Vital and Health Statistics

Sample Design: Third National Health and Nutrition Examination Survey

Series 2:
Data Evaluation and Methods Research
No. 113

This report presents a detailed description of the sample design for the Third National Health and Nutrition Examination Survey, 1988–94, including a brief description of research that led to the choice of the final design. The National Health and Nutrition Examination Survey (NHANES) is one of the major surveys of the National Center for Health Statistics, Centers for Disease Control. Information on the health and nutritional status of the noninstitutionalized population of the United States is collected through the NHANES household interviews and standardized physical examinations.

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

Hyattsville, Maryland
September 1992
DHHS Publication No. (PHS) 92-1387

National Center for Health Statistics

Manning Feinleib, M.D., Dr.P.H., *Director*

Jacob J. Feldman, Ph.D., *Associate Director for Analysis and Epidemiology*

Gail Fisher, Ph.D., *Associate Director for Planning and Extramural Programs*

Peter L. Hurley, *Associate Director for Vital and Health Statistics Systems*

Robert A. Israel, *Associate Director for International Statistics*

Stephen E. Nieberding, *Associate Director for Management*

Charles J. Rothwell, *Associate Director for Data Processing and Services*

Monroe G. Sirken, Ph.D., *Associate Director for Research and Methodology*

David L. Larson, *Assistant Director, Atlanta*

Office of Research and Methodology

Monroe G. Sirken, Ph.D., *Associate Director*

Kenneth W. Harris, *Special Assistant for Program Coordination and Statistical Standards*

Lester R. Curtin, Ph.D., *Chief, Statistical Methods Staff*

James T. Massey, Ph.D., *Chief, Survey Design Staff*

Andrew A. White, Ph.D., *Chief, Statistical Technology Staff*

Division of Health Examination Statistics

Robert S. Murphy, *Director*

Kurt Maurer, Ph.D., *Deputy Director*

Ronette R. Briefel, Dr.P.H., *Coordinator for Nutrition Monitoring and Related Research*

Clifford Johnson, *Chief, Nutrition Statistics Branch*

Katherine M. Flegal, Ph.D., *Chief, Medical Statistics Branch*

Christopher Sempos, Ph.D., *Chief, Longitudinal Studies Branch*

Vicki L. Burt, *Chief, Survey Planning and Development Branch*

Jean Findlay, *Acting Chief, Survey Operations Branch*

Robert Krasowski, *Chief, Computer Systems and Programming Branch*

Contents

List of text tables

List of appendix tables

Symbols

- - -	Data not available
. . .	Category not applicable
–	Quantity zero
0.0	Quantity more than zero but less than 0.05
Z	Quantity more than zero but less than 500 where numbers are rounded to thousands
*	Figure does not meet standard of reliability or precision
#	Figure suppressed to comply with confidentiality requirements

Sample Design: Third National Health and Nutrition Examination Survey

by Trena M. Ezzati, James T. Massey, Office of Research and Methodology, National Center for Health Statistics; Joseph Waksberg, Adam Chu, Westat, Inc.; and Kurt R. Maurer, Division of Health Examination Statistics, National Center for Health Statistics

Introduction

The National Center for Health Statistics (NCHS) conducts several large-scale national health surveys. The National Health and Nutrition Examination Survey (NHANES), designed to assess the health and nutritional status of the noninstitutionalized population of the United States, consists of adult, youth, and family questionnaires followed by standardized physical examinations in specially equipped mobile examination centers (MEC's). The Third National Health and Nutrition Examination Survey (NHANES III) is the seventh in a series of surveys using health examination procedures that have been conducted since 1960 by NCHS. The target populations, the sample designs, and the data collection procedures for the previous health examination surveys—NHANES I, NHANES II, and a special survey of the Hispanic population (Hispanic HANES)—have been described in previous reports (1–8).

The target population of NHANES III is the civilian noninstitutionalized population 2 months of age and older. The survey is being conducted from 1988 through 1994 and includes a sample of approximately 40,000 persons. The household interview includes demographic, socioeconomic, dietary, and health history questions; the examination component consists of examinations by a physician, a dentist, and health technicians. In addition, a home examination consisting of an abbreviated set of physical measurements is offered to persons who are unable or unwilling to travel to the MEC for a complete examination (9). The interviews and examinations are performed by permanent staff employed by Westat, Inc., the data collection contractor for NHANES III.

As in previous NHANES, this survey has the following major goals:

- To produce national population health parameters;
- To estimate the national prevalence of selected diseases and disease risk factors;
- To investigate secular trends in selected diseases and risk factors;
- To contribute to the understanding of disease etiology; and
- To investigate the natural history of selected diseases.

A list of the major target conditions of NHANES III is shown in table I and the examination components of the survey are shown in tables II and III.

Previous NHANES have included only persons aged 6 months through 74 years. Because a growing proportion of the U.S. population consists of older Americans who experience greater morbidity and disability, NCHS imposed no upper age limit for the NHANES III. Also, because of the need for better information on the growth and development of young children, infants 2–5 months old are included in the NHANES III for the first time. Older persons and children are oversampled so that estimates of their health status can be made with acceptable precision.

Although not excluded from the target sample, small numbers of black and Mexican-American persons were included in previous NHANES. Therefore, reliable estimates of their health and nutritional status could not be produced for subgroups of these domains, that is, by age, sex, or other important demographic or socioeconomic breakdowns. To resolve this problem, NHANES III was designed to include a large sample of both the black and the Mexican-American populations so that reliable estimates of health and nutritional characteristics can be produced for these two largest minority groups of the U.S. population.

This report describes the broad design requirements for NHANES III, the research undertaken to develop the sample design specifications, the major research, the estimation procedures, and the methods used for estimating sampling errors. Documentation of the survey content, procedures, and methods to assess nonsampling errors are reported in another publication (9). Much of this report is based on survey requirement documents and the final sample design report prepared by Westat, Inc., as part of an NHANES III methods research contract (10).

Design specifications

Survey objectives

A desirable first step in designing a survey is to define the analytical objectives. As in the previous NHANES, a primary purpose of NHANES III is to produce a broad range of descriptive health and nutrition statistics for sex, race, ethnic, 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 III was designed to produce cross-sectional data and because respondents will be followed over time for future interviews and/or examinations, a set of cross-sectional and longitudinal objectives was developed. These objectives are:

- To produce estimates of means and proportions with a reasonable level of precision for a broad range of health and nutritional variables by sex, race, ethnic, and age subgroups of the civilian noninstitutionalized U.S. population;
- To determine differences between subgroup estimates with specified type I and type II errors;
- To monitor secular trends in health and nutritional status for subgroups of interest; and
- To investigate the etiology and natural history of selected diseases through followup of a cohort of initial respondents.

Domain and precision specifications

A primary interest of NHANES III is to estimate with acceptable precision the health and nutritional status of subgroups of the population. The subdomains for which separate analyses are expected to be carried out in NHANES III are shown in table A. The set of subdomains consists of sex-age groups for the largest race-ethnic subgroups in the U.S. population. The analytical domains in table A consist of the age groups shown separately for males and females. Therefore, 52 subdomains—twice the number of age groups are shown.

The sample for NHANES III was designed with specified precision for each of the major subdomains in table A. The goal is to have approximately equal precision for each of the analytical domains. The minimum precision requirements are:

- A prevalence statistic of 10 percent should have a relative standard error (RSE) less than 30 percent; and
- Differences of at least 10 percent in health or nutrition statistics between any two subdomains should be detected with a type I error of no more than 0.05 and a type II error of no more than 0.10.

Table A. Analytical subdomains classified by race-ethnicity and age: Third National Health and Nutrition Examination Survey, 1988–94

Black	*White and all other*	*Mexican-American*[1]
2–35 months	2–11 months	2–35 months
	12–35 months	
3–5 years	3–5 years	3–5 years
6–11 years	6–11 years	6–11 years
12–19 years	12–19 years	12–19 years
20–39 years	20–29 years	20–39 years
	30–39 years	
40–59 years	40–49 years	40–59 years
	50–59 years	
60 years and over	60–69 years	60 years and over
	70–79 years	
	80 years and over	

[1]Mexican-Americans can be any race.

NOTE: The analytical domains are for males and females separately.

To meet the predesignated precision requirements, the sample size for each of the defined subdomains for all groups was determined to be 560 or greater. For the subdomains with considerable oversampling in Mexican-American density strata, the minimum of 560 was increased to compensate for the design effect introduced by the variability in sampling rates among density strata. In addition, the total examined sample size for both black and Mexican-American persons is required to be 9,000, with 12,000 for white and all other persons (table B). The examined sample sizes were inflated by about one-third to account for expected nonresponse to the examination portion of the survey to determine the total sample sizes for the survey. The total sample sizes by race and ethnicity are also shown in table B.

Superimposed on the minimum cell size of 560 was the requirement that the total examined number of infants and young children (separately for males and females) in the 3 age groups—under 1 year, 1–2 years, and 3–5 years—needed to be at least 1,000 to provide adequate sample sizes for updating the growth charts for NHANES III.

Table B. Expected total and examined person sample sizes by race-ethnicity: Third National Health and Nutrition Examination Survey, 1988–94

Race-ethnicity	*Total*	*Examined persons*
Total	40,000	30,000
Black	12,000	9,000
White and all other	16,000	12,000
Mexican-American[1]	12,000	9,000

[1]Mexican-Americans can be any race.

Operational requirements

NHANES is unique because the examination component of the survey involves a number of logistical and cost concerns that also have to be considered in the design of the survey. Therefore, the following operational requirements were imposed upon the design of the survey:

- The number of sample persons selected at each stand (survey location) should be between 300 and 600, with an average of approximately 450, yielding an expected 340 examined persons. Research has shown this number to be the approximate optimum number to give as many primary sampling units (PSU's) as possible, while keeping the sample size in each area large enough to justify the costs associated with moving and setting up the mobile examination centers (MEC's);
- The minimum time to complete field work at any stand is 4 weeks; and
- The data collection period should be 6 years. However, because it is not desirable to wait the full 6 years to have updated data since the last national survey (NHANES II, 1976–80), the first 3-year period (phase 1) should contain a national probability sample of the eligible population so that some estimates of health and nutrition can be produced when the first 3 years of data collection are completed. (The subdomain estimates produced from phase 1 will not be as detailed as those for the full 6-year survey because the analytical data requirements apply only to the full 6-year data collection period.)

In addition to these operational requirements, another factor considered in the final design was how to select as large an average number of sample persons per household as possible, thereby maximizing the response rates and reducing screening costs. This was also an important design consideration because previous experience with both NHANES and Hispanic HANES indicated that response rates increase when a large sample of persons are selected within households. One of the factors thought to be responsible for the increased response rates in multiple-sample person households is that each person is given a remuneration for his or her time and participation.

Methods research

In planning for NHANES III, new design features for the survey and a broad reexamination of the basic methodology for sampling and respondent contact were evaluated. First, the use of NHANES III as a baseline for a long-term longitudinal study of adults and children as well as a cross-sectional study created a broad range of analytical possibilities for NHANES and also created several design and cost issues. Second, because previous NHANES focused on statistics for the total U.S. population, relatively small numbers of black and Mexican-American persons were included. Therefore, estimates of their health and nutritional status for separate age-sex subdomains were often inadequate due to their unreliability.

NCHS recognized the need for information on subgroups of the population when the special Hispanic HANES was conducted in 1982–84 (8). Including adequate numbers of black and Mexican-American persons in NHANES III serves to update the Hispanic HANES and enhances comparison of the health of minorities with that of white and all other persons, thus allowing the investigation of factors related to differential health status. Therefore, the primary focus of the sample design research was on the development of efficient procedures to oversample selected subgroups of the U.S. population so that reliable health and nutritional statistics are available.

To evaluate these design issues, an 18-month methodological contract was awarded to Westat, Inc., in May 1986. The development of an efficient sample design for NHANES III was one of the major tasks carried out as part of this contract. A more detailed discussion of the other methodological issues examined as part of this research contract is described in another report (11).

The original design specifications required separate estimates for black, Mexican-American, and Puerto Rican populations in addition to white and all other populations. However, research indicated that it would be costly to locate a probability sample of the Puerto Rican population of sufficient size to provide useful data. Because of cost constraints and the limit on the number of persons that could be examined in the 6-year time period, separate estimates for Puerto Rican persons were not included in the final design specifications. Puerto Rican persons, if selected for the sample, are included in the "white and all other" estimation cell.

Similarly, the cost and feasibility implications of producing separate estimates for the increasing numbers of Asian-Americans in the United States were also evaluated as part of the design research. As in the case of Puerto Ricans, the cost appeared to be very high and oversampling of Asian-Americans was not considered in the final design.

The three major options (including combinations of them) evaluated as part of the design research were:

- An independently selected sample of the civilian noninstitutionalized population;
- A sample of interviewed and examined persons from prior NHANES and a supplementary sample of the current civilian noninstitutionalized population; and
- A sample of either persons or addresses selected for the National Health Interview Survey (NHIS), another major survey conducted by the NCHS, supplemented by an independent sample of the current civilian noninstitutionalized population.

As described earlier, the sample design was required to produce statistically reliable estimates of means and proportions for a broad range of health and nutritional variables by age, sex, race, and ethnicity. To accomplish this objective, a number of related issues were addressed. The following key issues were examined:

- The subdomains that would need to be oversampled to meet the precision requirements and a determination of the extent of oversampling required;
- The data collection costs associated with each of the three design options;
- The treatment of prior nonrespondents and persons who died or moved if a linked NHANES or NHIS design were employed;
- The approximately optimum design of a supplementary sample, if such a sample were needed;
- The major features of an independent sample design including:
  - the number of PSU's to be selected;
  - the approximate optimum segment size for the area sample; and
  - the within-PSU selection procedures that would efficiently produce the desired sample sizes.

