycles for 4-year estimates, the 4-year weights are calculated as one-half times the 2-year weights.

The 4-year sample weights for NHANES 1999–2002 may be combined with the 2-year weights for the later survey cycles (NHANES 2003–2006) to create 6- or 8-year weights: for 6-year weights for 1999–2004, the 6-year weight is equal to two-thirds times the 4-year 1999–2002 weight for survey years 1999–2000 and 2001–2002, and one-third times the 2-year weight for survey years 2003–2004. For 8-year weights from 1999–2006, the 8-year weight is equal to one-half times the 4-year 1999–2002 weights and one-fourth times the 2-year weights for survey years 2003–2004 and 2005–2006. 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. While 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.

To combine 2001–2002 with 2003–2004 and 2005–2006 to produce 6-year estimates, only 2-year weights are used and the 6-year weights are calculated as one-third times the 2-year weights. Future years of data can continue to be added using the same methods as above.

Variance Estimation

Sampling errors should be calculated for all survey estimates to aid in determining statistical reliability of those estimates. For complex sample surveys, exact mathematical formulas for variance estimates are usually not available. Variance approximation procedures are required to provide reasonable, approximately unbiased, and design-consistent estimates of variance. The 2-year NHANES samples are limited in analytic capabilities. Even though each 2-year sample is nationally representative, it was selected from only 30 PSUs and the sample sizes for some specific race and Hispanic origin-sex-age-income-pregnancy 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 (that is, 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. The first section (''Variance Estimation for NHANES 1999–2002'') summarizes the variance estimation procedures unique to the 1999–2002 samples, while the second section (''Variance Estimation for NHANES 2003–2006'') describes the creation of variables necessary for variance estimation on the public- and restricted-use data files for the 2003–2006 samples. Two variance approximation procedures which 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 of the replication approach is to select subsamples repeatedly from the whole sample, calculate the statistic of interest for each of these subsamples (or ''replicates''), and then 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 the derivation of replicates from a full sample. The jackknife procedure retains most of the sample in each replicate, whereas the BRR approach retains only 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 only provided 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.

For the linearization approach, nonlinear estimates are approximated by linear ones 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 the use of the Taylor Series Linearization methods for variance estimation in all NHANES surveys. SUDAAN, Stata, and the SAS survey procedures can be used to obtain variance estimated by this method.

Variance Estimation for NHANES 1999–2002

Variance estimation for publicly released data

In considering the most appropriate variance estimation technique for the NHANES 1999–2001 PSUs, an issue of concern was the fact that it was not feasible to account for the effect of subsampling counties from the NHIS PSUs on the variances of the estimates. The NHANES 1999–2001 sample of PSUs was essentially selected using systematic sampling from a sorted list with no explicit stratification. As a result, the delete-one jackknife method was used to create replicate weights because this method is generally used when explicit stratification has not been used to select the sample, even though systematic sampling may have been used. Further, because the delete-one jackknife method provides a completely unstratified variance estimate, it introduces some amount of between-stratum variation that does not exist. This additional variation was desirable as a means to compensate for the fact that variance estimates could not explicitly account for the subsampling from the NHIS PSUs.

With the delete-one jackknife method, each replicate was created by deleting all cases from each study location at a time, thereby making it possible to identify all records from a particular study location. The risk of location identification, coupled with the fact that the data files contain some geographic data and other characteristics of the area, led to concerns about disclosure risks in the release of the NHANES 1999–2000 data files. As a result, NCHS initiated research to examine the disclosure risks of NHANES before the release of these data. Various methods for splitting each PSU into two dissimilar pseudo-PSUs for the purpose of replication were considered, and the most optimal method, named the cluster-split method, was used to create replicates for the 1999–2000 sample (3).

Beginning in 2002, NHANES became a stratified design with two PSUs per stratum for the 2-year samples. Given this and the great number of replicates needed for continual 2-year releases, NCHS decided that in future data releases, only PSU and strata indicators will be released for variance estimation. As a result, a new method of variance estimation had to be developed for use with the publicly released data.

Continued research resulted in the creation of the masked variance units described in the next section, and these units were created for the 1999–2002 sample. While the replicate weights for the 1999–2000 sample still exist on the public-use file, it is recommended that the masked variance units be used with the Taylor Series Linearization method in the analysis of any data in the NHANES 1999–2002 sample.

Variance Estimation for NHANES 2003–2006

Variance estimation for publicly released data

PSUs are selected from strata defined by geography and proportions of minority populations as described in the ''Stratification and selection of PSUs for 2002–2006'' section. 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.

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 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. They 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 variance estimates. MVUs have been created for all 2-year survey cycles from NHANES 1999–2000 through 2005–2006. They can also be used for analyzing any combined 4-, 6-, or 8-year data set.

Many surveys swap data values between cases for disclosure limitation. Rather than swapping individual values, however, the NHANES procedure used in NHANES 2005–2006 (16) 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. Any PSUs with swapped segments are no longer completely associated with a single real PSU; thus, 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 2003–2006, estimates were first calculated for every segment in all study locations for comparative purposes. These estimates provide general descriptions of the segments such as percentage of Hispanic sample participants, prevalence of home ownership, and obesity prevalence that should be similar for swapped segments. Study locations that were the most at risk for data disclosure (locations with smaller populations or in rural areas) were then 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 on variance of swapping the pair of segments. The distance measure to determine the effect of swapping the pair segments on variance was calculated as

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

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, $v(\bar{x}_l|S^r)$ is the variance of the estimate after swapping, and $v(\bar{x}_l|S)$ is the variance of the estimate before swapping. Note that for the 1999–2002 data, the simpler distance measure, $D = \Sigma_{l=1}^{q} \left| (\bar{x}_l|S^r) - \bar{x}_l|S) \right|$, was used. This measure solely considers the differences in the selected estimates, rather than incorporating the variances of those estimates.

