Third National Health and Nutrition Examination Survey
(NHANES III), 1988-94

NHANES III SECOND EXAM FILE DOCUMENTATION

Series 11, No. 3A

July 1999

Table of Contents

Introduction

The National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC) collects, analyzes, and disseminates data on the health status of U.S. residents. The results of surveys, analyses, and studies are made known through a number of data release mechanisms including publications, mainframe computer data files, CD-ROMs (Search and Retrieval Software, Statistical Export and Tabulation System (SETS)), and the Internet.

The National Health and Nutrition Examination Survey (NHANES) is a periodic survey conducted by NCHS. The third National Health and Nutrition Examination Survey (NHANES III), conducted from 1988 through 1994, was the seventh in a series of these surveys based on a complex, multi-stage sample design. It was designed to provide national estimates of the health and nutritional status of the United States' civilian, noninstitutionalized population aged two months and older.

The following table summarizes the NHANES III data which are currently available on CD-ROM, including this release.

Table 1. Available NHANES III CD-ROMs

CD-ROM Name	Release Date	Size in Megabytes	Data Files / Description
NHANES III, 1988-94, Series 11, No. 3A, ASCII Version (this release)	July 1999	33	Second exam sample files for dietary recall, examination, laboratory, additional laboratory analytes and documentation
NHANES III, 1988-94, Series 11, No. 2A, ASCII Version	April 1998	407	Dietary recall (replacement), electrocardiography, laboratory (additional analytes), and vitamins/medicines data files and documentation
NHANES III, 1988-94, Series 11, No. 1, Revised SETS Version 1.22a	October 1997	285	Adult and youth household questionnaire, examination, and laboratory data files and documentation, plan and operation, analytic and reporting guidelines, weighting and estimation methodology, field operations, non-response bias
NHANES III, 1988-94, Series 11, No. 1A, ASCII Version	July 1997	454	Adult and youth household questionnaire, dietary recall, examination, and laboratory data files and documentation
NHANES III, 1988-94, Series 11, No. 1, SETS Version 1.22a *	July 1997	285	Adult and youth household questionnaire, examination, and laboratory data files and documentation
NHANES III Reference Manuals and Reports October 1996	October 1996	152	Plan and operation, analytic and reporting guidelines, weighting and estimation methodology, field operations, non-response bias

* Do not use this CD-ROM It had technical problems and has been superseded by the revised SETS version 1.22a, Series 11, No. 1, released in October 1997.

This CD-ROM, Series 11 No. 3A, contains data obtained from a second exam of selected survey participants who had a primary exam. This release does not replace the previous NHANES III data releases series 11 Nos. 1A and 2A).

Table 2. Location of the interview and examination components in the NHANES III public use data files

Topic	Data File						
	HA	HY	EXAM	LAB	DIET	VMS	ECG
Sample weights	X	X	X	X	.	.	X
Age/race/sex	X	X	X	X	.	.	X
Ethnic background	X	X	.	.	.	.	.
Household composition	X	X	.	.	.	.	.
Individual characteristics	X	X	.	.	.	.	.
Health insurance	X	X	.	.	.	.	.
Family background	X	X	.	.	.	.	.
Occupation of family head	X	X	.	.	.	.	.
Housing characteristics	X	X	.	.	.	.	.
Family characteristics	X	X	.	.	.	.	.
Orientation	X	X	.	.	.	.	.
Health services	X	X	.	.	.	.	.
Selected health conditions	X	X	X	.	.	.	.
Diabetes questions	X	.	.	.	.	.	.
High blood pressure and cholesterol questions	X	.	.	.	.	.	.
Cardiovascular disease questions	X	.	.	.	.	.	.
Musculoskeletal conditions	X	.	.	.	.	.	.
Physical functioning questions	X	.	.	.	.	.	.
Gallbladder disease questions	X	.	.	.	.	.	.

Table 2. (continued) Location of the interview and examination components in the NHANES III public use data files

	Data File						
Topic	HA	HY	EXAM	LAB	DIET	VMS	ECG
Kidney conditions	X	.	.	.	.	.	.
Respiratory and allergy questions	X	X	.	.	.	.	.
Diet questions	X	.	.	.	.	.	.
Food frequency	X	.	X	.	.	.	.
Vision questions	X	X	.	.	.	.	.
Hearing questions	X	X	.	.	.	.	.
Dental care and status	X	X	.	.	.	.	.
Tobacco	X	.	X	.	.	.	.
Occupation	X	.	.	.	.	.	.
Language usage	X	X	.	.	.	.	.
Exercise	X	.	.	.	.	.	.
Social support/residence	X	.	.	.	.	.	.
Vitamin/mineral/medicine usage	X	X	X	.	.	.	.
Blood pressure measurement	X	.	X	.	.	.	.
Birth	.	X	X	.	.	.	.
Infant feeding practices/diet	.	X	.	.	.	.	.
Motor and social development	.	X	.	.	.	.	.
Functional impairment	X	X	.	.	.	.	.
School attendance	.	X	.	.	.	.	.
Cognitive function	.	X	X	.	.	.	.