The major conclusions drawn from the research formed the basis of the final NHANES III sample design. They are:

1. The NHANES or NHIS samples for prior years are not efficient methods of establishing the sample for NHANES III. An independent single sampling frame, consisting of an area sample, should be used. Area samples should be used both in the selection of PSU's and of households. An area sample supplement to the National Health Interview Survey (NHIS) would be necessary because a single year of

NHIS is not large enough to provide the needed sample for most of the subdomains, especially if only part of the sample located in about 89 stands is used. (Only about 89 stands can be included in NHANES III if the total target sample size is about 40,000 persons and if average sample size per PSU is to be 450 persons.) Similarly, an area supplement would be needed if an earlier NHANES was used as the sampling frame because of the inadequate sample size of black and Mexican-American persons and some of the age-sex subdomains for the "white and all other" category. Therefore, the cost and complexity of working with multiple frames appeared to negate the value of the alternative frames.

2. The Mexican-American sample should be viewed as a supplementary sample in the moderate to high dense Mexican areas, superimposed on a basic self-weighting sample that is large enough to satisfy the sample size requirements for all the black and white and other sex-age subdomains.
3. The supplementary sample should use geographical stratification with strata defined on the basis of the proportion of Mexican-American persons in the block group-enumeration districts in the 1980 census. Optimum allocation of the supplementary sample to the geographic strata should be used, taking costs and variances into account.
4. To establish the sampling and subsampling rates, sex-age-race-ethnicity subdomains that need approximately the same number of screened households to obtain the target sample should be grouped with a single sampling rate used for all subdomains within a group. This will significantly reduce the number of separate sampling rates that have to be applied.
5. The optimum allocation sampling rates for the various Mexican-American strata should be set to provide the number of households to be screened as required for the rarest sex-age subdomain for Mexican-Americans. The subsamples of the screened households needed for the less rare subdomains are to be obtained by progressively cutting back the rates in the most dense strata to equal the rate in the next remaining stratum. This will minimize the variation in weights for each subdomain, within the limits of what can be done without increasing the level of screening.
6. The procedure for subsampling persons and households from the screening sample should be carried out by subdividing the screening sample into random, although unequal, subsets. In one random group, all persons in the screening sample are designated as sample persons. The sampling rate reflected by this random group is the rate applicable to the largest subdomain in each race/ethnicity group. In a second random group, all persons except those in the largest subdomain are designated as sample persons. In the third random group, all except the two most common subdomains are designated, and so forth. The number of random subsets and their sizes are calculated to ensure getting the desired sample size. This procedure maximizes the average number of sample persons per household while adhering to the sampling rates required for each subdomain.
7. The measures of size used in the selection of both PSU's and segments will be $\sum P_{ji} M_j r_i$, where P_{ji} is the proportion of the population of PSU j (or segment j), in the ith subdomain, M_j is the total population of PSU j (or segment j) and r_i is the sampling rate in subdomain i. The summation is over the subdomains. For measures of size, the subdomains will be collapsed to the three race-ethnic groups. The reason for this is that the race-ethnicity composition of an area is fairly stable over time but the age-sex distributions can vary greatly, particularly because of the aging of the population since the most recent decennial census.
8. A two-PSU-per-stratum sample design should be used for noncertainty PSU's in NHANES III. The two PSU's should be selected with probabilities proportional to size (PPS) and without replacement. The measures of size should be the ones described above. After the selection of the noncertainty PSU's, one PSU should be assigned to the first 3-year period (phase 1) and the other one to the second 3-year period (phase 2).
9. No special action will be taken to supplement the sample of low-income persons since oversampling poverty areas on the basis of geography is not very efficient and any other method is expensive. However, it is likely that even without oversampling for persons below the poverty level, there will be enough black and Mexican-American persons and total persons in the sample below the poverty level to permit a reasonable amount of analysis for collapsed sex-age subdomains.
10. The sample design should be reviewed at the end of the first year of field operations, and periodically thereafter, to ascertain whether modifications are necessary. It is particularly important that this be done in 1991, when the 1990 census data become available.

In addition to these major design decisions, the research also provided much of the data needed for the detailed specification of the sample design, for example, stratification variables for the selection of PSU's, the optimum segment size, the method of calculating measures of size for both PSU's and segments, the sampling rates to be used in the various density strata, and the level of screening required.

Sample design

Summary

The general structure of the NHANES III sample design is the same as that of the previous National Health and Nutrition Examination Surveys (4–7). Each of these surveys used a stratified multistage probability design. A summary of the major design parameters for the two previous NHANES and the special Hispanic HANES, as well as NHANES III, is given in table C. The NHANES III sample has been designed to be self-weighting within a PSU for subdomains and almost self-weighting nationally for each of the subdomain groups (but not for the total population).

The NHANES III sample represents the total civilian noninstitutionalized population 2 months of age or older in the 50 States of the United States. The first stage of the design consists of selecting a sample of 81 primary sampling units (PSU's), which are mostly individual counties. In some cases, adjacent counties were combined to keep PSU's above a minimum size. The PSU's were stratified and selected with probability proportional to size (PPS). Thirteen large counties (strata) were chosen with certainty (probability of one). For logistical and operational reasons, these 13 certainty PSU's were divided into 21 stands (survey locations). After the 13 certainty strata were designated, the remaining PSU's were grouped into 34 strata, and two noncertainty PSU's were selected per stratum. The selection was done with PPS and without replacement. Each noncertainty PSU is also referred to as a "stand." The NHANES III can thus be considered as consisting of 81 PSU's or 89 stands.

The 89 stands in the sample were randomly subdivided into two sets, one consisting of 44 stands and the other consisting of 45 stands. One set of stands was allocated to the first 3-year survey period (1988–91) and the other set allocated to the second 3-year period (1991–94). Therefore, unbiased estimates (from the point of view of sample selection) of health and nutritional characteristics can be independently produced for both phase 1 and phase 2.

For most of the sample, the second stage of the design consists of area segments comprised of city or suburban blocks, combinations of blocks, or other area segments in places where block statistics were not produced in the 1980 census. In the first phase of NHANES III, the area segments are used only for a sample of persons who lived in housing units built before 1980. For units built in 1980 and later, the second stage consists of sets

Table C. Health and Nutrition Examination Surveys by selected sample design parameters

Parameter	*NHANES I*	*NHANES II*	*Hispanic HANES*	*NHANES III*
Age of civilian noninstitutionalized target population	1–74 years	6 months–74 years	6 months–74 years	2 months and over
Geographical areas	United States (excluding Alaska and Hawaii)	United States (including Alaska and Hawaii)	Southwest for Mexican-Americans; NY, NJ, CT for Puerto Ricans; Dade County, FL for Cubans	United States (including Alaska and Hawaii)
Average number of sample persons per household	1	1	2–3	2–3
Number of survey locations	100	64	17 in Southwest; 9 in NY, NJ, CT; 4 in Dade County	89
Domains for oversampling	Low income; children ages 1–5 years; women ages 20–44 years; persons ages 65 years and over	Low income; children ages 6 months–5 years; persons ages 60–74 years	Dade County: persons ages 6 months–19 years and 45–74 years; Southwest and NY, NJ, and CT: persons ages 6 months–19 years and 45–74 years	Children 2 months–5 years; persons 60 years and over; Mexican-American and black persons
Sample size	32,000	28,000	16,000	40,000 (expected)
Examined sample size	24,000	20,000	12,000	30,000 (expected)
Years covered	1971–74	1976–80	1982–84	1988–94

NOTE: NHANES is the National Health and Nutrition Examination Survey.

of addresses selected from building permits issued in 1980 or later. These are referred to as new construction segments. In the second phase, 1990 census data and maps are being used to define the area segments. Because the second phase follows within a few years of the 1990 census, new construction will not account for a significant part of the sample and the entire sample comes from the area segments.

The third stage of sample selection consists of households and certain types of group quarters, such as dormitories. All households and eligible group quarters in the sample segments are listed and a subsample is designated for screening to identify potential sample persons. The subsampling rates enable production of a national, approximately equal, probability sample of households in most of the United States, with higher rates for the geographic strata with high Mexican-American concentrations. Within each geographic stratum, there is an approximate equal probability sample of households across all 89 stands. The screening rate in each stratum is designed to produce the desired number of sample persons for the rarest age-sex domain in the race-ethnic group defining the geographic stratum.

Persons within the sample of households or group quarters are the fourth stage of sample selection. All eligible members within a household are listed, and a subsample of individuals is selected based on sex, age, and race-ethnicity. The definitions of the sex, age, race-ethnic classes, subsampling rates, and designation of potential sample persons within screened households are developed to provide approximately self-weighting samples for each subdomain within geographic strata and simultaneously to maximize the average number of sample persons per sample household. Experience in previous NHANES indicated that this increased the overall participation rate.

Although the exact sample sizes will not be known until data collection is completed, estimates have been made. A summary of the expected sample sizes at each stage of the design is shown in the following table:

Number of PSU's	81
Number of stands (survey locations)	89
Number of segments	2,138
Number of households to be screened	106,000
Number of households with persons	20,000
Number of persons	40,600
Number of interviewed sample persons	35,000
Number of examined sample persons	30,100

Stratification and selection of PSU's

The sampling frame for NHANES III was composed of all of the counties, parishes, and independent cities in the United States (including Alaska and Hawaii), all of which are referred to as "counties" for convenience. From these counties, primary sampling units (PSU's) were formed. Most PSU's consist of single counties although a few were made up of small groups of contiguous counties. The PSU's for NHANES III are defined as individual counties. This definition reduces the amount of travel necessary to visit the mobile examination center (MEC) for the examination component of the survey and achieves as high a response rate as possible. Combinations of counties were used only where counties were so small, in terms of population, that their probabilities of selection would have been lower than what is required for some of the domains. If selected for the sample, they would introduce considerable variability in the sampling weights. Consequently, these small counties were combined with one or more adjacent counties to form more efficient sampling units. For the same reason, the independent cities in Virginia were combined with one or more nearby counties to define the PSU's for sampling. Of the approximately 3,100 counties and county equivalents in the United States, 2,812 PSU's (most of which consisted of individual counties) were defined for NHANES III. After stratification, the sample of PSU's was selected from the 2,812 PSU's.

The sampling frame, measures of size, and stratification variables used 1980 census data. The frame of PSU's was constructed by merging information from two data sources. The first of these was the Bureau of the Census STF-1C file, which contained information for each county or county-equivalent. The information used for the PSU sampling was: (a) region; (b) metropolitan status; (c) 1980 population; and (d) 1980 population for Mexican-American and black persons. The second source was the Bureau of the Census file "Population Estimates by County with Components of Change, 1981–1985." Data from the latter source were used to update the 1980 population figures in the STF-1C file as discussed in the next section on calculation of PSU measures of size. This information was useful for stratification of PSU's and for determining the probabilities of selecting the sample of PSU's.

Calculation of measures of size

Multistage area samples selected with probability proportional to size (PPS) generally use a single variable as the measure of size, for example, total population or total housing units. This was the practice followed in NHANES I and II. However, when subdomains are to be sampled at different rates, but a self-weighting sample is desired for each subdomain, such measures of size can result in highly variable workloads. This applies to the sample selection at both the PSU and the segment level.

For example, assume that there are L subdomains that are to be sampled at rates $r_1, r_2, \ldots, r_L$. In the jth PSU, let the proportions of the population in the various subdomains be denoted by $P_{j1}, \ldots, P_{jL}$, where $P_{j1} + P_{j2} + \ldots + P_{JL} = 1$ for each j. If the measures of size of the PSU were N_h, where N_h is the total population of the PSU, then the populations in the subdomains in the jth PSU are $P_{h1}N_h, \ldots P_{hL}\,N_h$, respectively.

With these measures, the probability of selection of a PSU will be kN_h. For self-weighting samples, the within-PSU rates for subdomain i need to be proportional to r_i/kN_h. Thus, the number of sample persons in the ith subdomain is $(P_{hi}N_h)(r_i/kN_h) = r_iP_{hi}/k$. The total sample size in the PSU will be as follows:

$$\sum_i r_i P_{hi}/k$$

The total sample size in a PSU will thus depend on P_{hi}, the percentage distribution of the subdomains within the PSU. Unless the P_{hi} is approximately the same in all PSU's, the workloads will vary. In the case of the particular subdomains established for NHANES III, the distributions by sex and age will probably not create any problems. Age and sex distributions are reasonably consistent in large areas such as counties. However, race and ethnicity vary considerably. The percentage, for example, of Mexican-Americans and black persons will range from close to zero in some counties to almost 100 percent in others.

For logistical and operational reasons, workloads per noncertainty PSU must be kept within a fairly narrow range. Using total population or households to establish the PSU probabilities of selection would create large differences between the desired uniform workloads and the actual sample sizes. Subsampling or sample supplements in each PSU would be required to bring the sample size in each PSU in conformance with the desired numbers. This would add substantially to the variances. Furthermore, it would introduce uncertainty into the total sample sizes by subdomain.

Using the quantity

$$M_h = \sum_i P_{hi} N_h r_i$$

as the sampling measure of size for PSU h avoids these complications. Under PPS sampling, the probability of selecting a PSU is then $k\Sigma P_{hi} N_h r_i$. Within a PSU, the sample size in the ith subdomain will be $(P_{hi} N_h r_i)/(k\Sigma P_{hi} N_h r_i)$. The total sample size in the PSU over all subdomains is then $(\Sigma P_{hi} N_h r_i)/(k\Sigma P_{hi} N_h r_i) = 1/k$ and is thus independent of both the N_h and the P_{hi}.