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 only swapped once. Consideration was also given to pairs of segments from at-risk study locations; these pairs were minimized where possible.

Further research by Park (15) indicated that variance estimates very generally tended to increase as more segments were swapped, although the variance for specific analysis variables could also be underestimated after swapping. For this reason, the amount of swapping (that is, the number of study locations determined to be at risk and the number of segments swapped per location) is limited.

Variance estimation in the RDC

For the current sample design, NHANES data are released to the public in 2-year data cycles. In addition to public-use data files, special data sets are available only through the NCHS' RDC. These special data sets are for (a) data items collected for an odd number of calendar years (1, 3, or 5 years); (b) data sets with geographically linked data 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.

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 MEC weight. This single-year weight can be combined with the MEC weight provided on the 2-year public-use file to create a 3-year MEC weight. All single-year sample weights were calculated in the same manner as the public-use 2-year weights described in the "Weighting the Sample Data" section of this report. 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 given the restrictions of the RDC such as the prohibition of publication of PSU-level estimates. Further, any subnational estimate that is generated out of an RDC analysis must be reviewed and approved by NCHS staff to protect the confidentiality of sample respondents.

More information on the RDC and lists of special NHANES data files are available on the NCHS website (http://www.cdc.gov/nchs/). Information on proposals for use of stored specimens is also available on the NCHS website.

References

1. Arnett R, Hunter E, Cohen S, Madans J, Feldman J. The Department of Health and Human Services Survey Integration Plan: Current progress. Proceedings of the Government Statistics Section, American Statistical Association 142–7. 1996.
2. Botman SL, Moore TF, Moriarity CL, Parsons VL. Design and estimation for the National Health Interview Survey, 1995–2004. National Center for Health Statistics. Vital Health Stat 2(130). 2000.
3. Dohrmann S, Curtin LR, Mohadjer L, Montaquila J, Le T. National Health and Nutrition Examination Survey: Limiting the risk of data disclosure using replication techniques in variance estimation. Proceedings of the Survey Research Methods Section, American Statistical Association 807–12. 2002.
4. Engel A, Murphy RS, Maurer K, Collins E. Plan and operation of the HANES I Augmentation Survey of Adults 25–74 Years, United States, 1974–1975. National Center for Health Statistics. Vital Health Stat 1(14). 1978.
5. Hunter EL, Arnett R. Survey "reinventing" at Health and Human Services. Chance 9(3):54–7. 1996.
6. National Center for Health Statistics. Plan and initial program of the Health Examination Survey. National Center for Health Statistics. Vital Health Stat 1(4). 1965.
7. National Center for Health Statistics. Plan, operation, and response results of a program of children's examinations. National Center for Health Statistics. Vital Health Stat 1(5). 1967.
8. National Center for Health Statistics. Plan and operation of a health examination survey of U.S. youths 12–17 years of age. National Center for Health Statistics. Vital Health Stat 1(8). 1969.
9. Miller HW. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–1973. National Center for Health Statistics. Vital Health Stat 1(10a). 1973.
10. National Center for Health Statistics. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–1973. National Center for Health Statistics. Vital Health Stat 1(10b). 1977.
11. McDowell A, Engel A, Massey JT, Maurer K. Plan and operation of the second National Health and Nutrition Examination Survey, 1976–80. National Center for Health Statistics. Vital Health Stat 1(15). 1981.
12. Maurer KR. Plan and operation of the Hispanic Health and Nutrition Examination Survey, 1982–84. National Center for Health Statistics. Vital Health Stat 1(19). 1985.
13. Ezzati TM, Massey JT, Waksberg J, et al. Sample design: Third National

Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Stat 2(113). 1992.
14. Montaquila JM, Bell B, Mohadjer L, Rizzo L. A methodology for sampling households late in a decade. Proceedings of the Survey Research Methods Section, American Statistical Association 311–15. 1999.
15. Park I. Primary sampling unit (PSU) masking and variance estimation in complex surveys. Survey Methodology 34(2):183–94. 2008.
16. Park I, Dohrmann S, Montaquila J, Mohadjer L, 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.

Appendix I. Glossary

Centers for Disease Control and Prevention (CDC)—CDC is one of the major operating components of the U.S. Department of Health and Human Services.

Computer-assisted personal interviewing (CAPI)—An interviewing technique in which the interviewer uses a laptop computer. The laptop displays the question text for the interviewer to read and provides any other necessary instructions to the interviewer. Interviewers record the respondent's answers using the keyboard. Software directs the interviewer to the next appropriate question based on the answers entered. For the National Health and Nutrition Examination Survey (NHANES), the screening interview and all subsequent interviews conducted in the household utilize CAPI. At screening, the CAPI system is also the mechanism with which eligible household members are selected for the survey.

U.S. Department of Health and Human Services (HHS)—The U.S. government's principal agency for protecting the health of all Americans and providing essential human services, especially for those who are least able to help themselves. CDC, including the National Center of Health Statistics (NCHS), operates under HHS authority.

Domain—A demographic group of analytic interest (analytic domain). Analytic domains may also be sampling domains if a sample design is created with the intention of meeting goals for those specific demographic groups. For NHANES, sampling domains are defined by race and Hispanic origin, income, age, and gender. See ''Sampling domain.''

Domain flags—See ''Sampling domain flags.''

Double sampling—A general term for a method used in a number of statistical applications, such as stratification and regression or ratio estimation. An application of double sampling is to update a sampling frame when the sample is to be selected with respect to a measure of size (MOS) but a reliable estimate of that measure is not available. For NHANES, double sampling or two-phase sampling, was used in second-stage units (SSUs or segments) late in the decade 2000–2010 when the population counts from the U.S. Census Bureau, used in the calculation of the MOS, were old and feared to be no longer representative of the study location. In the NHANES study locations for which an accurate MOS is not available, a larger-than-needed sample of segments was selected in the first phase. After field staff determined the number of dwelling units (DUs) in the first-phase sample of segments, an updated MOS that reflected the ratio of the actual number of DUs to the expected number of DUs was calculated. The final sample of segments was selected by subsampling from the first-phase segments using the updated MOS.