Table 2. (continued) Location of the interview and examination components in the NHANES III public use data files

Topic	Data File						
	HA	HY	EXAM	LAB	DIET	VMS	ECG
Alcohol and drug use	.	.	X	.	.	.	.
Reproductive health	.	.	X	.	.	.	.
Diagnostic interview schedule	.	.	X	.	.	.	.
Activity	.	.	X	.	.	.	.
Physician's examination	.	.	X	.	.	.	.
Height and weight	.	.	X	.	.	.	.
Body measurements	.	.	X	.	.	.	.
Dental examination	.	.	X	.	.	.	.
Allergy skin test	.	.	X	.	.	.	.
Audiometry	.	.	X	.	.	.	.
Tympanometry	.	.	X	.	.	.	.
WISC and WRAT	.	.	X	.	.	.	.
Spirometry	.	.	X	.	.	.	.
Bone densitometry	.	.	X	.	.	.	.
Gallbladder ultrasonography	.	.	X	.	.	.	.
Central nervous system function evaluation	.	.	X	.	.	.	.
Fundus photography	.	.	X	.	.	.	.
Physical function evaluation	.	.	X	.	.	.	.
Fasting questions	.	.	.	X	.	.	.

Table 2. (continued) Location of the interview and examination components in the NHANES III public use data files

Topic	Data File						
	HA	HY	EXAM	LAB	DIET	VMS	ECG
Laboratory tests on blood and urine	.	.	.	X	.	.	.
Total nutrient intakes	.	.	X	.	.	.	.
Individual foods	.	.	.	.	X	.	.
Combination foods	.	.	.	.	X	.	.
Ingredients	.	.	.	.	X	.	.
Prescription Medicines	X	X	.	.	.	X	.
Vitamins and Minerals	X	X	.	.	.	X	.
Electrocardiography	.	.	.	.	.	.	X

Data File Definitions

HA - Household Adult Data File
HY - Household Youth Data File
EXAM - Examination Data File
LAB - Laboratory Data File and Second Laboratory Data File
DIET - Dietary Recall Data Files
VMS - Vitamin Mineral Supplement Data File
ECG - Electrocardiography Data File

This document includes the documentation for the NHANES III Second Exam Combined Foods File and also contains a general overview of the survey and the use of the data files. The general overview includes five sections. The first section, entitled "Guidelines for Data Users," contains important information about the use of the data files. The second section, "Survey Description," is a brief overview of the survey plan and operation. The third section, "Sample Design and Analysis Guidelines," describes some technical aspects of the sampling plan and discusses some analytic issues particularly related to the use of data from complex sample surveys. The "Data Preparation and Processing Procedures" section describes the editing conventions and the codes used to represent the data. The last and fifth section, "General References," includes a reference list for the survey overview sections of the document.

Public Use Data Files for the third National Health and Nutrition Examination Survey will also be available from the National Technical Information Service (NTIS). A list of NCHS public use data tapes available for purchase from NTIS may be obtained from the Data Dissemination Branch at NCHS. Information regarding a bibliography (on disk) of journal articles

citing data from all the NHANES and the availability of NHANES III data in CD-ROM/SETS software format can be obtained from the Data Dissemination Branch at:

Data Dissemination Branch
National Center for Health Statistics
Room 1018
6525 Belcrest Road
Hyattsville, Maryland 20782

Phone: (301)458-4636

URL:http://www.cdc.gov/nchswww

NTIS can be contacted at:

NTIS - Computer Products Office
5285 Port Royal Road
Springfield, Virginia 22161
(703) 487-4807

Copies of all NHANES III questionnaires and data collection forms are included in the Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94 (NCHS, 1994; U.S. DHHS, 1996). This publication, along with detailed information on NHANES procedures, interviewing, data collection, quality control techniques, survey design, nonresponse, and sample weighting can be found on the NHANES III Reference Manuals and Reports CD-ROM (U.S. DHHS, 1996). Information on how to order this CD-ROM is also available from the Data Dissemination Branch at NCHS at the address and telephone number given above.