The formula for the PSU measures of size given above was somewhat oversimplified. For example, the NHANES III design included provisions for stratifying areas by concentration of Mexican-Americans, and sampling households within strata at varying rates depending on the race-ethnic distribution within the stratum. Consequently, the sampling rates will actually depend on the density stratum, as well as on the sex-age and race-ethnicity subdomains. Specifically, in terms of the different subdomains and minority strata, the measure of size assigned to PSU h was:

$$M_h = \sum_i \sum_k \sum_l r_{ikl} N_{hikl}$$

where
- h = PSU
- i = Mexican-American density stratum (as defined in the next section)
- k = race-ethnicity subdomain
- l = sex-age subdomain
- N_{hikl} = estimated 1990 U.S. population in the (k,l)th subdomain in density stratum i in PSU h
- r_{ikl} = U.S. sampling rate in density stratum i for the (k,l)th race-ethnicity-sex-age subdomain.

The exact values of the N_{hikl} were not known at the time the PSU's were selected. Hence the required population counts were estimated from 1980 population figures and adjusted by the most recent Bureau of the Census projections of the county's population. Thus, for NHANES III, the measure of size assigned to PSU h was:

$$M_h = \sum_i \sum_k \sum_l r_{ikl} N_{hikl} = \frac{C_h}{C'_h} \sum_i \sum_k C'_{hik} \sum_l r_{ikl} \frac{C_{\cdot\cdot kl}}{C_{\cdot\cdot k}}$$

where
- C'_h = 1980 population for PSU h
- C'_{hk} = 1980 population count for race-ethnicity subdomain in PSU h
- C_h = most recent (1985) population count for PSU h
- C'_{hik} = 1980 population count for race-ethnicity subdomain k in minority stratum i in PSU h
- $C_{\cdot\cdot kl}$ = projected 1990 total U.S. population count for race-ethnicity-sex-age subdomain (k,l)
- $C_{\cdot\cdot k}$ = projected 1990 total population count for race-ethnicity subdomain k.

A further simplification was necessary because the values of C'_{hik} were not available. Assuming that the distribution of a minority among geographical strata is identical among PSU's, then $C'_{hik} = C'_{hk} P_{ik}$, where P_{ik} is the U.S. proportion in the ith stratum for the kth race-ethnicity subdomain. Hence, the measure of size for PSU h was calculated as follows:

$$M_h = \frac{C_h}{C'_h} \sum_k C'_{hk} A_k$$

where

$$A_k = \sum_{i,l} P_{ik} r_{ikl} \frac{C_{kl}}{C_k}$$

The derivation of the r_{ikl} used to compute the measures of size is described in the section on selection of segments.

Minimum measures of size

The selection probability of a PSU determined the maximum rate at which persons residing in that particular PSU could be selected for NHANES III. If the measure of size of a PSU was too small, the required sampling rates for some subdomains could not be achieved. Consequently, special weighting procedures would be required for these PSU's, and the resulting variability in weights would increase sampling errors.

To ensure that all required sampling rates could be achieved, the measure of size M_h of a PSU had to satisfy the following inequality:

$$M_h \geq M r_{ikl}/2$$

where h denotes the PSU, i denotes the Mexican-American density stratum, k denotes race-ethnicity, l denotes the sex-age subdomain, M is the measure of size of the stratum in which the PSU is located, and r_{ikl} is the sampling rate for the (k,l)th subdomain in density stratum i. The factor of 2 in the right-hand side of the above inequality reflects the fact that two PSU's were selected per stratum. Counties for which $M_h < M r_{ikl}/2$ were combined with a neighboring county to form PSU's satisfying the minimum size requirement.

The procedures used to construct the frame required establishing the set of PSU's prior to stratification, thus the value of M was not known at the time decisions on combining PSU's were made. Therefore some approximations were necessary. Also, the values of r_{ikl} depend on the density strata, and many PSU's do not have high-density strata. For these PSU's, it was unnecessary to have criteria that guarded against situations that will not occur. The data files used to select the PSU's did not provide direct information on whether all density strata occurred in a county, or whether some strata were not present and thus could be ignored in the inequality above.

Because the strata sizes were to be made as close to equal as possible, a good approximation to M was obtained by assuming they were equal. Let m_c represent the number of certainty stands in the certainty PSU's. The number of noncertainty PSU's is then $89 - m_c$. (Note: m_c is the number of stands (survey locations), in the certainty PSU's, not the number of PSU's. Some certainty PSU's contain multiple stands.)

If M_{nc} is the total measure of size of the noncertainty PSU's, the value of M used to check the PSU measures of size was:

$$M = 2M_{nc} / (89 - m_c)$$

The maximum values of r_{ikl} (screening sampling rate) for the six density strata (defined by proportion of Mexican-American population) are referred to as $\hat{r}_i$, and are shown in table D. The highest value is 6/930, for stratum 6 (the highest density Mexican-American stratum). The next highest value is 5/930 for stratum 5, and so forth.

The goal was to avoid combining counties unnecessarily. Increasing the geographic area of a PSU was likely to make a high participation rate more difficult to achieve due to the distance to the mobile examination center (MEC). The approach taken was to keep individual counties as PSU's when they would not require weighting, or, at most, only a small amount. However, combinations were made when the alternative was to impose large weights. The procedure avoided using the high rates required by the minority density strata for the minimum measures unless there was a reasonable chance that the density strata would actually occur.

The value of r_{ikl} that was used for the minimum measure of a county depended on the proportion of the Mexican-American population in the county. The value used for a county is shown in the following table:

Percent of Mexican-American persons in county	*Value of* r_{ikl}
Less than 1	1/930
1–2 9	2.5/930
3	5/930

As a result of these rules, the minimum was smaller than it should be for a few counties. Using these rules, about 300 counties were found to have measures of size that were too small and thus were combined with neighboring counties. The effect on the sample was minimal.

Table D. Screening sampling rates and number of screened households, by density stratum: Third National Health and Nutrition Examination Survey, 1988–94

Density stratum	*Screening sampling rate* ($\hat{r}_i$)	*Number of screened households* *Including reserve*	*Excluding reserve*
Total	...	158,927	105,950
1	1/930	82,770	55,180
2	1.8/930	11,569	7,713
3	3/930	22,319	14,880
4	4/930	14,558	9,705
5	5/930	14,287	9,524
6	6/930	13,424	8,949

Selection of certainty PSU's

After assigning measures of size to each PSU, the 13 largest counties (in terms of the measure of size) were included in the sample with certainty, that is, they were designated as self-representing. The cutoff used to identify the certainty PSU's represented approximately three-eighths of the average stratum size for the noncertainty PSU's. (Since two PSU's were selected from each noncertainty stratum, a PSU can be considered as representing half a stratum. A certainty PSU should thus be close to half a stratum size or larger. The three-eighths is equivalent to three-fourths of half a stratum.)

For operational purposes, the largest certainty PSU's were subdivided so that each part would have approximately the same workload (450 sample persons) as in the noncertainty PSU's. The 13 certainty PSU's were thus converted to 21 stands. Most of the certainty PSU's consisted of single stands, as was the case for noncertainty PSU's. The additional eight stands came from three large PSU's.

The 13 certainty counties selected for NHANES III are shown in table 1, along with the number of stands (survey locations) designated for the county and the measures of size (expected sample size). Most of the certainty counties are in California or Texas, reflecting the substantial oversampling of Mexican-Americans.

Stratification and selection of noncertainty PSU's

Because there were 13 certainty PSU's designated, an additional 68 noncertainty PSU's were necessary to produce an 81-PSU sample. As indicated earlier, a two-PSU-per-stratum sample design was planned. This implied the creation of 34 strata. After selecting the certainty PSU's, the remaining noncertainty PSU's were stratified by region, within region by metropolitan status [Standard Metropolitan Statistical Area (SMSA) vs. (non-SMSA)], within metropolitan status by race-ethnicity and finally by income. Within these groups of PSU's, a total of 34 detailed strata of approximately equal aggregate size were created. The definitions of the 34 noncertainty strata are shown in table 2. Also listed in this table are the sample PSU's selected from each stratum.

To facilitate variance estimation, the noncertainty PSU's were selected by using the Durbin procedure (12) and involved the following steps:

(1) Initially, one PSU was sampled from each stratum following the usual PPS procedures. These PSU's were excluded

from the master file (frame) of the PSU's before selecting the second round of PSU's. The PSU's selected in the first round are denoted by the subscript j in the notation below.

(2) Next, within a particular stratum, Durbin probabilities P_{Di}, were computed for each PSU in the frame (other than those previously selected) as

$$P_{Di} = p_i[(1-2p_i)^{-1} + (1-2p_j)^{-1}]$$

where

p_i = probability of selecting the ith PSU in the stratum ($i \neq j$)

p_j = corresponding probability of the jth previously-selected PSU

For those PSU's that were selected in the first round, P_{Dj} was set equal to 0.

(3) The restricted frame of PSU's was then sorted by stratum and one PSU was selected with probability proportionate to P_{Di}.

The two PSU's sampled in each stratum are identified in table 2. The total of 81 sample PSU's for NHANES III is shown in the map included with this report.

Allocation of PSU's to time periods

To permit separate analyses for the two 3-year periods (1988–91 and 1991–94), as well as for the entire 6-year survey period, the sample of PSU's was randomly allocated to the two 3-year periods shown in tables 1 and 2. The allocation to the two periods was made in a way that retained as much of the original stratification as possible. Because two noncertainty PSU's were selected per stratum, one of the PSU's was randomly assigned to the first time period (phase 1), and the other was assigned to the second period (phase 2). In making the assignments to periods for the certainty PSU's, the PSU's were sequenced in a way that brought similar PSU's together (that is, by region, race-ethnicity-income class, using ascending-descending sequences), and then the PSU's were alternately assigned to the two time periods. In the three multiple-stand PSU's, half the st ınds in each PSU were assigned to each phase; when there was an odd number of stands, one of the phases was chosen at random to have an additional stand.

Selection of segments

The second stage of the design involved stratification within each of the 81 sample PSU's and selecting a sample of segments (clusters of housing units). The within-PSU sampling procedures were designed to achieve the target number of sample persons by age-sex-race-ethnicity. The sample sizes shown in table 3 are the desired numbers of examined persons. A much larger sample has to be identified and contacted because some sample persons refuse the examination portion of the survey. Table 4 shows the expected number of sample persons to be identified to produce the examined sample sizes in table 3. Table 4 was prepared by inflating the sample sizes in table 3 by the reciprocals of the expected response rates. Individual race-ethnicity, sex, and age response rates were used. The Mexican-American response rates were based on the Hispanic HANES experience. The other subdomains used NHANES II response rates.

Two sampling frames were used to select the sample of housing units within each of the sample PSU's. For phase one of the survey, the larger area segment frame is based on the 1980 census of the population and is only for a sample of persons who lived in housing units built before 1980. For units built in 1980 and later, the second stage consists of sets of addresses selected from building permits issued in 1980 or later. These are referred to as new construction segments. For phase two of the survey, only 1990 census information is being used. Since phase two will be carried out in late 1991 through 1994, the 1990 census data will be current during the entire interview period, and a new construction frame is not necessary.

Stratification within PSU's

The sample size for rare subdomains can be increased by differential sampling within PSU's. In NHANES III, to reduce the high cost of screening necessary to locate the desired Mexican-Americans for the sample, area segments consisting of census block groups (BG's) and enumeration districts (ED's) will be stratified by the percent of the population that is Mexican-American, with a higher rate of selection used in strata containing 3 percent or greater Mexican-American population. Households will also be sampled at variable rates depending on the concentration of Mexican-Americans within the stratum.

Essentially, the procedure will involve a basic sample and a supplemental sample as follows:

(a) The basic sample will be a national self-weighting sample that is large enough to provide a self-weighting sample for all sex-age subdomains of both black and white and all other persons (that is, non-black and non-Mexican-American persons), and for a few sex-age subdomains for Mexican-American persons.

(b) The supplementation necessary for the increased sample in most sex-age groups for Mexican-Americans will be restricted to those BG-ED's with high Mexican-American populations.

(c) The supplementation will introduce variations in sampling rates that increase design effects. There will be an increase in the number of sample Mexican-Americans for those sex-age cells that were at the 560 level in the original plan to keep the precision at the specified level. This increase will be compensated for by a reduction in sample size in those Mexican-American estimation cells that exceed the 560 level. Therefore, the total sample size will remain at the level specified.

(d) The amount of supplementation in the high-density strata will be large enough to supply the number of sample persons needed for the rarest cell, Mexican-American males 60 years of age and over. Other sex-age domains will be subsampled to minimize the variability in sampling rates. This will be done by reducing the sampling rates in the highest density minority stratum until it is the same as the rate in the next highest stratum. For subdomains requiring further reduction, the rates in the two highest strata will be reduced, and so on.

A detailed discussion of the analysis leading to the approach summarized in items (a)–(d) above is described in another report (11). The principal findings of that analysis are summarized below.

Amount of screening without stratification

Without differential sampling, the screening costs for the NHANES III would be extremely high. Table 5 shows the screening levels that would be necessary to locate the required numbers of examined sample persons if there were no stratification of BG-ED's. For example, to obtain the desired number of female Mexican-American infants 2–11 months old, slightly over 172,000 households would need to be screened under a self-weighting design. For female black infants 2–11 months old, the corresponding screening level is about 67,500. Further, research showed that at the level of 67,500 households, the screening sample would be large enough to provide the desired sample sizes for all age-sex classifications for both black and white and other persons.

Optimizing sample size among strata

In table 6 there are projections of the 1990 distribution of Mexican-American persons according to the degree of concentration in BG-ED's, adjusted to reflect the changes in the distribution expected to occur in the 1980–90 decade. These distributions were derived from tabulations of the 1980 Bureau of the Census Master Area Reference File (MARF). The MARF tape does not identify Mexican-American persons separately from other Hispanic persons. Hence, the Hispanic distribution in the United States excluding New York, New Jersey, Connecticut, and Florida was used to approximate the Mexican-American distribution. The Hispanic community in the four excluded States is predominantly Puerto Rican and Cuban. The data were used only to develop the general sample design strategy. The detailed 1989 Bureau of the Census STF-1 tapes were used for sample selection; they contained separate counts of Mexican-American persons, as well as other Hispanic subgroups.