Dwelling unit (DU)—Also ''housing unit.'' A house, an apartment, a mobile home or trailer, a group of rooms, or a 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 any other individuals in the building and which have direct access from outside the building or through a common hall. In this report, the term is used generally to mean those DUs eligible (that is, excluding institutional group quarters), or that could be eligible (that is, vacant at the time of sampling, but could be occupied once screening begins) for the survey.

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. Those living in group quarters are usually 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: institutionalized group quarters and noninstitutionalized group quarters.

Institutionalized group quarters—Group quarters providing formally authorized and supervised care or custody in institutional settings, such as correctional facilities, nursing facilities or skilled nursing facilities, in-patient hospice facilities, mental (psychiatric hospitals) facilities, group homes for juveniles, and residential treatment centers for juveniles. Institutionalized group quarters are not included in the NHANES sample. (For group quarters included in NHANES, see ''Noninstitutionalized group quarters.'')

Household—The group of persons living in an occupied dwelling unit.

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

Masked variance units (MVUs)—A collection of SSUs aggregated into groups for the purpose of variance estimation designed to not reveal the identity of the selected primary sampling units (PSUs). For NHANES, rather than using the units as sampled, some pseudo-units are created by swapping segments between PSUs. The resulting units produce variance estimates that closely approximate the variances that would have been estimated using the ''true'' design variance estimates. MVUs have been created for all 2-year survey cycles from

NHANES 1999–2000 through 2005–2006. They can also be used for analyzing any combined 4-, 6-, or 8-year data set.

Maximum sampling rate—The largest probability of selection assigned to a demographic group within a survey design. This value within certain strata and demographic groups was used in determining 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-ethnicity-income groups of interest.

National Center for Health Statistics (NCHS)—NCHS is the principal health statistics agency in the United States. It 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, the National Health Interview Survey, NHANES, the National Health Care Surveys, and the National Survey of Family Growth, among others. CDC's NCHS is part of the U.S. Department of Health and Human Services (HHS).

National Health Interview Survey (NHIS)—A continuing household survey with the purpose of securing accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States. NHIS is one of NCHS' major data collection programs. The study locations selected for NHANES 1999–2001 were linked to NHIS at the county level, as they were selected from a frame that included all counties in two panels of the NHIS PSUs.

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

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

Office of Management and Budget (OMB)—OMB reviews survey designs, materials, and questionnaires proposed for use by government agencies. The review is conducted by the OMB's Office of Information and Regulatory Affairs.

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 have such a large MOS 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'').

Public-use 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 assure that the identities of the respondents are protected. This file is disseminated by NCHS to encourage widespread use of the survey data.

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. PPS sampling is used in NHANES in the selection of PSU and SSUs.

Race and Hispanic origin—The terms race and Hispanic origin are used in this report as they were used in sample selection, and usually referred to a single term: race-ethnicity. The racial and ethnic groups this term refers to are Mexican-American persons, non-Mexican-American black persons, and a third group consisting of all other persons.

Release group—A systematic subsample of a study location's screening sample, with the screening sample sequenced by segment number and a temporary geographically based sequence number. Each release group contained cases from all segments, except as limited by release group and segment size. In most study locations the 50% release group (that is, group A) was released to the interviewers first. The yield from this group was monitored and used to project estimates of the total yield of sampled persons (SPs) expected from this group. Based on these figures, additional groups (or portions of groups) were released as needed. The sample was monitored on a daily basis to determine whether additional release groups were required.

Replicates—Subsamples selected repeatedly from a sample used in some variance estimation approaches. With these approaches, the statistic of interest is calculated for each of the subsamples, 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 1999–2006 multiyear samples; the delete-one jackknife method was used in the creation of replicate weights for the 1999–2000 sample.

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 interviews are considered interview respondents. Persons agreeing to complete the in-home interviews and a mobile examination center (MEC) examination are considered MEC respondents.

Response rate—The number of survey respondents divided by the number of persons selected into the sample. The response rates referred to in this document specifically 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 file—An electronic data set containing respondent records from a survey which contain some

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

Sampling domain—See "Domain." There were 53 sampling domains in NHANES 1999, and 76 sampling domains in NHANES 2000–2006. Table B of this report contains the specific sampling domains for these years.

Sampling domain flags—Strings of "0" and "1" attached to each sampled DU in the CAPI system. There was one string for each race-ethnicity-income group, with each digit of the string representing one of the specific age-gender sampling domains. If the digit corresponding to an age-gender domain in a race and Hispanic origin string contained a "1," then all persons in that DU with the matching demographic characteristics were included in the sample. A "0" and "1" in each string was set based on the sampling rates.

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 and Hispanic origin and sex-age-income 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.

Sampling variance—The sampling variance is a measure of the variation of a statistic, such as a proportion or a mean, which is due to having taken a sample instead of collecting data from every person in the full population. It measures the variation of the estimated proportion or mean over repeated samples. The sampling variance is zero when the full population is observed, as in a census. For NHANES, the sampling variance estimate is a function of the sampling design and the population parameter being estimated (that is, a proportion or a mean). Many common statistical software packages compute "population" variances by default; these may underestimate the sampling variance. Estimating the sampling variance requires special software, such as those discussed in this report.

Sampling weight—For a respondent in NHANES, this is the estimated number of persons in the target population that he or she represents. For example, if a man in the sample represents 12,000 men in his race-ethnicity-income-age category, then his "sampling weight" is 12,000. The NHANES sampling weights were adjusted for different sampling rates (of the race-ethnicity-income-age-gender groups), different response rates, and different coverage rates among persons in the sample, so that accurate national estimates can be made from the sample. Because it is the product of all these adjustments, it is sometimes called the "final" sampling weight.