NHANES III Second Exam Sample

The NHANES III Second Exam Sample was a sub-study of NHANES III, conducted for research purposes. These research files are intended to provide additional data for use with special statistical methods to improve estimates from the main survey data and for methodologic investigations. Following this description of the Second Exam Sample is information on the overall survey which is also relevant for the Second Exam Sample, including: general guidelines for data users, a description of the survey, sample design, analysis guidelines and a description of the data preparation and processing procedures.

Sample design and survey description

No statistical sampling design was applied for the second exam. However, a nonrandom sample of about five percent was obtained by selecting approximately 20
participants from the roughly 400 sample persons examined at each survey location. The following general guidelines were used by the MEC staff to select participants for the second exam:

1) select mainly adults, 2)half between the ages of 20-39 years, and half over 40 years; 3) select about half men and half women. The sample obtained consists of 2,603 persons, with 1,205 males (46 percent)and 1,398 females (54 percent).

Age group	2nd # of Exams	Percentage of 2nd Exams
< 12	212	8
12-19	231	9
20-39	809	31
40-59	578	22
> 60	773	30

The second exams were scheduled after the first or primary exams, when possible at the same time of day as the first exam. The second exams were conducted over the same time period as the primary exams for a particular survey location by the
same MEC staff, although priority was given to scheduling and completing primary exams. The second exams were administered following the same protocols as for the primary exam, with the following exceptions: the food frequency questionnaire
was not administered to adolescents 12-16 years; the WISC/WRAT was not administered to youths 6-16 years, and hand/knee x-rays were not re-administered on adults aged 60 and over.

Analytic Issues

Due to the research nature of these data, special caution should be used in analysis. All analyses should include thorough investigation of the potential selection bias of this small non-random sub-sample. Careful attention to identifying and evaluating differences in important characteristics (e.g., age and race-ethnicity) between the subsample and the main sample should be considered along with other issues.

The second exam data can be linked to the primary exam data and the household interview data using the unique identifier (SEQN). This is necessary to obtain the demographic data for the sample. NCHS recommends that the survey design variables (e.g., sample weights) not be linked with the second exam data, since the survey design variables were created for the full sample. There are no sample weights or other design variables specifically created for the second exam sample. There are weights labeled as "replicate...weight," but these are Fay's BRR Replicate Interview Weights. These weights are to be applied to the primary exam sample, with software which uses the balanced repeated replication (BRR) method. They should not be used with the Second Exam Sample.

Because the second exams were identical to the primary exams, with the exceptions noted above, the file structure for the second exams is the same as for the primary exam files. The variable nomenclature is the same with the following important distinction: the first or primary exam variable names have a 'p' in the third position while the second or "replicate" exam variable names have a 'r' in the third position (e.g., 'BMPWT' or 'BMRWT').

GUIDELINES FOR DATA USERS

Please refer to the following important information before analyzing data.

NHANES III Background Documents

- The Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94, (NCHS, 1994; U.S. DHHS, 1996) provides an overview of the survey and includes copies of the survey forms.

- The sample design, nonresponse, and analytic guidelines documents on the NHANES III Reference Manuals and Reports CD-ROM (U.S. DHHS, 1996) discuss the reasons that sample weights and the complex survey design should be taken into account when conducting any analysis.

- Instruction manuals, laboratory procedures, and other NHANES III reference manuals on the NHANES III Reference Manuals and Reports CD-ROM(U.S. DHHS, 1996) are also available for further information on the details of the survey.

Analytic Data Set Preparation

- Most NHANES III survey design and demographic variables are found only on the Adult and Youth Household Data Files available on the first release. In preparing a data set for analysis, other data files must be merged with either or both of these files to obtain many important analytic variables.

- All of the NHANES III public use data files are linked with the common survey participant identification number (SEQN). Merging information from multiple NHANES III data files using this variable ensures that the appropriate information for each survey participant is linked correctly.

- NHANES III public use data files do not have the same number of records on each file. The Household Questionnaire Files (divided into two files, Adult and Youth) contain more records than the Examination Data File because not everyone who was interviewed completed the examination. The Laboratory Data File contains data only for persons aged one year and older. The Individual Foods Data File based on the dietary recall has multiple records for each person rather than the one record per sample person contained in the other data files.

- For each data file, SAS program code with standard variable names and labels is provided as separate text files on the CD-ROM that contains the data files. This SAS program code can be used to create a SAS data set from the data file.

- Modifications were made to items in the questionnaires, laboratory, and examination components over the course of the survey; as a result, data may not be available for certain variables for the full six years. In addition, variables may differ by phase since some changes were implemented between phases. Users are encouraged to read the Notes sections of this document carefully for information about changes.