The data in table 6 indicate that most Mexican-American persons live in areas with a high concentration of Mexican-Americans. For example, almost 50 percent of all Mexican-Americans live in areas (BG-ED's) that are more than 25 percent Mexican-American. The fact that Mexican-Americans tend to be concentrated in certain areas indicates that it is possible to achieve worthwhile reductions in screening by oversampling the more highly concentrated areas. The optimum allocation of the sample among the geographical strata shown in table 6 will be restricted to the high-density Mexican-American BG-ED's, with the sample in the less dense strata restricted to the part of the basic 67,500 self-weighting sample located in the less dense strata. To facilitate the allocation process, two subsets of strata will be defined. Subset "a" will consist of the least concentrated areas (strata 1 and 2), and subset "b" will consist of the more highly concentrated areas (strata 3–10).

The following terms are defined:

n_a = expected number of sample persons for a given sex-age-race group in subset a from the basic (67,500) screening sample

n_b = corresponding number to be selected from subset b to meet specified precision requirements.

For a particular age-sex-race group, the precision requirement was expressed in terms of the variance of an estimated proportion, p'. In the subsequent analysis, V is the desired variance of an estimated proportion based on a self-weighting sample of size n, where n is the desired number of sample persons necessary to meet the sample size targets. These sample sizes are shown in table 7 for two selected subgroups. The subgroups given in this table represent two of the rarest groups. They were selected to give an indication of the amount of supplementation necessary with stratification. In the analysis leading up to the optimum allocation, a number of different subgroups were considered in the analysis. However, the derivations of the final sampling rates were based on the optimum allocation for elderly Mexican-American males. The sample sizes include the inflation factors necessary to compensate for nonresponse.

The value of n_a for a particular subgroup is determined from the 67,500-household screening sample. In table 7 the n_a is shown in the column headed "Expected number of sample persons in low density strata from basic sample. The value of n_b can then be determined as follows. First, however, it should be noted that the precision requirement on p' is:

$$\sigma^2(p') = W_a^2\sigma^2(p'_a) + W_b^2\sigma^2(p'_b) = V$$

where W_a = proportion of the population of interest in subset a

W_b = proportion of the population of interest in subset b

p'_a = estimated proportion in subset a

p'_b = estimated proportion in subset b

The required values of V for a 10-percent item are shown in the next-to-last column of table 7. (A 10-percent item was used in variance calculations. However, the optimum allocation holds for all estimated proportions provided that the proportion does not vary widely from stratum to stratum.) It should be noted that $\sigma^2(p'_a)$ is fixed because it is determined by the number of cases supplied by the basic screening sample. Thus, the overall precision requirement can be expressed as a requirement on the variance of p'_b, that is,

$$\sigma^2(p'_b) = \frac{V - W_a^2\sigma^2(p'_a)}{W_b^2} = V_1$$

Using these values of V_t (shown in the last column of table 7), it follows that to minimize the cost in subset b, the optimum allocation of the sample to stratum h (in subset b) is given by the usual formula (13).

$$n_h = n_b \frac{W_h S_h / \sqrt{C_h}}{\sum_j W_j S_j / \sqrt{C_j}}$$

where W_j is the proportion of the group of interest in stratum j (in subset b), C_j is the corresponding relative unit cost, S_j is the standard deviation of the item being estimated (assumed to be the same in all strata), and n_b is the total sample to be allocated to subset b. To meet the specified precision requirement, n_b was computed from the following formula:

$$n_b = \frac{(\sum_j W_j S_j / \sqrt{C_j})(\sum_j W_j S_j \sqrt{C_j})}{V_t}$$

In table 8 there is a summarization of the n_b for two selected subdomains, and the corresponding expected reduction in screening levels with geographical stratification.

The analysis discussed above was carried through using a number of different assumptions concerning the overall level of screening. The actual within-stratum sampling rates used in NHANES III have been derived under the assumption that the screening will be at a sufficiently high level to produce the required numbers of Mexican-American males aged 80 years or more, the group requiring the most screening.

Stratification of segments

Area segments consist of city or suburban blocks (as defined in the most recent census), combinations of two or more blocks, or other area segments in places where block statistics were not produced in the 1980 census. (In PSU's with examinations scheduled for 1991 or later, 1990 census data are being used.)

Most of the United States was blocked for the 1980 census. In areas where block statistics were available, segments are single blocks when the measure of size (MOS) of the blocks exceeds a certain minimum. Blocks that are below the minimum are combined with other blocks that are in close geographical proximity. The combinations are carried out as a computer operation. Within each PSU, the blocks reported on the 1980 census STF-1B file in each minority and density stratum are sorted by tract, block group (BG), and block number. Blocks with MOS below the minimum are combined with succeeding blocks until the desired MOS is achieved. The combinations are kept to the same BG. When the combinations approach the end of a BG without reaching the minimum, earlier blocks within the same BG are added. Consequently, the combinations consist of blocks in close geographical proximity, and in most cases, they are adjacent blocks. As a result of the method of combination, some large blocks that could have been segments by themselves are combined with small blocks.

In the nonblocked part of the United States (mostly rural areas), 1980 census enumeration districts (ED's) generally comprise the segments. They are always the first stage of selection, with small ED's combined in the same way as small blocks. Where ED's are unusually large and it appears that they create unreasonable workloads for the person performing the listing operation, they are "chunked." In a few cases the chunking can be done as an office operation, from information available on maps. More often, however, a field visit is necessary to subdivide the ED into a number of smaller geographical areas. One such area is selected at random. This random selection is taken into account in recording the probability of selection of the segment. Maps of all segments are prepared and they define the areas that are subsequently listed.

The new construction sample utilizes a three-stage sample design: (1) PSU's; (2) clusters of building permits issued during one or several adjoining months by a building permit office; (3) and housing units within the clusters. The sampling rates at the various stages are arranged to provide a self-weighting sample of new construction. All new construction is classified into the nondensity stratum and uses the sampling rates for that stratum. The measures of size are based on the assumption that all residents are in the "white and all other" category and that the housing units are occupied by average-size households.

The source of the data used to establish measures of size for the building permit offices and, within each office, the measures for each month or year starting with 1980 is the Bureau of the Census C-40 reports, "Construction Reports—Housing Authorized by Building Permits and Public Contracts." The selection of offices and time periods is performed as an office operation by Westat, Inc. A segment is defined as all residential permits issued per month in a building permit office reporting monthly to the Bureau of the Census or in a year for annual reporters. Where the monthly or annual permits are below a predesignated minimum, consecutive time periods are combined. Within each PSU, the places are listed in sequence, and within place there is a listing of the total segments for each month or year (a segment is based on the number of housing units authorized). With a random start, a systematic sample of segments is selected.

To sample specific housing units and obtain the addresses of the selected units, field visits to the sample building permit offices are necessary. Interviewers visit the offices, list all of the permits for the months specified, subsample permits to obtain the equivalent of a single measure following instructions provided, and obtain the addresses of the sample units. When the sample units are located in large apartment houses, the entire building is subsequently listed and subsampled. The same procedures for listing and subsampling are used for the building permit sample and the area sample.

The procedures for selecting the segment sample involve both explicit and implicit modes of stratification. The PSU and the minority-density geographical strata comprise explicit stratification. The six density strata within the PSU's are shown in table D.

To keep combined blocks within a single BG, the stratification is done on the basis of the characteristics of the BG or ED in which the segments are located rather than on the specific block or blocks in a segment. This stratification is only applied to area segments. The new construction segments are included in stratum 1, the nondensity stratum. Within the geographical strata, there is implicit stratification created by sorting the area segments by tract number, BG or ED number within tract, and segment number within BG or ED, and selecting a systematic sample with PPS. The new construction segments are sorted by month and year the permits were issued. The sort order generally introduces a partial effect of stratifying by socio-economic level.

Measure of size of segments

To describe the procedure for creating measures of size for segments, the following notation is used:

The subscripts for the indexes are:

h = PSU
i = Mexican-American density stratum
j = segment
k = race-ethnicity subdomains
l = sex-age subdomains

The parameters used for the computations are:

N_{hijkl} = estimated population in the segment
$N_{h.kl}$ = total population in k, lth subdomain in PSU across all segments
$N_{..kl}$ = total in U.S. population in k, lth subdomain
n_{hijkl} = sample size in segment
r_{ikl} = $n_{.ikl}/N_{ikl}$
= U.S. sampling rate in stratum i for k, lth race-ethnicity, sex-age subdomain
M_h = measure of size of hth PSU
M_{hij} = measure of size of h, i, jth segment
M = total measure of stratum in which PSU is located

The measure of size of a segment is calculated in the same way as for PSU's. Research on intraclass correlations and unit costs indicated that an average of 14 examined sample persons per segment is reasonably close to an optimum for most statistics in NHANES. Also, as noted earlier, operational requirements make it necessary to have a fairly constant number of examined sample persons per stand, about 340 in most cases. This implies having about 24 segments per stand. There were 24 segments selected in each noncertainty PSU. The number in certainty stands varied a little from this number depending on the measure of size of the stand. M_{hij} is denoted as the measure of size of a segment where

$$M_{hij} = \sum_{kl} r_{ikl} N_{hijkl}$$

The conditional probability of selection of a segment (j) in stratum (i) within noncertainty PSU (h) is therefore (24) M_{hij}/M_h.

The sampling rates within a segment are

$$r_{ikl} M/48M_{hij}$$

The sample size in the segment is then

$$\frac{M}{48M_{hij}} \sum_{kl} r_{ikl} N_{hijkl} = M/48$$

As $M/2$ is approximately 340, the average segment sample size is 14. Defining the measures of size as indicated above produces an approximately constant number of sample persons per segment. A similar strategy was followed in the certainty PSU's, designed to produce about the same sample size per segment and the same ultimate probability in each subdomain as in the noncertainty PSU's.

As indicated at the beginning of this section on measure of size of segments, N_{hijkl} is defined as the number of persons in a segment in each race-ethnicity-sex-age subdomain, and $N_{h\ kl}$ is the total number in the subdomain in the PSU. The current population of the PSU or segment is not known in such detail; the sample selection therefore uses 1980 data, except for the census updates of the total county population. The Bureau of Census's estimate of the current population of a county is denoted by C_h and the 1980 count by C'_h. (The 1985 census updates were used for the values of C_h.) Similarly, the current U.S. population in a subdomain is denoted by $C_{\ kl}$ and the 1980 count by $C'_{\ kl}$. The race-ethnicity totals are $C_{.k}$ and C'_{kl}. The estimates of the current population of a PSU by race-ethnicity are estimated by:

$$N_{h\ k} = \frac{C_h}{C'_h} C'_{h\ .k}$$

For purposes of sample selection, it is assumed that the age-sex distribution within a race-ethnicity group in a segment conforms to the current U.S. distribution rather than resembling the 1980 distribution in the segment. Consequently,

$$N_{hijkl} = \frac{C_{\ kl} C_h}{C_{\ k} C'_h} C'_{hijk}$$

The following simplifications can be made in the computations of the measures of size.

$$\begin{aligned} M_{hij} &= \sum_{kl} r_{ikl} N_{hijkl} \\ &= \frac{C_h}{C'_h} \sum_k C'_{hijk.} \sum_l r_{ikl} \frac{C_{.kl}}{C_{\ k.}} \\ &= \frac{C_h}{C'_h} \sum_k C'_{hijk} A_{ik} \end{aligned}$$

$$\text{where} \quad A_{ik} = \sum_l r_{ikl} \frac{C_{\ kl}}{C_{\ l}}$$

The values of A_{ik} used in calculating the measures of size are shown in table 9. The measures of size also used the sampling rates r_{ikl}, which differ in the various high-density Mexican-American areas (table 10).

Departures from self-weighting sample

The development above makes a number of assumptions in demonstrating that the measures of size will provide a self-weighting sample with equal size samples in all PSU's and segments. The assumptions, of course, do not apply exactly. Deviations from exact sample sizes in segments are permitted to retain the self-weighting features of the sample, except for a few unusual outliers. The number of sample persons per PSU, however, is fixed in advance and can not be changed. To retain the preassigned workload, some variation among PSU's in sampling rates was necessary, primarily caused by the following factors:

- There is an assumption of equal size strata, equal to $M^*/66$, where M^* is the total measure of size for all PSU's in the United States. In practice there is some variability in stratum sizes. For a self-weighting sample, the variable stratum sizes should be reflected in variable sample sizes per PSU, and thus, per segment.
- For equality in workloads, it is necessary for the current proportion of the population in each race-ethnic group in a PSU and in a segment to be the same as in 1980.
- The proportion of each race-ethnic group living in high-density Mexican-American areas is estimated by using 1980 data with some attrition based on earlier experience.
- The measures of size treat the age-sex distribution for a race-ethnic group as being identical in all segments and PSU's.
- Assumptions are made about the nonresponse and coverage rate for each subdomain, and it is assumed these rates apply in all PSU's. These assumptions may not hold exactly.

Number of segments and probability of selection

The discussion above indicates that there are 24 segments in most stands, and that the within segment rates provide a uniform sampling rate across all PSU's if the sizes of the strata used in the PSU selection are equal. Although the strata sizes can not be made equal, the range is fairly low. The sample selection is based on 24 segments per stand. Because the measures of size of the certainty stands were not equal to half the noncertainty strata, all of them did not have 24 segments. The sample for the two phases consists of 2,138 segments in all 89 stands in the sample.