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 or screening interview consisted of a household roster, collecting the income level of the household, and the race and Hispanic origin, age, and gender of all members. In NHANES, only persons aged 18 and over can answer the screener.

Screening—The process of conducting, or attempting to conduct, the screening interview in the DUs contained in the groups released. Occupied DUs (households) are "screened" through the screening interview. Other units can also be "screened"; the process for these units is simply verification that they are vacant, or simply not DUs. See "Screener."

Screening sample—The sample of DUs selected for a study location.

Secondary sampling unit (SSU)—The second-stage selection unit in a multistage area probability sample. For NHANES, these are typically referred to as "segments."

Segment—A group of housing units located near one another, all of which were considered for selection into the sample. For NHANES, segments consist of a census block or groups of blocks. For NHANES, the selection of segments comprises 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 of course be selected with differing probabilities, but such differences in selection probabilities at various stages cancel out. NHANES 1999–2006 is a self-weighting sample of persons within each sampling domain.

Simple random sample—A sample in which all members of the population are selected directly and have an equal chance to be selected for the sample. The NHANES sample is not a simple random sample. The NHANES sample was stratified, selected in stages, and employed unequal chances of selection for the respondents, by race and Hispanic orign, income, age, and gender. Such designs are referred to as "complex" and require special software to estimate the variance of statistics computed from a sample with a complex design.

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 PSUs selected with certainty were divided into multiple study locations and fielded at different times.

Survey location—See "Study location."

Strata and stratification—The partitioning of a population of PSUs into mutually exclusive categories

(strata). Typically, stratification is used to increase the precision of survey estimates for subpopulations important to the survey's objectives. For the selection of PSUs fielded in 2002–2006, PSUs were stratified based on region, MSA status, and various population demographics. In NHANES 1999–2006, strata were designated for within-PSU selection based on the density of Mexican-American persons living in that area.

Target population—The population to be described by estimates from the survey. In NHANES 1999–2006, the target population was the resident civilian noninstitutionalized population of the United States, which excluded 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 District of Columbia.

Two-phase segment selection—See ''Double sampling.''

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

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

Weight—See ''Sampling weight.''

Appendix II. Tables and Figure

Table I. Derivation of expected screening requirements for National Health and Nutrition Examination Survey, 2002–2007

Race and Hispanic origin and income-sex-age sampling domain			Projected population in 2004	Target number of exams for 1 year	Projected number of households screened to have one examined person	Projected number of households screened to attain target no. of exams over 6 years in self-weighting area sample	Projected number of examined persons in "basic" area sample	Number of additional examined persons to attain target
Black	Male	Under age 1 year	587,103	48	213	60,639	285	0
	and	1–2 years	1,137,668	85	113	57,434	510	0
	female	3–5 years	1,696,528	85	76	38,699	510	0
	Male	6–11 years	1,791,468	85	68	34,933	510	0
	Male	12–15 years	1,352,908	92	97	53,486	552	0
	Male	16–19 years	1,223,145	92	110	60,639	552	0
	Male	20–39 years	4,611,191	85	32	16,066	510	0
	Male	40–59 years	4133420	85	38	19,504	510	0
	Male	60 years and over	1,694,838	85	87	44,310	510	0
	Female	6–11 years	1,733,088	85	72	36,782	510	0
	Female	12–15 years	1,316,008	85	95	48,439	510	0
	Female	16–19 years	1,248,516	85	102	52,273	510	0
	Female	20–39 years	5,623,481	85	25	12,998	510	0
	Female	40–59 years	5,042,331	85	30	15,313	510	0
	Female	60 years and over	2,478,671	85	68	34,558	510	0
Mexican American	Male	Under age 1 year	576,027	92	209	60,639	290	262
	and	1–2 years	1,127,767	90	114	61,383	533	7
	female	3–5 years	1,604,386	85	80	40,897	510	0
	Male	6–11 years	1,512,695	85	77	39,392	510	0
	Male	12–15 years	935,649	92	125	68,932	486	66
	Male	16–19 years	868,279	92	142	78,549	426	126
	Male	20–39 years	3,736,086	85	38	19,307	510	0
	Male	40–59 years	2,292,162	85	60	30,663	510	0
	Male	60 years and over	728,735	95	202	115,176	300	270
	Female	6–11 years	1,411,518	85	85	43,154	510	0
	Female	12–15 years	847,952	92	139	76,896	435	117
	Female	16–19 years	808,686	92	149	82,442	406	146
	Female	20–39 years	3,654,437	85	35	17,859	510	0
	Female	40–59 years	2,219,359	85	55	28,070	510	0
	Female	60 years and over	880,495	92	165	91,067	368	184

See footnote at end of table.

Table I. Derivation of expected screening requirements for National Health and Nutrition Examination Survey, 2002–2007—Con.