- Extremely high and low values have been verified whenever possible, and numerous consistency checks have been performed. Nonetheless, users should examine the range and frequency of values before analyzing data.

- Some data were not ready for release at the time of this publication due to continued processing of the data or analysis of laboratory specimens. A listing of those data are available in the general information section of each data file.

- Confidential and administrative data are not being released to the public. Additionally, some variables have been recoded to help protect the confidentiality of the survey participants. For example, all age-related variables were recoded to 90+ years for persons who were 90 years of age and older.

- Some variable names may differ from those used in the Phase 1 NHANES III Provisional Data Release and some variables included in the Phase 1 provisional release may not appear on these files.

- Although the data files have been edited carefully, errors may be detected. Please notify NCHS staff (301-458-4636) of any errors in the data file or the documentation.

Analytic Considerations

- NHANES III (1988-94) was designed so that the survey's first three years, 1988-91, its last three years, 1991-94, and the entire six years were national probability samples. Analysts are encouraged to use all six years of survey results.

- Sample weights are available for analyzing NHANES III data. One of the following three sample weights will be appropriate for nearly all analyses: interviewed sample final weight (WTPFQX6), examined sample final weight (WTPFEX6), and mobile examination center (MEC)- and home-examined sample final weight (WTPFHX6). Choosing which of these sample weights to use in any analysis depends on the variables being used. A good rule of thumb is to use "the least common denominator" approach. In this approach, the user checks the variables of interest. The variable that was collected on the smallest number of persons is the "least common denominator," and the sample weight that applies to that variable is the appropriate one to use for that analysis. For more detailed information, see the Analytic and Reporting Guidelines for NHANES III (U.S. DHHS, 1996).

Referencing or Citing NHANES III Data

- In publications, please acknowledge NCHS as the original data source. For instance, the reference for the NHANES III Laboratory Data File on this CD-ROM is:

  U.S. Department of Health and Human Services (DHHS). National Center for Health Statistics. Third National Health and Nutrition Examination Survey, 1988-1994, NHANES III Second Laboratory Data File

(CD-ROM, Series 11, No. 3A). Hyattsville, MD.: Centers for Disease Control and Prevention, 1999.

- Please place the acronym "NHANES III" in the titles or abstracts of journal articles and other publications in order to facilitate the retrieval of such materials in bibliographic searches.

SURVEY DESCRIPTION

The third National Health and Nutrition Examination Survey (NHANES III) was the seventh in a series of large health examination surveys conducted in the United States beginning in 1960. Three of these surveys, the National Health Examination Surveys (NHES), were conducted in the 1960's (NCHS, 1965; NCHS, 1967; NCHS, 1969). In 1970, an expanded nutrition component was added to provide data with which to assess nutritional status and dietary practices, and the name was changed to the National Health and Nutrition Examination Survey (Miller, 1973; Engel, 1978; McDowell, 1981). A special survey of Hispanic populations in the United States was conducted during 1982-1984 (NCHS, 1985).

The general structure of the NHANES III sample design was similar to that of the previous NHANES. All of the surveys used complex, multi-stage, stratified, clustered samples of civilian, noninstitutionalized populations. NHANES III was the first NHANES without an upper age limit; in fact, the age range for the survey was two months and older. A home examination option was employed for the first time in order to obtain examination data for very young children and for elderly persons who were unable to visit the mobile examination center (MEC). The home examination included only a subset of the components used in the full MEC examination since it would have been difficult to collect some types of data in a home setting. A detailed description of design specifications and copies of the data collection forms can be found in the Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-1994 (NCHS, 1994; U.S. DHHS, 1996).

NHANES III was conducted from October 1988 through October 1994 in two phases, each of which comprised a national probability sample. The first phase was conducted from October 18, 1988, through October 24, 1991, at 44 locations. The second phase was conducted from September 20, 1991, through October 15, 1994, at 45 different locations. In NHANES III, 39,695 persons were selected over the six years; of those, 33,994 (86%) were interviewed in their homes. All interviewed persons were invited to the MEC for a medical examination. Seventy-eight percent (30,818) of the selected persons were examined in the MEC, and an additional 493 persons were given a special, limited examination in their homes.

Data collection began with a household interview. Several questionnaires were administered in the household: Household Screener Questionnaire, Family Questionnaire, Household Adult Questionnaire, and Household Youth Questionnaire.