The actual probability of selection of a segment depends on the measure of size of the segment, the measure of the PSU, and the total measure of the stratum from which the PSU is selected. The following terms apply in determining the probability of selection:

M_{ahij} = measure of size of a segment in the hth PSU in stratum a

P_{ah} = measure of the PSU that was used in sample selection

P_a = stratum size

The probability of selection of the PSU was $2P_{ah}/P_a$. The probability of a segment within the PSU is:

$$\frac{24M_{ahij}}{\sum_{ij} M_{ahij}}$$

The overall probability of selection of a segment is:

$$48\left(\frac{M_{ahij}}{\sum_{ij} M_{ahij}}\right)\left(\frac{P_{ah}}{P_a}\right)$$

It can be noted that P_{ah} is approximately equal to

$$\sum_{ij} M_{ahij}$$

The segment measures of size implicitly include provision for the required oversampling in minority density strata. For example, if the sampling rates in one density stratum are twice those in another, then the measures are twice as large.

Minimum measure of size of segments

One of the goals of the sample design is to create equal probabilities of selection for each domain, within each density stratum, within a PSU. To create equal probabilities, the within-segment sampling rate for a domain in noncertainty PSU's should be:

$$S_{hijkl} = (r_{ikl})\,\frac{\sum_{ij} M_{ahij}}{48M_{ahij}}\,\frac{P_a}{P_{ah}}$$

In certainty areas it is:

$$S_{hijkl} = (r_{ikl})\,\frac{\sum_{ij} M_{ahij}}{24M_{ahij}}$$

To avoid creating special weights, the within-segment sampling rates need to be ≤1 for all values of r_{ikl}. The most restrictive constraint is for domains with the highest value of r_{ikl} in the density strata. These maximum sampling rates are indicated by $\hat{r}_i$ as shown in table D. The values of r_{ikl} for all values of i, k, l are shown in table 10. These values of r_{ikl} were used when NHANES III began. Periodic reviews of the sample yield will be made during the survey. If the reviews indicate that some subdomains may deviate from the desired sample sizes, adjustments in sampling rates will be made. A summary report of the rates used in each stand will be made when NHANES III is completed.

In noncertainty areas, the within-segment sampling requires that

$$S_{hijkl} = \hat{r}_i\quad\frac{\sum_{ij} M_{ahij}}{48M_{ahij}}\,\frac{P_a}{P_{ah}} \le 1$$

and

$$M_{ahij} \geq \frac{\hat{r}_i \sum M_{ahij}}{48} \frac{P_a}{P_{ah}}$$

In certainty areas, the requirement is

$$M_{ahij} \geq \hat{r}_i \frac{\sum M_{ahij}}{24}$$

In some of the certainty counties, the 24 in the denominator is replaced by the number of segments designated for the PSU, in most cases 48 for two-stand counties and 72 for three-stand counties.

Controlling sample size per PSU

To implement the sample segment selection, the minimum measure is made 50 percent greater to permit a reserve 50-percent sample to be selected. The procedure for controlling the sample sizes in the PSU's and calculating the weighting factors is described below.

The sample size in each PSU that will result from a self-weighting sample in each domain within each density stratum is derived. This number is based on several assumptions that are expected to hold only approximately. However, once calculations are prepared of the sample sizes, they are treated as quotas and the number of sample persons in each PSU must adhere to the quota. The calculation of the sample size goals within a PSU is described below.

The probability of selecting a noncertainty PSU is $2M_{ah}/M_a$ where M_{ah} is the measure of size of the hth PSU in stratum a. For any domain-density stratum, a constant sampling rate in the United States is desired. This rate is denoted by r_{ikl}. Within the sample PSU, the sampling rate is $2r_{ikl}M_a/M_{ah}$. The total number of sample persons in a PSU is

$$\sum_{ikl} r_{ikl} N_{ikl} M_a / 2M_{ah}$$

where N_{ikl} is the total population in the PSU in class i, k, l. On the assumption that the population distribution will be approximately the same as in 1980, $M_{ah} = \sum r_{ikl} N_{ikl}$ and the sample size will be $M_a/2$. In certainty PSU's, the sample size is approximately equal to $\sum r_{ikl} N_{hikl} = M_h$. The quotas (goals) assigned to the PSU's are thus proportionate to the measures of size of the strata from which the PSU is selected. If M is the total measure of size of all PSU's in the United States, the quotas for the PSU's are calculated as shown below. (The total quota shown in the formula below, 40,561, is taken from table 4.)

For noncertainty PSU's, the quota is

$$\frac{M_a}{2M}(40{,}561)$$

For certainty PSU's, the quota is

$$\frac{M_h}{M}(40{,}561)$$

The stratification established for the PSU selection keeps the values of M_a within fairly narrow bounds. Thus, in noncertainty PSU's the quotas do not vary substantially from the average of 456, which is 40,561 divided by the 89 total stands in the sample. There will be greater variation among the certainty PSU's.

As there is a constant number of segments per PSU (24) in noncertainty areas, the variation in quotas per PSU is also reflected in segment sample sizes. In addition, since 1980 the changes of the population distribution among segments is likely to be greater than among PSU's. Thus, more variation can be expected in the average segment size than in PSU's, but even this should be within a moderate range.

The approximate equality that exists in sample sizes per PSU and segment does not occur in the screening sample. Considerable variation can be expected. The amount of screening per segment varies considerably among the density strata. About half of the screening will be in the minority density strata; therefore, the amount of screening in a PSU is partially based on what part of the population lives in high-density strata.

The number of sample persons in a generated PSU depends upon several factors that include the race-ethnicity breakdown in the PSU, the age distributions, and the proportion of Mexican-American persons living in the various density strata. A set of assumptions is made about these factors to permit the sampling operations to proceed. However, there is no way of knowing in advance whether the assigned quota for a particular PSU is lower or higher than what would arise from self-weighting samples within the various domains and density strata. Consequently, it is necessary to have a sample selection procedure that can produce samples either somewhat larger or smaller than those arising from the application of the self-weighting sampling rates.

Initially a screening sample that uses sampling rates 50 percent larger than those for self-weighting samples is designated in each PSU. This screening sample in each PSU is then divided into a group of subsamples, referred to as "stop-rule" groups. Each subsample is a systematic subsample of the screeners, with the screeners sequenced prior to subsampling in the following order: density stratum, segment number, and household number. Each subsample thus cuts across all segments and density strata. The stop-rule groups and the percent of households in each subsample are shown in table E.

The 50-percent subsample is released first to the interviewers. When the initial assignment is about 75 percent complete, the resulting yield is analyzed and used to project estimates of the total number of sample persons expected from the initial assignment. Based on these estimates, additional subsamples

Table E. Percent of Mexican-Americans in segment by density stratum number: Third National Health and Nutrition Examination Survey, 1988–94

Density stratum number	*Percent of Mexican-Americans in segment*
1	Nondensity, less than 3
2	3–4.9
3	5–9 9
4	10–19.9
5	20–49.9
6	50 or more

are released. As additional households are screened, the decision on the number of subsamples required is reevaluated to ascertain if more households will be necessary to achieve the target number of sample persons. If so, additional subsamples are released. The reevaluation is done continuously.

A count is kept of the number and size of subsamples used. This will provide the information necessary to calculate the PSU sample weights that will reflect the deviations in sample sizes from self-weighting samples used within the PSU's.

Selection of households and persons

The third stage of sample selection consists of households and certain types of group quarters. All households in the sample segments are listed, and a subsample of households and group quarters is 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 households in most of the United States, and higher rates for the geographical strata with high minority concentrations. Within each geographical stratum, there is an approximately equal probability sample of households across all 81 PSU's.

A constant sampling rate for the screened households is desired within each density stratum (subject to the stop-rule modification). The screening sample in each density stratum must equal the highest rate among all subdomains for the screening sample to yield the desired number of sample persons in all subdomains. These screening rates denoted by r_i are shown in table D.

Applying these sampling rates to the expected number of occupied housing units in the various density strata provides an estimate of the number of households to be screened. They are shown in the last column of table D.

For domains with the maximum sampling rates in a density stratum, subsampling is not required. For other domains, however, subsampling reduces the screening sample to the rates shown in table 10. The subsampling rates are the ratios of the sampling rates for a domain divided by $\hat{r}_i$. These are shown in table 11.

Within-segment sampling rates

To achieve equal probability of selection within a density stratum, the subsampling rate within a segment must be

$$S_{hijkl} = r_{ikl}/\text{prob}_{hij}$$

where S_{hijkl} is the selection probability for the k, lth subdomain in the jth segment in the ith density stratum in the hth PSU and r_{ikl} is the overall probability of selection, as reported in table 10. Prob_{hij} is the probability of selection of the jth segment. As stated earlier, the probability of a segment in a noncertainty PSU is

$$\text{Prob}_{hij} = 48 \frac{M_{ahij}}{\sum_{ij} M_{ahij}} \frac{P_{ah}}{P_a}$$

where P_{ah} is the measure of the PSU and P_a is the measure of the stratum from which the PSU was selected. In certainty PSU's, the probability was

$$24 \frac{M_{hij}}{\sum_i M_{hij}}$$

Parameters used in computing measures of size

One of the goals in computing measures of size is to create approximately equal workloads among PSU's and segments. Except for differences arising from variation in the size of the strata used for PSU selection, which were fairly small, equality is achieved by the following measures of size:

For PSU's

$$M_h = \sum_i P_{hi} N_h r_i$$

where M_h is the measure of the hth PSU; P_{hi} is the proportion of the population in the PSU that is in the ith combination of subdomains and density strata; and r_i is the sampling rate for that subdomain-density stratum.

For segments

$$M_{hij} = \sum_{kl} r_{ikl} N_{hijkl}$$

As noted earlier, the values of P_{hi}, N_h, and N_{hijkl} must be approximated and use the following formulas:

$$M_{hij} = \frac{C_h}{C'_h} \sum_k C'_{hijk} A_{ik}$$

and

$$M_h = \frac{C_h}{C'_h} \sum_k C'_{hk} A_k$$

where $$A_{ik} = \sum_l r_{ikl} \frac{C_{kl}}{C_k}$$

and

$$A_k = \sum_{il} P_{ik} r_{ikl} \frac{C_{kl}}{C_l}$$

with

C_h = current population of PSU h

C'_h = 1980 population of PSU h

C'_{hk} = 1980 PSU population in kth race-ethnicity subdomain

r_{ikl} = sampling rate in the ith density stratum for k, lth subdomain

P_{ik} = U.S. proportion in the ith density stratum for the kth race-ethnicity subdomain

C_{kl} = projected 1990 U.S. population in the k, lth subdomain

$C_{..k}$ = projected 1990 U.S. population in the kth race-ethnicity subdomain.

These approximations are used because it is likely that race-ethnicity composition in 1980 is an accurate projection of the 1990 data, but the 1980 age-sex breakdowns may be poor projections of the 1990 data. The assumption that the national age-sex breakdowns within a race-ethnicity group apply to each PSU and segment seems more reasonable than that the 1980 age-sex distributions are retained in 1990.

The values of the parameters used to calculate A_{ik} and A_k are shown in tables 9 and 12, respectively.

Selection of sample persons

After the sample of screened households are identified, a sample of persons to be interviewed and examined from individual households is selected. All eligible members (persons 2 months of age and older) within a household are listed and a subsample of individuals is selected based on sex, age, and race-ethnicity. Sample persons are selected at rates established to ensure that the target sample sizes by subdomain will be achieved. This means that young persons, elderly persons, black persons, and Mexican-Americans are oversampled. The sample is also selected to maximize the average number of sample persons per household because it appeared to increase the overall participation rate in previous surveys.

The 52 analytical subdomains were collapsed into 16 groups with a common sampling rate used for each group. Table 10 shows the sampling rates used for the 16 groups of subdomains in the six density strata. These sampling rates are designed to provide a 50-percent reserve sample, as well as a provision for the expected nonresponse in each subdomain.

Sampling rates were calculated for the subdomain in each race-ethnicity group that requires the highest sampling rate to achieve the desired sample size. The calculation is based on the optimum allocation method described in the "Selection of segments" section. These subdomains are in the collapsed classes assigned to the lowest domain numbers for each race-ethnicity group. (The collapsed classes were numbered in descending order of sampling rates, and thus the one with the highest sampling rate appears first in table 10.) These maximum rates determine the screening sample. In each density stratum, a sample of households to be screened is selected at the highest one of these rates that appeared for that density stratum. All screened persons in the subdomain used for optimum allocation are retained in the sample. The screened persons in other subdomains are subsampled to bring the samples down to the desired levels. The subsampling rates were designed to minimize the variability in sampling rates among strata, but still achieve the desired sample sizes, and thus the required precision. This was accomplished by progressively reducing the sampling rates in the highest density domains to equal the ones in the lower density domains to the extent this could be done and the desired sample sizes still be attained.

There was considerable subsampling needed to reduce the screened sample of 106,000 households (which contain about 285,000 persons) to about 41,000 sample persons. 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 NHANES and Hispanic HANES indicates that response rates improve when larger sample sizes within households are used. Therefore, a method of subsampling was developed to increase the average sample size per household. The sampling procedure, described later in this report, appears to maximize the number of sample persons per household. (Conversely, it minimizes the number of households containing sample persons.)

Assuming that a screening sample has been designated and persons are to be subsampled, the persons are classified into L subdomains with subsampling rates $r_1 \ldots r_L$. The subdomains are ordered by subsampling rate so that $r_i \le r_{i+1}$. The screening is the minimum amount necessary to achieve the sample size for the rarest subdomain so that $r_L = 1$.

Dividing the screened households into L unequal size random subsets, the sizes of the subsets are proportionate to: r_1, $r_2 - r_1, r_3 - r_2, \ldots r_{i+1} - r_i, \ldots r_L - r_{L-1}$. The sum of these proportions is $r_L = 1$, so all screened households are assigned to one of the sets.

The subsampling is then carried out as follows:

(a) In the first random subset (r_1 of all screened households), all persons in a household are designated as sample persons.
(b) In the second set (corresponding to $r_2 - r_1$), all persons in a household are sample persons except those in subdomain 1.
(c) In the third set, all persons are sample persons except those in subdomains 1 and 2.
(d) The procedure is continued in this way to subdomain L, which excludes all persons other than those in the Lth subdomain.