Race and Hispanic origin and income-sex-age sampling domain			Projected population in 2004	Target number of exams for 1 year	Projected number of households screened to have one examined person	Projected number of households screened to attain target no. of exams over 6 years in self-weighting area sample	Projected number of examined persons in "basic" area sample	Number of additional examined persons to attain target
Other, low income	Male	Under age 1 year	528,844	45	223	60,143	270	0
	and	1–2 years	1,045,330	54	105	33,911	324	0
	female	3–5 years	1,591,622	54	79	25,548	324	0
	Male	6–11 years	1,691,439	27	67	10,837	162	0
	Male	12–15 years	1,217,776	27	93	15,052	162	0
	Male	16–19 years	1,137,953	27	105	17,003	162	0
	Male	20–29 years	1,995,232	27	64	10,390	162	0
	Male	30–39 years	1,550,762	27	78	12,617	162	0
	Male	40–49 years	1,552,642	27	81	13,041	162	0
	Male	50–59 years	1,240,377	27	104	16,915	162	0
	Male	60–69 years	1,018,725	27	115	18,580	162	0
	Male	70–79 years	653,253	27	242	39,202	162	0
	Male	80 years and over	345,484	16	622	59,738	96	0
	Female	6–11 years	1,640,043	27	69	11,177	162	0
	Female	12–15 years	1,185,777	27	95	15,459	162	0
	Female	16–19 years	1,289,219	27	137	22,143	162	0
	Female	20–29 years	2,780,447	27	42	6,808	162	0
	Female	30–39 years	2,131,932	27	56	9,076	162	0
	Female	40–49 years	1,901,695	27	65	10,525	162	0
	Female	50–59 years	1,699,254	27	7	11,261	162	0
	Female	60–69 years	1,459,540	27	93	15,103	162	0
	Female	70–79 years	1,423,241	27	103	16,761	162	0
	Female	80 years and over	1,115,698	27	140	22,620	162	0
Other, non-low income	Male	Under age 1 year	2,174,859	70	58	24,508	420	0
	and	1–2 years	4,298,893	70	34	14,386	420	0
	female	3–5 years	6,545,504	70	22	9,148	420	0
	Male	6–11 years	7,000,436	70	21	8,834	420	0
	Male	12–15 years	5,040,067	71	27	11,501	426	0
	Male	16–19 years	4,986,659	76	28	12,602	456	0
	Male	20–29 years	11,670,740	79	13	5,980	474	0
	Male	30–39 years	13,051,517	81	14	6,562	486	0
	Male	40–49 years	16,293,817	82	10	5,072	492	0
	Male	50–59 years	13,165,863	79	13	6,242	474	0
	Male	60–69 years	8,020,536	80	20	9,460	480	0
	Male	70–79 years	5,323,446	79	33	15,437	474	0
	Male	80 years and over	2,627,703	70	67	28,166	420	0

See footnote at end of table.

Table I. Derivation of expected screening requirements for National Health and Nutrition Examination Survey, 2002–2007—Con.

Race and Hispanic origin and income-sex-age sampling domain			Projected population in 2004	Target number of exams for 1 year	Projected number of households screened to have one examined person	Projected number of households screened to attain target no. of exams over 6 years in self-weighting area sample	Projected number of examined persons in "basic" area sample	Number of additional examined persons to attain target
Other, non-low income—Con.	Female	6–11 years	6,668,178	70	22	9,275	420	0
	Female	12–15 years	4,821,197	68	28	11,515	408	0
	Female	16–19 years	4,621,916	70	28	11,629	420	0
	Female	20–29 years	11,094,039	75	12	5,590	450	0
	Female	30–39 years	13,128,858	79	12	5,624	474	0
	Female	40–49 years	16,348,739	79	10	4,583	474	0
	Female	50–59 years	13,472,653	75	12	5,524	450	0
	Female	60–69 years	8,613,202	72	20	8,558	432	0
	Female	70–79 years	6,151,219	67	30	12,112	402	0
	Female	80 years and over	3,834,159	68	55	22,428	408	0
Total	Female	. . .	280,025,082	4,966	6,275	2,169,145	28,615	1,178

. . . Category not applicable.

Table II. Final sampling rates and base weights for National Health and Nutrition Examination Survey, 2002–2006