At the MEC, an examination was performed, and five automated questionnaires or interviews were administered: MEC Adult Questionnaire, MEC Youth Questionnaire, MEC Proxy Questionnaire, 24-Hour Dietary Recall, and Dietary Food Frequency (ages 12-16 years). The health examination component included a variety of tests and procedures. The examinee's age at the time of the interview and other factors determined which procedures were administered. Blood and urine specimens were obtained, and a number of tests and measurements were performed including body measurements, spirometry, fundus photography, x-rays, electrocardiography, allergy and glucose tolerance tests, and ultrasonography. Measurements were taken of bone density, hearing, and physical, cognitive, and central nervous system

functions. A physician performed a limited standardized medical examination and a dentist performed a standardized dental examination. While some of the blood and urine analyses were performed in the MEC laboratory, most analyses were conducted elsewhere by contract laboratories.

A home examination was conducted for those sample persons aged 2-11 months and aged 20 years or older who were unable to visit the mobile examination center. The home examination consisted of an abbreviated version of the tests and interviews performed in the MEC. Depending on age of the sample person, the components included body measurements, blood pressure, spirometry, venipuncture, physical function evaluation, and a questionnaire to inquire about infant feeding, selected health conditions, cognitive function, tobacco use, and reproductive history.

SAMPLE DESIGN AND ANALYSIS GUIDELINES

Sample Design

The general structure of the NHANES III sample design is the same as that of the previous NHANES. Each of these surveys used a stratified, multi-stage probability design. The major design parameters of the two previous NHANES and the special Hispanic HANES, as well as NHANES III, have been previously summarized (Miller, 1973; McDowell, 1981; NCHS, 1985; NCHS, 1994). The NHANES III sample was designed to be self-weighting within a primary sampling unit (PSU) for subdomains (age, sex, and race-ethnic groups). While the sample was fairly close to self-weighting nationally for each of these subdomain groups, it was not representative of the total population, which includes institutionalized, non-civilian persons that were outside the scope of the survey.

The NHANES III sample represented the total civilian, noninstitutionalized population, two months of age or over, in the 50 states and the District of Columbia of the United States. The first stage of the design consisted of selecting a sample of 81 PSU's that were mostly individual counties. In a few cases, adjacent counties were combined to keep PSU's above a minimum population 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 operational reasons, these 13 certainty PSU's were divided into 21 survey locations. After the 13 certainty strata were designated, the remaining PSU's in the United States were grouped into 34 strata, and two PSU's were selected per stratum (68 survey locations). The selection was done with PPS and without replacement. The NHANES III sample therefore consists of 81 PSU's or 89 locations.

The 89 locations were randomly divided into two groups, one for each phase. The first group consisted of 44 and the other of 45 locations. One set of PSU's was allocated to the first three-year survey period (1988-91) and the other set to the second three-year period (1991-94). Therefore, unbiased estimates (from the point of view of sample selection) of health and nutrition characteristics can be independently produced for both Phase 1 and Phase 2 as well as for both phases combined.

For most of the sample, the second stage of the design consisted of area segments composed 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 were 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 consisted of sets 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 were used to define the area segments. Because the second phase followed within a few years of the 1990 Census, new construction did not account for a significant part of the sample, and the entire sample came from the area segments.

The third stage of sample selection consisted of households and certain

types of group quarters, such as dormitories. All households and eligible group quarters in the sample segments were listed, and a subsample was designated for screening to identify potential sample persons. The subsampling rates enabled 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 populations. Within each geographic stratum, there was a nearly equal-probability sample of households across all 89 stands.

Persons within the sample of households or group quarters were the fourth stage of sample selection. All eligible members within a household were listed, and a subsample of individuals was selected based on sex, age, and race or ethnicity. The definitions of the sex, age, race or ethnic classes, subsampling rates, and designation of potential sample persons within screened households were developed to provide approximately self-weighting samples for each subdomain within geographic strata and at the same time to maximize the average number of sample persons per sample household. Previous NHANES indicated that this increased the overall participation rate. Although the exact sample sizes were not known until data collection was completed, estimates were made. Below is a summary of the sample sizes for the full six-year NHANES III at each stage of selection:

Number of PSU's	81
Number of stands (survey locations)	89
Number of segments	2,144
Number of households screened	93,653
Number of households with sample persons	19,528
Number of designated sample persons	39,695
Number of interviewed sample persons	33,994
Number of MEC-examined sample persons	30,818
Number of home-examined sample persons	493

More detailed information on the sample design and weighting and estimation procedures for NHANES III can be found in the Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94 (NCHS, 1994; U.S. DHHS, 1996) and in the Analytic and Reporting Guidelines: Third National Health and Nutrition Examination Survey (NHANES III), 1988-94 (U.S. DHHS, 1996).