This procedure produces the correct subsampling rates for all subdomains. Furthermore, it maximizes the number of sample persons per selected household.

In the field, the sampling procedure is implemented in the following way. At the time of the sample selection for screening, a computer-generated message is attached to the screening questionnaire for each household. This message gives instructions to the interviewer regarding the persons who are to become sample persons (for example, if there are Mexican-Americans in the household, include all Mexican-Americans in the household; if there are black persons present, include all children under 6 years of age and males ages 60 years or over). After all persons in the household are listed on the screener and the age, sex, race, and ethnicity (Mexican-American) status are determined, the interviewer consults the sampling label to determine which person(s), if any, are to receive the extended household adult or youth medical history interview and the physical examination in the mobile examination center. The interviewers do not have to carry out any subsampling operation; they are instructed only on the persons to include.

Because the sampling messages are prepared in advance of the screening, they are based on the expected distribution of the screened population by sex, age, race, or ethnicity, rather than the distribution actually achieved. Thus, there may be small deviations in the sample from the desired number in each subdomain. Such deviations are inevitable when subsampling rates must be established before the screening is completed. The proportion of the screened households assigned to each sampling message is shown in table 13.

As part of the implementation of the subsampling, all screened households are put into random groups, the proportion in each group corresponds to those shown in table 13. A separate randomization is necessary in each density stratum.

Instead of unrestricted randomization, a pseudorandom procedure was used that guarantees that all households within each sequence of 1,000 consecutive households are assigned different random numbers. (The random number assigned determines the sampling message to be received.) Furthermore, each sequence of numbers in a segment is assigned random numbers that are fairly evenly spaced in the interval from 000 to 999.

Within each PSU, the list of households to be screened is sorted. The first sorting is the three combinations of the stop-rule groups shown in table F. They consist of subsample A (50 percent), subsamples B–D (26 percent), and subsamples E–Z (24 percent) of the full sample. Within each group there are further sorts by density stratum, segment number within stratum, and household number within segment. To start, a random number from 000 to 999 is assigned to the first household. A separate random number is used in each PSU. The number 311 is then used as a skip interval and is added successively, starting with the random numbers. The cumulative total, modulo 1,000, provides the random number that determines the sampling message.

Table F. Percent and cumulative percent of households in stop-rule subsamples: Third National Health and Nutrition Examination Survey, 1988-94

Subsample	*Percent*	*Cumulative percent*
A	50	50
B	10	60
C	8	68
D	8	76
E	5	81
F	5	86
G	3	89
H	3	92
I	2	94
J	2	96
K	2	98
L	1	99
Z	1	100

Further subsampling is done to ensure the random selection of a sufficient number of sample persons with morning and afternoon-evening blood draws in the mobile examination center. A 50-percent subsample of persons is designated to have blood drawn in the morning (that is, morning examination appointments). This group is referred to as the "standard" sample. The remaining 50 percent, referred to as the "modified" sample, is given afternoon or evening appointments. The selection of the 50-percent samples is also done by households, rather than by individual sample persons because previous NHANES and Hispanic HANES showed that response rates are improved when sample persons in a household are given appointments at the same time.

The random numbers generated for each household are used to designate the 50-percent sample. Odd numbers correspond to the standard sample and even numbers correspond to the modified group. Information on whether the households are in the standard or modified samples is included in the sampling message.

There are also 50-percent subsamples of persons for two specific tests—allergy and central nervous system (CNS) testing. However, these are selected at the mobile examination center using systematic sampling.

Estimation procedures

Weighting the sample data

The goal of NHANES III is to produce data on the health and nutritional status of the civilian noninstitutionalized population aged 2 months and older residing in the United States. Due to the stratified multistage design of NHANES III, standard statistical methods for analyzing the data collected are inappropriate. An approach to analyzing data from large complex surveys such as NHANES has been published (14).

If each individual in the NHANES III sample had the same probability of selection, then all sample persons could be considered to represent the same number of people in the population. In that case the data could be analyzed without regard to sample weights. However, the complex design of NHANES III includes different probabilities of selection for various subgroups of the population. As a result, the responses of surveyed sample persons (referred to as sample elementary units) must be inflated or weighted in order to produce national estimates. Weights are needed to estimate population means, medians, and other descriptive statistics. Consequently, a weight is assigned to each unit in the sample that then permits estimation of population totals. An estimator $\hat{X}$, for any given population total X can be expressed as a weighted sum over all sample elementary units, defined as:

$$\hat{X} = \sum_{u} x(u) W_f(u) \qquad (1)$$

where u represents a sample elementary unit, $x(u)$ is the health or nutrition characteristic or response of interest for unit u, and $W_f(u)$ is the final weight for sample unit u. The estimation method described in this section applies to statistics derived from the NHANES III examinations. Similar methods can be used to define sample weights for estimates based solely on the NHANES III household interview data.

The purpose of weighting the sample data is to permit analysts to produce estimates of statistics that would have been obtained if the entire sampling frame had been surveyed. Sampling weights are used to ensure that estimation procedures are consistent with the sample design. Weighting the sample data accomplishes the following objectives:

1. To estimate population values from the sample data;
2. To adjust for differential probabilities of selection among various subgroups (race, ethnicity, sex, and age subdomains) and persons living in different density strata);
3. To reduce biases arising because nonrespondents may be different from respondents;
4. To compensate, to the extent possible, for inadequacies in the sampling frame resulting from, for example, omissions of some housing units in the listing of the area segments or omissions of persons with no fixed addresses (noncoverage); and
5. To reduce sampling variances by using auxiliary information, that is known with a high degree of accuracy, in the estimation procedure.

National inflation weights

NHANES III uses a multistage sample design involving clustering, stratification, and varying probabilities of selection to provide an economical design with satisfactory precision for specified subdomains of interest. Because of the complex sample design, complex estimation techniques are also required to produce relatively unbiased estimates of population values from the survey data. The NHANES III estimator takes into account the selection procedures of the complex survey design to define and develop the final sample weight W_f for each sample elementary unit. The final weight is the product of the following three component weights:

1. Inverse of the probabilities of selection;
2. Nonresponse adjustment;
3. Ratio adjustment (poststratification).

The first component weight (w_l) reflects the unequal probabilities of selection. For a multi-stage sample, the overall probability that a person is selected into the sample is the product of the conditional probabilities of selection at each stage of selection. The ultimate probability of selecting a sample person in NHANES III depends on the following three factors: (1) the person's age, sex, race, and ethnicity; (2) the density stratum (based on the percent Mexican-American population in the stratum); and (3) the PSU. The reciprocal or inverse of the product of these three factors, defined as w_1, is the first component weight of W_f in equation (1). The probabilities that reflect the effect of the age-sex-race and/or ethnicity domains and density stratum are shown in table 10, which shows the numerators of the probabilities of selection. The denominator is 930 for each domain and density stratum. The third factor in the first component weight is the PSU which also affects the probability of selection because the "stop-rule" procedure, described in the section, "Controlling sample size per PSU," determines what part of the self-weighting sample is used. This part can vary from PSU to PSU. The PSU probability is the total percentage of the full sample in the set of stop-rule subsamples used in the PSU.

The overall probability of selection for each person is the product of the appropriate rate in table 10 and the PSU rate. The first component weight (w_1) is the reciprocal of this probability. There are two special cases that should be noted in the calculation of w_1 for NHANES III. First, there are examinations that are only carried out for half of the sample—morning (standard) or modified (afternoon-evening) blood draws, allergy tests, and central nervous system tests. The probabilities of these items are half the overall probabilities described above. Secondly, since NHANES III is actually two national samples, the first 44 stands or survey locations (phase 1) comprise a national sample and the second 45 stands (phase 2) also comprise a national sample; thus, for analyses of only phase 1 or phase 2 data, the probabilities are further multiplied by a factor of one-half.

In an ideal hypothetical sampling situation having no nonsampling error components—for example, no frame problems, nonresponse, or interviewer effects—the first component weight will provide an unbiased estimator for the population total X, that is, equation 1, with w_1 substituted for W_f yields the estimator:

$$\hat{X} = \sum_{u} x(u) w_1(u) \qquad (2)$$

Such an estimator is referred to as a Horvitz-Thompson estimator.

However, sample surveys rarely get all selected persons to participate; that is, unit nonresponse occurs. Historically, in NHANES, approximately 25 percent of the sample people refuse to participate in the examination portion of the survey. This examination nonresponse could bias the Horvitz-Thompson estimator of equation (2). An analysis of nonresponse bias in the NHANES II has been described (15). A weighting adjustment for nonresponse is believed to reduce potential bias for most items measured in NHANES. The first component weight, w_1, is therefore adjusted to compensate for differential nonresponse rates by grouping respondents and nonrespondents into homogeneous groups. These groups are generally referred to as weighting classes. In NHANES, these weighting classes are typically formed on the basis of selected age groups, sex, race-ethnicity, income, SMSA residence, and geographic region. The examination nonresponse adjustment weight, referred to as w_2, is a ratio calculated as follows:

$$w_2 = \frac{\sum_{i=1}^{n_s(g)} W_{(1i)}}{\sum_{i=1}^{n_r(g)} W_{(1i)}} \qquad (3)$$

where W_{1i} is the first component weight for the ith sample person, $n_s(g)$ is the total sample size in the gth nonresponse weighting class cell, and $n_r(g)$ is the number of examined persons in the gth cell. The summation is carried out separately for each weighting class cell.

A second type of nonresponse that occurs in sample surveys is item nonresponse. Item nonresponse occurs when a sampled unit provides usable information for some items but not for others. Item nonresponse in NHANES III occurs when sample persons refuse or cannot respond to certain questionnaire items or do not participate in certain examination components. Examination item nonresponse can also be the result of insufficient time to complete the entire 4-hour examination. One method of accounting for item nonresponse is to impute for the missing data. A commonly used imputation procedure of adjusting data sets for missing values is the "hot-deck" method (16).

The final component weight (w_3) is called poststratification which is done to align the estimated population counts (for the 52 age-sex-race-ethnic analytical subdomains) from the sample to the total population or census count at the midpoint of the data collection period. The population figures used as controls in the poststratification for the entire NHANES III will be 1991 estimates; for phase 1, 1990 estimates will be used; and for phase 2, 1993 estimates will be used. The poststratification component weight is calculated as follows:

$$w_3 = \frac{N_k}{\sum_{i=1}^{n_k} w_{1i} w_{2i}} \qquad (4)$$

where N_k is the census estimate of the civilian, noninstitutional population in domain k, and the summation in the denominator is over all sample persons in domain k.

The final examination weight (w_f) for each sample person (i) is the product of the above three component weights:

$$w_{fi} = w_{1i} w_{2i} w_{3i} \qquad (5)$$

The adjusted weights will be smoothed even further to ensure that there are no extremely small or large sampling weights.

For each sample person, a final examination weight reflecting the unequal probabilities of selection, adjustments for nonresponse, and poststratification will be included in the public-use data tapes. Similarly, an interview weight and any special subsample weights will also be included in the data tapes. Most of the commonly used computer software packages have an option for incorporating sample weights in cross-tabulations and statistical analyses.

Variance estimation

The NHANES III is based on a complex sample design. The assumption of simple random sampling for estimating variances is not appropriate because it would result in estimates of variances for most items that are lower than those actually present. Design effects are often used to gauge the effects of the various sampling techniques, such as clustering and stratification, and they provide an indication of the success of the complex sample in controlling the variances of the estimates compared with simple random samples. A design effect is defined as the ratio of the actual variance of an estimate from a complex sample to the expected variance of the same estimate, if the sample were drawn from a simple random sample. When the design effect is close to 1.0, the complex sample design is determined to have little effect on the variances. Analysts could consider assuming simple random sampling for data analysis. The use of average design effects is discussed in the last section on variances for subdomains.

Because of adjustments for nonresponse and poststratification in NHANES III, precise formulas for computing sampling variances from the complex survey are not available. However, there are several methods that can provide good approximations for the sampling variance. For a variance approximation to be satisfactory, the variance estimates must reflect all the major features of the sample design used in the survey, including the weighting of the sample data. The three methods generally used for variance estimation with complex samples are Taylor-linearization, balanced repeated replication (BRR), and jackknife (17,18). Generally, the different approximations give similar estimates for sampling variances. No approximation method is substantially better, in all circumstances, than any other method. Software available for linearization methods may not provide for poststratification. This can lead to serious overstatements of the variances for some statistics. In addition, linearization methods often do not take into account the variance effects of nonresponse adjustments because of the difficulty in expressing the adjustment methods in algebraic form. BRR and jackknife can handle poststratification and nonresponse adjustments more easily, although complete reweighting must be done for replicates.

For NHANES III, two PSU's were selected from each noncertainty stratum, making it possible to compute virtually unbiased estimates of variances. NHANES III analytical data tapes will include variance strata and replicate weights that can be used for either BRR or linear approximation, thereby allowing users choices in variance estimation procedures. Computer software packages that compute appropriate standard errors of estimates from surveys with a complex sample design are available. Previously, NCHS used the NCHS BRR program (19) to calculate variances for data collected in NHANES. Recently, analysts of NHANES data have also used two SAS procedures, SESUDAAN and SURREGR (20,21), which depend on a Taylor-series approximation. Other software packages available include SUPER CARP (22), a program developed at Iowa State University that depends on a Taylor-series approximation; Wesvar (23), which was developed by Westat, Inc., that can be used for BRR or jackknife; and OSIRIS, developed by the University of Michigan (24), which contains procedures for either linearization, BRR, or jackknife.