Race and Hispanic origin and income-sex-age sampling domain			Numerator of sampling rate[1]				Base weights by domain			
			Density stratum 1 (less than 10%)	Density stratum 2 (10%–29%)	Density stratum 3 (30%–59%)	Density stratum 4 (60% or more)	Density stratum 1 (less than 10%)	Density stratum 2 (10%–29%)	Density stratum 3 (30%–59%)	Density stratum 4 (60% or more)
Black	Male	Under age 1 year	1.00	1.00	1.00	1.00	1,181.78	1,181.78	1,181.78	1,181.78
	and	1–2 years	0.95	0.95	0.95	0.95	1,247.72	1,247.72	1,247.72	1,247.72
	female	3–5 years	0.64	0.64	0.64	0.64	1,851.77	1,851.77	1,851.77	1,851.77
	Male	6–11 years	0.58	0.58	0.58	0.58	2,051.41	2,051.41	2,051.41	2,051.41
	Male	12–15 years	0.88	0.88	0.88	0.88	1,339.84	1,339.84	1,339.84	1,339.84
	Male	16–19 years	1.00	1.00	1.00	1.00	1,181.78	1,181.78	1,181.78	1,181.78
	Male	20–39 years	0.26	0.26	0.26	0.26	4,460.50	4,460.50	4,460.50	4,460.50
	Male	40–59 years	0.32	0.32	0.32	0.32	3,674.15	3,674.15	3,674.15	3,674.15
	Male	60 years and over	0.73	0.73	0.73	0.73	1,617.30	1,617.30	1,617.30	1,617.30
	Female	6–11 years	0.61	0.61	0.61	0.61	1,948.31	1,948.31	1,948.31	1,948.31
	Female	12–15 years	0.80	0.80	0.80	0.80	1,479.43	1,479.43	1,479.43	1,479.43
	Female	16–19 years	0.86	0.86	0.86	0.86	1,370.92	1,370.92	1,370.92	1,370.92
	Female	20–39 years	0.21	0.21	0.21	0.21	5,513.22	5,513.22	5,513.22	5,513.22
	Female	40–59 years	0.25	0.25	0.25	0.25	4,679.81	4,679.81	4,679.81	4,679.81
	Female	60 years and over	0.57	0.57	0.57	0.57	2,073.66	2,073.66	2,073.66	2,073.66
Mexican American	Male	Under age 1 year	1.00	1.50	2.20	3.00	1,181.78	787.85	537.17	393.93
	and	1–2 years	1.00	1.00	1.00	1.05	1,181.78	1,181.78	1,181.78	1,125.51
	female	3–5 years	0.67	0.67	0.67	0.67	1,752.24	1,752.24	1,752.24	1,752.24
	Male	6–11 years	0.65	0.65	0.65	0.65	1,819.19	1,819.19	1,819.19	1,819.19
	Male	12–15 years	1.00	1.00	1.00	1.50	1,181.78	1,181.78	1,181.78	787.85
	Male	16–19 years	1.00	1.25	1.50	1.50	1,181.78	945.43	787.85	787.85
	Male	20–39 years	0.32	0.32	0.32	0.32	3,711.67	3,711.67	3,711.67	3,711.67
	Male	40–59 years	0.51	0.51	0.51	0.51	2,337.11	2,337.11	2,337.11	2,337.11
	Male	60 years and over	1.00	1.50	2.20	3.00	1,181.78	787.85	537.17	393.93
	Female	6–11 years	0.71	0.71	0.71	0.71	1,660.61	1,660.61	1,660.61	1,660.61
	Female	12–15 years	1.00	1.25	1.25	1.60	1,181.78	945.43	945.43	738.61
	Female	16–19 years	1.00	1.35	1.60	1.60	1,181.78	875.39	738.61	738.61
	Female	20–39 years	0.29	0.29	0.29	0.29	4,012.72	4,012.72	4,012.72	4,012.72
	Female	40–59 years	0.46	0.46	0.46	0.46	2,552.99	2,552.99	2,552.99	2,552.99
	Female	60 years and over	1.00	1.50	1.60	2.00	1,181.78	787.85	738.61	590.89
Other, low income	Male	Under age 1 year	0.99	0.99	0.99	0.99	1,191.53	1,191.53	1,191.53	1,191.53
	and	1–2 years	0.56	0.56	0.56	0.56	2,113.24	2,113.24	2,113.24	2,113.24
	female	3–5 years	0.42	0.42	0.42	0.42	2,804.99	2,804.99	2,804.99	2,804.99
	Male	6–11 years	0.18	0.18	0.18	0.18	6,612.62	6,612.62	6,612.62	6,612.62
	Male	12–15 years	0.25	0.25	0.25	0.25	4,760.85	4,760.85	4,760.85	4,760.85
	Male	16–19 years	0.28	0.28	0.28	0.28	4,214.64	4,214.64	4,214.64	4,214.64
	Male	20–29 years	0.17	0.17	0.17	0.17	6,897.10	6,897.10	6,897.10	6,897.10
	Male	30–39 years	0.21	0.21	0.21	0.21	5,679.75	5,679.75	5,679.75	5,679.75
	Male	40–49 years	0.22	0.22	0.22	0.22	5,494.95	5,494.95	5,494.95	5,494.95
	Male	50–59 years	0.28	0.28	0.28	0.28	4,236.68	4,236.68	4,236.68	4,236.68
	Male	60–69 years	0.31	0.31	0.31	0.31	3,856.90	3,856.90	3,856.90	3,856.90
	Male	70–79 years	0.65	0.65	0.65	0.65	1,828.03	1,828.03	1,828.03	1,828.03
	Male	80 years and over	0.99	0.99	0.99	0.99	1,199.60	1,199.60	1,199.60	1,199.60
	Female	6–11 years	0.18	0.18	0.18	0.18	6,411.69	6,411.69	6,411.69	6,411.69

See footnote at end of table.

Table II. Final sampling rates and base weights for National Health and Nutrition Examination Survey, 2002–2006—Con.

Race and Hispanic origin and income-sex-age sampling domain			Numerator of sampling rate[1]				Base weights by domain			
			Density stratum 1 (less than 10%)	Density stratum 2 (10%–29%)	Density stratum 3 (30%–59%)	Density stratum 4 (60% or more)	Density stratum 1 (less than 10%)	Density stratum 2 (10%–29%)	Density stratum 3 (30%–59%)	Density stratum 4 (60% or more)
	Female	12–15 years	0.25	0.25	0.25	0.25	4,635.75	4,635.75	4,635.75	4,635.75
	Female	16–19 years	0.37	0.37	0.37	0.37	3,236.31	3,236.31	3,236.31	3,236.31
	Female	20–29 years	0.11	0.11	0.11	0.11	10,526.80	10,526.80	10,526.80	10,526.80
	Female	30–39 years	0.15	0.15	0.15	0.15	7,896.05	7,896.05	7,896.05	7,896.05
	Female	40–49 years	0.17	0.17	0.17	0.17	6,808.54	6,808.54	6,808.54	6,808.54
	Female	50–59 years	0.19	0.19	0.19	0.19	6,363.46	6,363.46	6,363.46	6,363.46
	Female	60–69 years	0.25	0.25	0.25	0.25	4,745.01	4,745.01	4,745.01	4,745.01
	Female	70–79 years	0.28	0.28	0.28	0.28	4,275.58	4,275.58	4,275.58	4,275.58
	Female	80 years and over	0.37	0.37	0.37	0.37	3,168.03	3,168.03	3,168.03	3,168.03
Other, non-low income	Male	Under age 1 year	0.40	0.40	0.40	0.40	2,923.98	2,923.98	2,923.98	2,923.98
	and	1–2 years	0.24	0.24	0.24	0.24	4,981.26	4,981.26	4,981.26	4,981.26
	female	3–5 years	0.15	0.15	0.15	0.15	7,833.83	7,833.83	7,833.83	7,833.83
	Male	6–11 years	0.15	0.15	0.15	0.15	8,111.62	8,111.62	8,111.62	8,111.62
	Male	12–15 years	0.19	0.19	0.19	0.19	6,231.07	6,231.07	6,231.07	6,231.07
	Male	16–19 years	0.21	0.21	0.21	0.21	5,686.54	5,686.54	5,686.54	5,686.54
	Male	20–29 years	0.10	0.10	0.10	0.10	11,982.62	11,982.62	11,982.62	11,982.62
	Male	30–39 years	0.11	0.11	0.11	0.11	10,921.02	10,921.02	10,921.02	10,921.02
	Male	40–49 years	0.08	0.08	0.08	0.08	14,130.14	14,130.14	14,130.14	14,130.14
	Male	50–59 years	0.10	0.10	0.10	0.10	11,480.78	11,480.78	11,480.78	11,480.78
	Male	60–69 years	0.16	0.16	0.16	0.16	7,574.95	7,574.95	7,574.95	7,574.95
	Male	70–79 years	0.25	0.25	0.25	0.25	4,642.10	4,642.10	4,642.10	4,642.10
	Male	80 years and over	0.46	0.46	0.46	0.46	2,544.28	2,544.28	2,544.28	2,544.28
	Female	6–11 years	0.15	0.15	0.15	0.15	7,726.62	7,726.62	7,726.62	7,726.62
	Female	12–15 years	0.19	0.19	0.19	0.19	6,223.44	6,223.44	6,223.44	6,223.44
	Female	16–19 years	0.19	0.19	0.19	0.19	6,162.56	6,162.56	6,162.56	6,162.56
	Female	20–29 years	0.09	0.09	0.09	0.09	12,819.78	12,819.78	12,819.78	12,819.78
	Female	30–39 years	0.09	0.09	0.09	0.09	12,741.09	12,741.09	12,741.09	12,741.09
	Female	40–49 years	0.08	0.08	0.08	0.08	15,635.92	15,635.92	15,635.92	15,635.92
	Female	50–59 years	0.09	0.09	0.09	0.09	12,973.67	12,973.67	12,973.67	12,973.67
	Female	60–69 years	0.14	0.14	0.14	0.14	8,373.95	8,373.95	8,373.95	8,373.95
	Female	70–79 years	0.20	0.20	0.20	0.20	5,916.60	5,916.60	5,916.60	5,916.60
	Female	80 years and over	0.37	0.37	0.37	0.37	3,195.13	3,195.13	3,195.13	3,195.13