Analysis Guidelines

Because of the complex survey design used in NHANES III, traditional methods of statistical analysis based on the assumption of a simple random sample are not applicable. Detailed descriptions of this issue and possible analytic methods for analyzing NHANES data have been described earlier (NCHS, 1985; Yetley, 1987; Landis, 1982; Delgado, 1990). Recent analytic and reporting guidelines that should be used for most NHANES III analyses and publications are contained in Analytic and Reporting Guidelines (U.S. DHHS, 1996). These recommendations differ slightly from those used by analysts for previous NHANES surveys. These suggested guidelines provide a framework to users for producing estimates that conform to the analytic design of the survey. All users are strongly urged to review these analytic and reporting guidelines before beginning any analyses of NHANES III data.

It is important to remember that this set of statistical guidelines is not absolute. When conducting analyses, the analyst needs to use his/her

subject matter knowledge (including methodological issues) as well as information about the survey design. The more one deviates from the original analytic categories defined in the sample design, the more important it is to evaluate the results carefully and to interpret the findings cautiously.

In NHANES III, 89 survey locations were randomly divided into two sets or phases, the first consisting of 44 and the other of 45 locations. One set of PSU's was allocated to the first three-year survey period (1988-91) and the other set to the second three-year period (1991-94). Therefore, unbiased national estimates of health and nutrition characteristics can be independently produced for each phase as well as for both phases combined. Computation of national estimates from both phases combined (i.e., total NHANES III) is the preferred option; individual phase estimates may be highly variable. In addition, individual phase estimates are not statistically independent. It is also difficult to evaluate whether differences in individual phase estimates are real or due to methodological differences. That is, differences may be due to changes in sampling methods or data collection methodology over time. At this time, there is no valid statistical test for examining differences between Phase 1 and Phase 2. Therefore, although point estimates can be produced separately for each phase, no test is available to test whether those estimates are significantly different from each other.

NHANES III is based on a complex, multi-stage probability sample design. Several aspects of the NHANES design must be taken into account in data analysis, including the sample weights and the complex survey design. Appropriate sample weights are needed to estimate prevalence, means, medians, and other statistics. Sample weights are used to produce correct population estimates because each sample person does not have the same probability of selection. The sample weights incorporate the differential probabilities of selection and include adjustments for noncoverage and nonresponse. A detailed discussion of nonresponse adjustments and issues related to survey coverage have been published (U.S. DHHS, 1996). With the large oversampling of young children, older persons, black persons, and Mexican-Americans in NHANES III, it is essential that the sample weights be used in all analyses. Otherwise, a misinterpretation of results is highly likely. Other aspects of the design that must be taken into account in data analyses are the strata and PSU pairings from the sample design. These pairings should be used to estimate variances and test for statistical significance. For weighted analyses, analysts can use special computer software packages that use an appropriate method for estimating variances for complex samples such as SUDAAN (Shah, 1995) and WesVarPC (Westat, 1996).

Although initial exploratory analyses may be performed on unweighted data using standard statistical packages and assuming simple random sampling, final analyses should be done on weighted data using appropriate sample weights. A summary of the weighting methodology and the type of sample weights developed for NHANES III is included in Weighting and Estimation Methodology (U.S. DHHS, 1996).

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 (the United States) had been surveyed. Sample weights can be considered as measures of the number of persons the particular sample

observation represents. Weighting takes into account several features of the survey: the specific probabilities of selection for the individual

domains that were oversampled as well as nonresponse and differences between the sample and the total U.S. population. Differences between the sample and

the population may arise due to sampling variability, differential undercoverage in the survey among demographic groups, and possibly other types of response errors, such as differential response rates or misclassification errors. Sample weighting in NHANES III was used to:

1. Compensate for differential probabilities of selection among subgroups (i.e., age-sex-race-ethnicity subdomains where persons living in different geographic strata were sampled at different rates);
2. Reduce biases arising from the fact that nonrespondents may be different from those who participate;
3. Bring sample data up to the dimensions of the target population totals;
4. Compensate, to the extent possible, for inadequacies in the sampling frame (resulting from omissions of some housing units in the listing of area segments, omissions of persons with no fixed address, etc.); and
5. To reduce variances in the estimation procedure by using auxiliary information that is known with a high degree of accuracy.