A new software package, "Software for SUrvey DAta ANalysis" (SUDAAN), has been developed by the Research Triangle Institute in cooperation with NCHS and other agencies of the Public Health Service (25). SUDAAN uses the first-order Taylor series approximation to determine estimates of standard errors for means and proportions (and differences in means and proportions) with appropriate corrections for complex survey designs, including poststratification. One advantage that this program has over other linearization software packages is that it allows analysts to incorporate the actual complex sample design of the survey in the calculation of standard errors, for example, the joint probabilities of selection for each pair of PSU's and whether the sample was selected with or without replacement. In addition, the software is available for personal computers as well as for mainframe computers.

Variance estimates for each phase

In the allocation of noncertainty PSU's to the two phases of NHANES III, one PSU in each stratum (selected at random) was assigned to phase 1 and the other one to phase 2. Thus, the sample design for each phase was a one-PSU-per-stratum selection. There is no completely unbiased method of estimating variances for such a design. A common approximation is to pair strata that are similar and to compute variances as if the two selected strata from each pair had been selected from a single stratum, with replacement. These pairings will also be indicated on the data tapes.

Variances for subdomains

For some subdomain analyses in NHANES III, estimates may be based on small sample sizes or come from a small number of PSU's. The variance estimates for these statistics are likely to be unstable, that is, the estimates of variances may themselves be subject to high variability. In this situation, the approach often used is to compute an average design effect to correct estimates of variances based on the assumption of simple random sampling. This was the recommended procedure for analysis of data collected in Hispanic HANES, which was a survey of three special Hispanic subgroups in selected areas of the United States (26, 27). This strategy may also be advisable for some subdomain estimates in NHANES III.

References

1. 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.
2. 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). 1968.
3. 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.
4. Miller HW. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–73. National Center for Health Statistics. Vital Health Stat 1(10a). 1978.
5. National Center for Health Statistics. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971–73. National Center for Health Statistics. Vital Health Stat 1(10b). 1977.
6. Engel A, Murphy RS, Maurer K, Collins E. Plan and operation of the HANES I Augmentation Survey of Adults 25–74 Years, United States, 1974–75. National Center for Health Statistics. Vital Health Stat 1(14). 1978.
7. 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.
8. 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.
9. National Center for Health Statistics. Plan and operation of the third National Health and Nutrition Examination Survey, 1988–94. Vital Health Stat. To be published.
10. Chu A, Waksberg J. NHANES III Sample Design—Final Report. Contract No. 282-86-0042. Rockville, Md., Westat, Inc., 1988.
11. Chu A, Lannom L, Morganstein D, Waksberg J. NHANES III Methods Research—Final Report. Contract No. 282-86-0042. Rockville, Md., Westat, Inc., 1989.
12. Durbin J: Design of multistage surveys for estimation of sampling errors. Applied Stat 16:152–64. 1967.
13. Hansen MH, Hurwitz WN, Madow WG. Sample Survey Methods and Theory. New York. John Wiley and Sons, Inc. 1953.
14. Landis JR, Lepkowski JM, Eklund SA, Stehouwer SA. A statistical methodology for analyzing data from a complex survey, the first National Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Stat 2(92). 1982.
15. Forthofer RN. Investigation of nonresponse bias in NHANES II. Amer J of Epidemiology 117:507–15. 1983.
16. Ford, BL. An Overview of Hot-Deck Procedures. In: Madow WG, Olkin I, and Rubin D, eds. Incomplete Data in Sample Surveys, Volume 2: Theory and Bibliographies. New York. Academic Press, Inc., 185–207. 1983.
17. Rust K. Variance estimation for complex estimators in sample surveys. J Official Stat 1: 381–97, 1985.
18. Wolter, KM. Introduction to variance estimation. New York: Springer-Verlag. 1985.
19. Jones G. The NCHS BRR Program. Unpublished document.
20. Shah, BV. SESUDAAN: Standard Errors Program for Computing of Standardized Rates from Sample Survey Data. Research Triangle Institute, Research Triangle Park, NC. 1981.
21. Holt, MM. SURREGR: Standard Errors of Regression Coefficients from Sample Survey Data. Research Triangle Institute, Research Triangle Park, NC. 1977. (Revised April 1982 by B.V. Shah.)
22. Hidiroglou MA, Fuller WA, and Hickman RD. SUPER CARP., Sixth Ed. Survey Section, Statistical Laboratory. Iowa State University. Ames, Oct., 1980.
23. Flyer P, Rust K, and Morganstein D. Complex survey variance estimation and contingency table analysis using replication. 1989 Proceedings of the Survey Research Methods Section of the American Statistical Association, 110–19. 1990.
24. Survey Research Center Computer Support Group: OSIRIS IV User's Manual, Institute for Social Research. University of Michigan, Ann Arbor. 1979.
25. SUDAAN: Software for Survey Data Analysis. Research Triangle Institute, Research Triangle Park, NC. 1989.
26. Kovar MG, Johnson C. Design effects from the Mexican-American portion of the Hispanic Health and Nutrition Examination Survey. 1986 Proceedings of the Survey Research Methods Section of the American Statistical Association, 396–9. 1987.
27. Lago J and others. Evaluation of design effects for the Mexican-American portion of the Hispanic Health and Nutrition Examination Survey. 1987 Proceedings of the Survey Research Methods Section of the American Statistical Association, 595–600. 1988.

List of detailed tables

Table 1. Number of stands (survey locations) and PSU measures of size by certainty counties: Third National Health and Nutrition Examination Survey, 1988–94

Certainty counties	*Number of stands (survey locations)*			*PSU[1] (and stratum) measure of size[2]*
	Both phases	*Phase 1*	*Phase 2*	
Total	21	10	11	14,452
Maricopa, AZ (Phoenix)	1	1	-	638
Los Angeles, CA	6	3	3	4,546
Orange, CA	1	1	-	726
San Diego, CA	1	-	1	781
Santa Clara, CA	1	-	1	540
Cook, IL (Chicago)	3	1	2	1,851
Wayne, MI (Detroit)	1	1	-	707
Kings, NY (Brooklyn)	1	-	1	673
Philadelphia, PA	1	1	-	524
Bexar, TX (San Antonio)	1	1	-	976
Dallas, TX	1	-	1	650
El Paso, TX	1	-	1	596
Harris, TX (Houston)	2	1	1	1,244

[1]Primary sampling unit.
[2]Expected sample size in PSU

Information on this page has been removed due to confidentiality requirements

Information on this page has been removed due to confidentiality requirements

Table 3. Number of examined persons, by race-ethnicity, sex, and age: Third National Health and Nutrition Examination Survey, 1988–94

	Number of examined persons			
Sex and age	Total	Black	White and all other	Mexican-American[1]
Both sexes	30,094	8,971	12,123	9,000
Male				
All ages	14,919	4,335	6,050	4,534
2–11 months	1,000	[170	660	[170
12–35 months	1,278	394]	490	394]
3–5 years	1,610	560	490	560
6–11 years	1,610	560	490	560
12–19 years	1,610	560	490	560
20–29 years	1,580	[511	490	[579
30–39 years	1,529	460]	490	579]
40–49 years	1,193	[350	490	[353
50–59 years	920	210]	490	220]
60–69 years	1,137	[300	490	[347
70–79 years	827	186	490	151
80 years and over	625	74]	490	61]
Female				
All ages	15,175	4,636	6,073	4,466
2–11 months	995	[141	683	[171
12–35 months	1,278	395]	490	393]
3–5 years	1,610	560	490	560
6–11 years	1,610	560	490	560
12–19 years	1,610	560	490	560
20–29 years	1,628	[599	490	[539
30–39 years	1,614	594]	490	530]
40–49 years	1,246	[396	490	[360
50–59 years	994	271]	490	233]
60–69 years	1,082	[279	490	[313
70–79 years	845	189	490	166
80 years and over	663	92]	490	81]

NOTE: Consolidated age classes for analyses are indicated by [ ] Sample sizes assume 560 examined persons per consolidated class, and 1,000 examined persons in the 2–11 months age class. For white and all other persons, the minimum sample size is 490 in order to satisfy the total sample size requirement

[1]Mexican-Americans can be any race.

Table 4. Number of sample persons, by race-ethnicity, sex, and age: Third National Health and Nutrition Examination Survey, 1988–94

Sex and age	Number of sample persons			
	Total	Black	White and all other	Mexican-American[1]
Both sexes	40,561	11,882	16,781	11,898
Male				
All ages	20,041	5,612	8,236	6,193
2–11 months	1,231	198	835	198
12–35 months	1,499	443	598	458
3–5 years	1,900	651	598	651
6–11 years	1,925	667	583	675
12–19 years	1,991	675	598	718
20–29 years	2,149	623	662	864
30–39 years	2,193	657	671	865
40–49 years	1,861	603	731	527
50–59 years	1,389	309	731	349
60–69 years	1,707	435	721	551
70–79 years	1,245	251	754	240
80 years and over	951	100	754	97
Female				
All ages	20,520	6,270	8,545	5,705
2–11 months	1,243	191	843	209
12–35 months	1,551	459	613	479
3–5 years	1,955	667	605	683
6–11 years	1,932	683	598	651
12–19 years	2,012	683	620	709
20–29 years	2,127	799	645	683
30–39 years	2,177	836	671	670
40–49 years	1,757	591	710	456
50–59 years	1,495	411	742	342
60–69 years	1,703	465	778	460
70–79 years	1,430	326	860	244
80 years and over	1,138	159	860	119

NOTE: Consolidated age classes for analyses are indicated by [ ]. Sample sizes assume 560 examined persons per consolidated class, and 1,000 examined persons in the 2–11 months age class. For white and all other persons, the minimum sample size is 490 in order to satisfy the total sample size requirement.

[1]Mexican-Americans can be any race.

Table 5. Number of households to be screened without stratification to meet minimum sample size requirement per consolidated class after allowance for nonresponse, by race-ethnicity, sex, and age: Third National Health and Nutrition Examination Survey, 1988–94

	Number of households		
Sex and age	*Black*	*White and all other*	*Mexican-American*[1]
Male			
2–11 months	67,500	63,536	150,212
12–35 months	63,881	18,800	146,834
3–5 years	63,648	12,563	144,710
6–11 years	35,652	6,400	86,145
12–19 years	32,821	5,320	79,402
20–29 years	27,706	4,208	79,258
30–39 years	32,456	3,830	79,258
40–49 years	43,441	5,398	79,258
50–59 years	37,075	7,888	84,290
60–69 years	60,532	8,463	194,468
70–79 years	56,257	14,488	194,695
80 years and over	56,763	39,839	194,118
Female			
2–11 months	67,500	67,500	172,366
12–35 months	67,941	20,251	167,881
3–5 years	66,686	13,362	165,851
6–11 years	37,240	6,882	90,497
12–19 years	33,726	5,788	84,764
20–29 years	30,292	4,119	67,219
30–39 years	31,999	3,798	67,219
40–49 years	33,909	5,032	67,219
50–59 years	34,423	7,502	78,092
60–69 years	43,800	7,799	145,862
70–79 years	45,288	11,285	145,585
80 years and over	45,292	22,728	143,899

NOTE: Consolidated age classes for analysis are indicated by [ ].

[1]Mexican-Americans can be any race

Table 6. Projections of the 1990 U.S. Mexican-American population by stratum number, according to the percent in block group or enumeration district, number in thousands, and percent distribution: Third National Health and Nutrition Examination Survey, 1988–94

Stratum number	*Percent of Mexican-Americans in BG/ED*[1] *stratum*	*Mexican-American population in thousands*	*Percent distribution of Mexican-Americans*
Total	...	11,603	100
1	less than 1	580	5
2	1–2.9	1,044	9
3	3–4.9	696	6
4	5–7.4	696	6
5	7.5–9.9	696	6
6	10–14.9	928	8
7	15–19.9	696	6
8	20–24.9	696	6
9	25–49.9	1,857	16
10	50+	3,714	32

[1]BG/ED is block group or enumeration district.

SOURCE: Tabulation of 1980 Master Area Reference File The counts of Mexican-Americans are derived from the Hispanic counts for the entire United States excluding New York, New Jersey, Connecticut, and Florida.

Table 7. Number of sample persons for self-weighting sample, level of screening, proportion of subgroup in low-density stratum, expected sample sizes, desired variances, and required variances in high-density stratum, by selected subdomains: Third National Health and Nutrition Examination Survey, 1988–94

Subdomain	Number of sample persons that meet precision target with self-weighting sample	Screening necessary to meet precision target	Proportion of subgroup population in low-density strata [1]	Expected number of sample persons in low-density strata from basic sample [2] (adjusted for nonresponse)	Desired variance of estimate	Required variance of estimate in high-density strata [1]
Mexican-American males, ages 80 years and over	97	194,000	0.14	4	0.00093	0.00067
Mexican-American males, ages 2–11 months	198	150,000	0.14	11	0.00045	0.00039

[1] Low-density and high-density strata are referred to as subset "a" and "b," respectively, in the section of the text "Optimizing sample size among strata." For this table, low density for Mexican-American has been defined as strata 1 and 2 in table 6, and high density is comprised of strata 3–10.
[2] Basic sample is assumed to be based on a screening sample of 67,500 households.

Table 8. Sample sizes and screening levels for self-weighting and stratified samples, by selected subdomains: Third National Health and Nutrition Examination Survey, 1988–94

Subdomain	Level of Screening (Self-weighting sample)				Level of Screening (Stratified sample with optimum allocation)			
	Number of identified sample persons to meet precision target (adjusted for nonresponse)	Total	Basic sample [1]	Supplement	Number of identified sample persons to meet precision target (adjusted for nonresponse)	Total	Basic sample [1]	Supplement
Mexican-American males ages 80 years and over	141	194,000	67,500	126,500	141	106,000	67,500	38,500
Mexican-American males ages 2–11 months	228	150,000	67,500	82,500	228	82,000	67,500	14,500

[1] Basic sample is assumed to be based on a screening sample of 67,500 households.