[1]Rates correspond to a 150% sample; sampling rates may be calculated by dividing the numerator by 1,182.

Table III. Description of interview, examination, and laboratory subsamples in National Health and Nutrition Examination Survey, 2003–2006

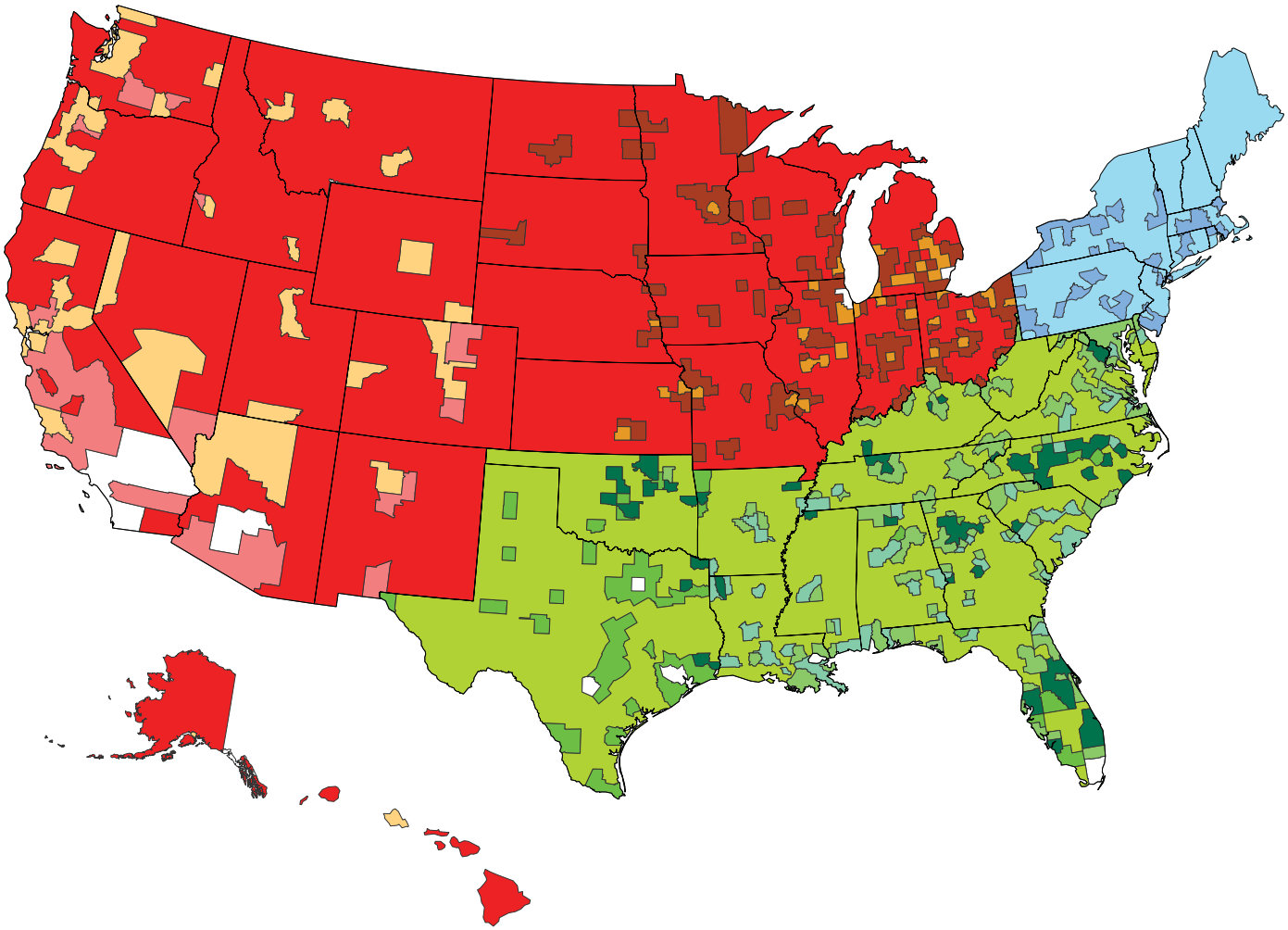
Characteristic of interest	Sample collected	Ages included	Sample fraction (of the age group)	Random groups included[1]
Composite international diagnostics (2003–2004)	Interview	20–39 years	1/2	1, 3, 5, 7, 9, 11
Hearing (2003–2004 only)	Examination	20–69 years	1/2	0, 2, 4, 6, 8, 10
Perfluorinated chemicals, nonpersistent pesticides	Blood	12 years and over	1/3	4, 6, 8, 9
Persistent organochlorine pesticides, dioxins, furans, PCBs	Blood	12 years and over	1/3	1, 2, 5, 11
Brominated flame retardants	Blood	12 years and over	1/3	0, 3, 7, 10
Volatile organic compounds	Tap water, blood	20–59 years (2003–2004)	1/2	4, 5, 6, 7, 8, 9
	Tap water	12 years and over (2005–2006)		
Perchlorate (2005–2006 only)	Tap water	12 years and over	1/2	0, 1, 2, 3, 10, 11
Non-persistent pesticides including organophosphate pesticide metabolites, iodine	Urine	6 years and over	1/3	4, 6, 8, 9
Heavy metals, speciated arsenic, mercury	Urine	6 years and over	1/3	1, 2, 5, 11