In NHANES III, the sample weighting was carried out in three stages. The first stage involved the computation of weights to compensate for unequal probabilities of selection (objective 1, above). The second stage adjusted for nonresponse (objective 2). The third stage used poststratification of the sample weights to Census Bureau estimates of the U.S. population to accomplish the third, fourth, and fifth objectives simultaneously. In NHANES III, several types of sample weights (see the sample weights table that follows) were computed for the interviewed and examined sample and are included in the NHANES III data file. Also, sample weights were computed separately for Phase 1 (1988-91), Phase 2 (1991-94), and total NHANES III (1988-94) to facilitate analysis of items collected only in Phase 1, only in Phase 2, and over six years of the survey. Three sets of pseudo strata and PSU pairings are provided to use with SUDAAN in variance estimation. Since NHANES III is based on a complex, multi-stage sample design, appropriate sample weights should be used in analyses to produce national estimates of prevalence and associated variances while accounting for unequal probability of selection of sample persons. For example, the final interview weight, WTPFQX6, should be used for analysis of the items or questions from the family or household questionnaires, and the final MEC examination weight, WTPFEX6, should be used for analysis of the questionnaires and measurements administered in the MEC. Furthermore, for a combined analysis of measurements from the MEC examinations and associated medical history questions from the household interview, the final MEC examination weight, WTPFEX6, should be used. We recommend using SUDAAN (Shah, 1995) to estimate statistics of interest and the associated variance. However, one can also use other published methods for variance estimation. Application of SUDAAN and alternative methods, such as the average design effect approach, balance repeated replication (BRR) methods, or jackknife methods for variance estimation, are discussed in Weighting and Estimation Methodology (U.S. DHHS, 1996).

Appropriate Uses of the NHANES III Sample Weights

Final interview weight, WTPFQX6

Use only in conjunction with the sample interviewed at home and with items collected during the household interview.

Final examination (MEC only) weight, WTPFEX6

Use only in conjunction with the MEC-examined sample and with interview and examination items collected at the MEC.

Final MEC+home examination weight, WTPFHX6

Use only in conjunction with the MEC+home-examined sample and with items collected at both the MEC and home.

Final allergy weight, WTPFALG6

Use only in conjunction with the allergy subsample and with items collected as part of the allergy component of the exam.

Final CNS weight, WTPFCNS6

Use only in conjunction with the CNS subsample and with items collected as part of the CNS component of the exam.

Final morning examination (MEC only) subsample weight, WTPFSD6

Use only in conjunction with the MEC-examined persons assigned to the morning subsample and only with items collected in the MEC exam.

Final afternoon/evening examination (MEC only) subsample weight, WTPFMD6

Use only in conjunction with the MEC-examined persons assigned to the afternoon/evening subsample and only with items collected in the MEC exam.

Final morning examination (MEC+home) subsample weight, WTPFHSD6

Use only in conjunction with the MEC- and home-examined persons assigned to the morning subsample and with items collected during the MEC and home examinations.

Final afternoon/evening examination (MEC+home) weight, WTPFHMD6

Use only in conjunction with the MEC- and home-examined persons assigned to the afternoon/evening subsample and with items collected during the MEC and home examinations.

DATA PREPARATION AND PROCESSING PROCEDURES

Automated data collection procedures for the survey were introduced in NHANES III. In the mobile examination centers, data for the interview and examination components were recorded directly onto a computerized data collection form. With the exception of a few independently automated systems, the system was centrally integrated. This operation allowed for ongoing monitoring of much of the data. Before the introduction of the computer-assisted personal interview (CAPI), the household questionnaire data were reviewed manually by field editors and interviewers. CAPI (1992-1994 only) questionnaires featured built-in edits to prevent entering inconsistencies and out-of-range responses. The multi-level data collection and quality control systems are discussed in detail in the Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-1994 (NCHS, 1994; U.S. DHHS, 1996). All interview, laboratory, and examination data were sent to NCHS for final processing.

Guidelines were developed that provided standards for naming variables, filling missing values and coding conventional responses, handling missing records, and standardizing two-part quantity/unit questionnaire variables. NCHS staff, assisted by contract staff, developed data editing specifications that checked data sets for valid codes, ranges, and skip pattern consistencies and examined the consistency of values between interrelated variables. Comments, collected in both interviews and examination components, were reviewed and recoded when possible. Responses to "Other" and "Specify" were recoded either to existing code categories or to new categories. The documentation for each data set includes notes for those variables that have been recoded and standardized and for those variables that differ significantly from what appears in the original data collection instrument. While the data have undergone many quality control and editing procedures, there still may be values that appear extreme or illogical. Values that varied considerably from what was expected were examined by analysts who checked for comments or other responses that might help to clarify unusual values. Generally, values were retained unless they could not possibly be true, in which case they were changed to "Blank but applicable." Therefore, the user must review each data set for extreme or inconsistent values and determine the status of each value for analysis.