Table 9. Values of A_{ik}[1] used in calculating measures of size, by race-ethnicity and density stratum: Third National Health and Nutrition Examination Survey, 1988–94

Density stratum (=i)	White and all other (k=1)	Black (k=2)	Mexican-American [2] (k=3)
1	0.119	0.583	0.844
2	0.119	0.583	1.177
3	0.119	0.583	1.418
4	0.119	0.583	1.502
5	0.119	0.583	1.539
6	0.119	0.583	1.576

NOTE: $A_{ik} = \sum_{l} r_{ikl} \frac{C_{.kl}}{C_{.k}}$

[1]Numbers are the numerators of A_{ik}; denominators are 930. See section of text entitled "Measure of size of segments" for a description of how A_{ik} is used in the calculation of the segment measure of size.
[2]Mexican-Americans can be any race.

Table 10. Sampling rates and expected sample sizes by Mexican-American density stratum, race-ethnicity, and sex-age domains: Third National Health and Nutrition Examination Survey, 1988–94

Race-ethnicity and domain[1]	*Mexican-American density stratum (percent Mexican-American in stratum)* 1 (< 3)	2 (3–4)	3 (5–9)	4 (10–19)	5 (20–49)	6 (50 or more)	Proportion of total	1990 population NHANES III estimate (thousands)	Sample size with reserve	Expected total sample size
White and all other										
Total	...	...	...	...	...	...	1.000	196,183	25,202	16,781
1	1.000	1.000	1.000	1.000	1.000	1.000	0.012	2,417	2,531	1,678
2	0.590	0.590	0.590	0.590	0.590	0.590	0.009	1,783	1,120	754
3	0.300	0.300	0.300	0.300	0.300	0.300	0.049	9,408	3,101	2,071
4	0.190	0.190	0.190	0.190	0.190	0.190	0.106	20,829	4,248	2,817
5	0.114	0.114	0.114	0.114	0.114	0.114	0.182	35,480	4,377	2,972
6	0.086	0.086	0.086	0.086	0.086	0.086	0.322	63,517	5,842	3,840
7	0.059	0.059	0.059	0.059	0.059	0.059	0.320	62,749	3,983	2,649
Black										
Total	...	...	...	...	...	...	0.999	28,180	17,672	11,882
8	1.000	1.000	1.000	1.000	1.000	1.000	0 075	2,118	2,273	1,515
9	0.900	0.900	0.900	0.900	0.900	0.900	0.102	2,881	2,781	1,880
10	0.670	0.670	0.670	0.670	0.670	0.670	0.071	2,008	1,442	950
11	0.550	0.550	0.550	0.550	0.550	0.550	0.198	5,584	3,300	2,262
12	0.470	0.470	0.470	0.470	0.470	0.470	0.553	15,589	7,876	5,275
Mexican-American[2]										
Total	...	..	...	...	...	...	1.000	11,653	17,691	11,898
13	1.000	1.800	3.000	4.000	5.000	6.000	0.037	430	1,942	888
14	1.000	1.800	3.000	3.700	3.700	3.700	0.066	771	2,584	1,371
15	1.000	1.800	2.800	2.800	2.800	2.800	0.118	1,375	3,679	2,130
16	0.800	1.000	1.000	1.000	1.000	1.000	0.779	9,077	9,487	7,509
Proportion of Mexican-Americans in stratum	0.140	0 060	0.120	0.140	0.220	0.320	...	236,016	60,565	40,561

NOTE: The numbers shown are the numerators of the sampling rates; the denominators are 930, which allows for the 50-percent reserve. These are the rates used when NHANES III started. Review of the sample yield carried out periodically indicated that the sample sizes in some subdomains would probably deviate from the desired numbers Changes were then made in the sampling rates A summary of the rates used in each survey site will be made when NHANES III is completed.

[1]Domain definitions:
1 is male or female under 1 year.
2 is male 80 years and over.
3 is male or female 1–3 years, female 80 years and over.
4 is male or female 3–5 years, 70–79 years
5 is male or female 50–69 years
6 is male or female 6–19 years, 40–49 years.
7 is male or female 20–39 years.
8 is male under 1 year, female under 6 years.
9 is male 1–5, 60 years and over.
10 is female 60 years and over
11 is male or female 6–11 years, male 40–59 years.
12 is male 12–39 years, female 12–59 years.
13 is male 60 years and over.
14 is female under 6 years
15 is male under 6 years, female 60 years and over.
16 is male or female 6–59 years

[2]Mexican-Americans can be any race

Table 11. Domain subsampling rates within screened sample, by Mexican-American density stratum, race-ethnicity, and sex-age domains: Third National Health and Nutrition Examination Survey, 1988–94

	Mexican-American density stratum (percent Mexican-American in stratum)					
Race-ethnicity and domain[1]	*1 (not dense) (<3)*	*2 (3–4.9)*	*3 (5–9.9)*	*4 (10–19.9)*	*5 (20–49.9)*	*6 (50 or more)*
White and all other						
1	1.000	0.560	0.330	0.250	0.200	0.170
2	0.590	0.330	0.200	0.150	0.120	0.100
3	0.300	0.170	0.100	0.080	0.060	0.050
4	0.190	0.110	0.060	0.050	0.040	0.030
5	0.114	0.063	0.038	0.029	0.023	0.019
6	0.086	0.048	0.029	0.022	0.017	0.014
7	0.059	0.033	0.020	0.015	0.012	0.010
Black						
8	1.000	0.560	0.330	0.250	0.200	0.170
9	0.900	0.500	0.300	0.230	0.180	0.150
10	0.670	0.370	0.220	0.170	0.130	0.110
11	0.550	0.310	0.180	0.140	0.110	0.090
12	0.470	0.260	0.160	0.120	0.090	0.080
Mexican-American[2]						
13	1.000	1.000	1.000	1.000	1.000	1.000
14	1.000	1.000	1.000	0.930	0.740	0.620
15	1.000	1.000	0.930	0.700	0.560	0.470
16	0.800	0.560	0.330	0.250	0.200	0.170
Maximum rate relative to stratum 1	1.000	1.800	3.000	4.000	5.000	6.000

[1]Domain definitions:
1 is male or female under 1 year.
2 is male 80 years and over.
3 is male or female 1–3 years, female 80 years and over.
4 is male or female 3–5 years, 70–79 years.
5 is male or female 50–69 years.
6 is male or female 6–19 years, 40–49 years.
7 is male or female 20–39 years.
8 is male under 1 year, female under 6 years.
9 is male 1–5, 60 years and over.
10 is female 60 years and over.
11 is male or female 6–11 years, male 40–59 years.
12 is male 12–39 years, female 12–59 years.
13 is male 60 years and over.
14 is female under 6 years.
15 is male under 6 years, female 60 years and over.
16 is male or female 6–59 years.

[2]Mexican-Americans can be any race.

Table 12. Values of parameters used to estimate A_k for primary sampling unit measure of size, by race-ethnicity domain: Third National Health and Nutrition Examination Survey, 1988–94

	$C_{.kl}/C_{..k}$	*Mexican-American density stratum* (i)	P_{ik}	$\sum_i P_{ik} r_{ikl} \frac{C_{.kl}}{C_{..k}}$
White and all other				
1	0.012	...	...	...
2	0.009	...	...	...
3	0.049	...	...	...
4	0.106	...	...	...
5	0.182	...	...	...
6	0.322	...	...	...
7	0.320	...	...	...
Black				
8	0.075	...	...	...
9	0.102	...	...	...
10	0.071	...	...	...
11	0.198	...	...	...
12	0.553	...	...	...
Mexican-American[1]				
13	0.037	1	0.140	0.118
14	0.066	2	0.060	0.071
15	0.118	3	0.120	0.170
16	0.276	4	0.140	0.210
17	0.340	5	0.220	0.338
18	0.163	6	0.320	0.504

White and all other: A_1 = 0.119/930
Black: A_2 = 0.583/930
Mexican-American: A_3 = 1.412/930

Where $A_k = \sum_{il} P_{ik} r_{ikl} \frac{C_{.kl}}{C_{..k}}$

[1]Mexican-Americans can be any race.

NOTE: Mexican-Americans are listed as having six age-sex domains although other tables show only four domains. At the time the PSU's were selected it was expected that six domains would be used, and the parameters for the PSU selections were selected on that basis. The number of domains were collapsed to four prior to selection of households and persons.

Table 13. Proportion of households with each sampling message label describing which household member to include in the sample, by density stratum and race-ethnicity domains: Third National Health and Nutrition Examination Survey, 1988–94

	Mexican-American density stratum (percent Mexican-American in stratum)						
Race-ethnicity and domain[1]	*1 (not dense) (<3)*	*2 (3–4.9)*	*3 (5–9.9)*	*4 (10–19.9)*	*5 (20–49.9)*	*6 (50 or more)*	*Cumulative domain definition*
White and all other							
Total	1.000	1.000	1.000	1.000	1.000	1.000	
1–7	0.059	0.033	0.020	0.015	0.012	0.010	All
1–6	0.027	0.015	0.009	0.007	0.005	0.004	Male or female under 20, 40 years and over
1–5	0.028	0.015	0.009	0.007	0.006	0.005	Male or female under 6, 50 years and over
1–4	0.076	0.043	0.025	0.019	0.015	0.013	Male or female under 6, 70 years and over
1–3	0.110	0.061	0.037	0.027	0.022	0.018	Male or female under 4, 80 years and over
1–2	0.290	0.161	0.097	0.073	0.058	0.048	Male or female under 2, male 80 years and over
1	0.410	0.228	0.136	0.102	0.082	0.069	Male or female under 2 years
None	-	0.444	0.667	0.751	0.800	0.833	No sample persons
Black							
Total	1.000	1.000	1.000	1.000	1.000	1.000	
8–12	0.470	0.261	0.157	0.118	0.094	0.078	All
8–11	0.080	0.045	0.027	0.020	0.016	0.014	Male or female under 12 years, male 40 years and over, female 60 years and over
8–10	0.120	0.066	0.040	0.030	0.024	0.020	Male or female under 6 years, male or female 60 years and over
8–9	0.230	0.128	0.077	0.058	0.046	0.038	Male or female under 6 years, male 60 years and over
8	0.100	0.056	0.033	0.025	0.020	0.017	Male under 1 year, female under 6 years
None	-	0.444	0.667	0.750	0.800	0.833	No sample persons
Mexican-American[1]							
Total	1.000	1.000	1.000	1.000	1.000	1.000	
13–18	0.800	0.556	0.333	0.250	0.200	0.167	All
13–15	0.200	0.444	0.600	0.450	0.360	0.300	Male or female under 6, male 60 years and over
13–14	-	-	0.067	0.225	0.180	0.150	Female under 5 years, male 60 years and over
13	-	-	-	0.075	0.260	0.383	Male 60 years and over
None	-	-	-	-	-	-	

[1]Mexican-Americans can be any race.

Appendix

Table I. Target diseases and conditions: Third National Health and Nutrition Examination Survey, 1988–94

Allergy
Arthritis
Cancer
Cardiovascular disease
Chronic obstructive pulmonary disease
Dental health
Diabetes
Gallbladder disease
Hearing
Infectious diseases
Kidney disease
Mental health
Osteoporosis

Table II. Home examination components for selected age groups: Third National Health and Nutrition Examination Survey, 1988–94

2–11 months	*20 years and over*
Recumbent length	Height
Weight	Weight
Mid-arm circumference	Mid-arm circumference
Triceps skinfold	Triceps skinfold
Head circumference	...
Food frequency (proxy)	...
...	Cognitive function[1]
...	Physical function[1]
...	Venipuncture
...	Interview questions
...	Spirometry

[1]60 years and over.

NOTE: Home examinations are not offered to sample persons 1–19 years.

Table III. Examination components as conducted in the mobile examination center for each age group: Third National Health and Nutrition Examination Survey, 1988–94

2 months–5 years	*6–19 years*	*20–39 years*	*40–59 years*	*60–74 years*	*75 years and over*
Physician exam	Physician exam	Physician exam	Physician exam	Physician exam	Physician exam
Venipuncture[1]	Venipuncture	Venipuncture	Venipuncture	Venipuncture	Venipuncture
...	...	...	Glucose tolerance test	Glucose tolerance test	...
...	Urine specimen	Urine specimen	Urine specimen	Urine specimen	Urine specimen
Body measurements	Body measurements	Body measurements	Body measurements	Body measurements	Body measurements
24-hour dietary recall	24-hour dietary recall	24-hour dietary recall	24-hour dietary recall	24-hour dietary recall	24-hour dietary recall
...	Food frequency[2]	...	...	...	...
...	...	...	Eye fundus photograph	Eye fundus photograph	Eye fundus photograph
...	...	...	ECG	ECG	ECG
...	Bioelectrical impedance[3]	Bioelectrical impedance	Bioelectrical impedance	Bioelectrical impedance	Bioelectrical impedance
...	Spirometry[4]	Spirometry	Spirometry	Spirometry	Spirometry
Dental exam[1]	Dental exam	Dental exam	Dental exam	Dental exam	Dental exam
...	...	Bone density	Bone density	Bone density	Bone density
...	...	...	..	Hand-knee x ray	Hand-knee x ray
...	...	Gallbladder ultrasound	Gallbladder ultrasound	Gallbladder ultrasound	...
...	Allergy	Allergy[5]	Allergy[5]	...	...
...	Audiometry and tympanometry	...	...	...	...
...	...	...	...	Physical function	Physical function
Proxy interview	Interview	Interview	Interview	Interview	Interview
...	...	Central nervous system test[5]	Central nervous system test[5]	...	...
...	Cognitive test[6]	...	...	...	...

[1]1 year or more.
[2]12–16 years.
[3]12–19 years.
[4]8–19 years.
[5]Procedure is for half-sample only.
[6]6–16 years.

Information on this page has been removed due to confidentiality requirements

DEPARTMENT OF
HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

BULK RATE
POSTAGE & FEES PAID
PHS/NCHS
PERMIT NO. G-281

DHHS Publication No. (PHS) 92-1387, Series 2, No. 113