[1]Each group is a random 1/12 sample.

Table IV. Variables used in the formation of nonresponse adjustment cells for weighting all samples from 1999 to 2006

Variables considered for nonresponse	Order and categories of variables cross-classified to form nonresponse adjustment cells				
	1999–2000	2001–2002	2003–2004	2005–2006	1999–2002
Interview weights					
Age in years	1: 0–5, 6–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over
Race and Hispanic origin	2: Mexican American, black, other	2: Mexican American, black, other	6: Mexican American, black, white, other	6: Mexican American, black, other	2: Mexican American, black, other
Gender	3: Male, female	3: Male, female	4: 20–59 only	4: 20–59 only	3: Male, female
Household size	. . .	4: 1–2, 3–4, 5–6, 7 or more	2: 1–2, 3–4, 5–6, 7 or more	2: 1–2, 3–4, 5–6, 7 or more	4: 1–4, 5–6, 7 or more
Pregnancy status	. . .	5: 15–19 only	5: 15–19 only	5: 15–19 only	5: 15–19 only
Urbanicity of study location	. . .	. . .	3: Urban, suburban, rural	3: Urban, suburban, rural	. . .
Examination weights					
Age in years	1: 0–5, 6–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over	1: Under age 1, 1–5, 6–11, 12–19, 20–59, 60 and over
Race and Hispanic origin	5: Mexican American, black, other	4: Mexican American, black, other	5: Mexican American, black, other	5: Mexican American, black, other	4: Mexican American, black, other
Gender	6: 20 and over only	6: 20 and over only	6: Under age 1, 20 and over only	6: Under age 1, 20 and over only	6: 20 and over only
Household size	3: 1–2, 3–4, 5–6, 7 or more	3: 1–2, 3–4, 5–6, 7 or more	4: 1–2, 3–4, 5–6, 7 or more	4: 1–2, 3–6, 7 or more	3: 1–2, 3–4, 5–6, 7 or more
Pregnancy status	. . .	. . .	. . .	. . .	. . .
Highest education level of the household reference person or spouse	2: Less than high school, high school, more than high school	2: Less than high school, high school, more than high school	3: Less than high school, high school, more than high school	3: Less than high school, high school, more than high school	2: Less than high school, more than high school
Self-reported health status	4: Excellent, very good, good, fair, poor, unknown	5: Excellent or unknown, very good, good, fair, poor	7: Excellent or very good, good or fair, poor or unknown	7: 60 and over only; Excellent or very good, good or fair, poor or unknown	5: Under age 6; Excellent or unknown, very good, good, fair, poor 6 and over; Excellent, very good, good,fair, poor,unknown
Level of activity limited	. . .	7: 60 and over only; yes, no, unknown	8: 20 and over only; yes or unknown, no	8: 20 and over only; yes, no or unknown	7: 60 and over only; yes, no, unknown
Urbanicity of study location	. . .	. . .	2: Urban, suburban, rural	2: Urban, suburban, rural	. . .

. . . Category not applicable.



Major stratum	Description[1]
1	Northeastern PSUs with less than 7.8% of the population in poverty
2	Northeastern PSUs with more than 7.8% of the population in poverty
3	Nonmetropolitan Midwestern and Western PSUs
4	Metropolitan Midwestern PSUs with population less than 9.8% black
5	Metropolitan Midwestern PSUs with population more than 9.8% black
6	Nonmetropolitan Southern PSUs
7	Metropolitan Southern PSUs with population less than 1.4% Mexican American and: - Less than 21% black; or, - Between 21% and 27.6% black and less than 10.8% of the population in poverty
8	Metropolitan Southern PSUs with population less than 1.4% Mexican American and: - More than 27.6% black; or, - Between 21% and 27.6% black and more than 10.8% of the population in poverty
9	Metropolitan Southern PSUs with population between 1.4% and 3.8% Mexican American
10	Metropolitan Southern PSUs with population more than 3.8% Mexican American
11	Metropolitan Western PSUs with population less than 14% Mexican American
12	Metropolitan Western PSUs with population more than 14% Mexican American
	PSUs selected with certainty

[1]Threshold values are approximate.
NOTE: PSU is primary sampling unit. PSUs selected for 2007 were not used.
SOURCE: CDC/NCHS, 1990 and 2000 censuses (data for major strata); variables used include census region (Northeast, Midwest, South, and West) and metropolitan statistical area status.

Figure. Major strata formed for selection of the 2002–2007 PSUs

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 5412
Hyattsville, MD 20782

1–800–232–4636
E-mail: cdcinfo@cdc.gov
Internet: http://www.cdc.gov/nchs

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES

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

CS228636
DHHS Publication No. (PHS) 2012–1355, Series 2, No. 155