Several editing conventions were used in the creation of final analytic data sets:

1. Standardized variables were created to replace all two-part quantity/unit questions using standard conversion factors. Standardized variables have the same name as the variable of the two-part question with an "S" suffix. For instance, MAPF18S (Months received WIC benefits) in the MEC Adult Questionnaire was created from the two-part response option to question F18, "How long did you receive benefits from the WIC program?," using the conversion factor 12 months per year.

2. Recoded variables were created by combining responses from two or more like variables, or by collapsing responses to create a summary variable for the purpose of confidentiality. Recoded variables have the original variable name with an R suffix. For example, place of birth

variable (HFA6X) in the Family Questionnaire was collapsed to a three level response category (U.S., Mexico, Other) and renamed HFA6XR. Generally, only the recoded variable has been included in the data file.

3. Fill values, a series of one or more digits, were used to represent certain specific conditions or responses. Below is a list of the fill values that were employed. Some of the fill values pertain only to questionnaire data, although 8-fill and blank-fill values are found in all data sets. Other fill values, not included in this list, are used to represent component-specific conditions.

   6-fills = Varies/varied. (Questionnaires only)

   7-fills = Fewer than the smallest number that could be reported within the question structure (e.g., fewer than one cigarette per day). (Questionnaires only)

   8-fills = Blank but applicable/cannot be determined. This means that a respondent was eligible to receive the question, test, or component but did not because of refusal, lack of time, lack of staff, loss of data, broken vial, language barrier, unreliability, or other similar reasons.

   9-fills = Don't know. This fill was used only when a respondent did not know the response to a question and said, "I don't know." (Questionnaires only)

   Blank fills = Inapplicable. If a respondent was not eligible for a questionnaire, test, or component because of age, gender, or specific reason, the variable was blank-filled. In the questionnaire, if a respondent was not asked a question because of a skip-pattern, variables corresponding to the question were blank-filled. For examination or laboratory components, if a person was excluded by a defined protocol (e.g., screening exclusion questions) and these criteria are included in the data set, then the corresponding variables were blank-filled for that person. For home examinees, variables for examination components and blood tests not performed as part of the home examination protocol were blank-filled.

4. For variables describing discrete data, codes of zero (0) were used to mean "none," "never," or the equivalent. Value labels for which "0" is used include: "has not had," "never regularly," "still taking," or "never stopped using." Unless otherwise labeled, for variables containing continuous data, "zero" means "zero.

5. Where there are logical skip patterns in the flow of the questionnaire or examination component, the skip was indicated by placing the variable label of the skip destination in parentheses as part of the value label of the response generating the skip. For example, in the Physical Function Evaluation, the variable PFPWC (in wheelchair) has a value label, "2 No (PFPSCOOT)" that means that the next item for persons not in a wheelchair would be represented by the variable, PFPSCOOT.

Variable Nomenclature

A unique name was assigned to every NHANES III variable using a standard convention. By following this naming convention, the origin of each variable is clear, and there is no chance of overlaying similar variables across multiple components. Variables range in length from three to eight characters. The first two variable characters represent the topic (e.g., analyte, questionnaire instrument, examination component) and are listed below alphabetically by topic. For questionnaires administered in the household, the remainder of the variable name following the first two characters indicates the question section and number. For example, data for the response to the Household Adult Questionnaire question B1 are contained in the variable HAB1. For most laboratory and examination variables, as well as some other variables, a "P" in the third position refers to "primary" and the remainder of the variable name is a brief description of the item. For instance, in the Laboratory Data File, information on the length of time the person fasted before the first blood draw is contained in the variable PHPFAST. The variable PHPFAST was derived as follows: characters 1-2 (PH) refer to "phlebotomy," character 3 (P) refers to "primary," characters 4-8 (FAST) refer to an abbreviation for "fasting."

CODE	TOPIC
AT	Alanine aminotransferase (from biochemistry profile)
AM	Albumin (from biochemistry profile)
AP	Alkaline phosphatase (from biochemistry profile)
AL	Allergy skin test
AC	Alpha carotene
AN	Anisocytosis
TM	Antimicrosomal antibodies
TA	Antithyroglobulin antibodies
AA	Apolipoprotein (AI)
AB	Apolipoprotein (B)
AS	Aspartate aminotransferase (from biochemistry profile)
LA	Atypical lymphocyte
AU	Audiometry
BA	Band
BO	Basophil
BS	Basophilic stippling
BC	Beta carotene
BX	Beta cryptoxanthin
BL	Blast
BU	Blood urea nitrogen (BUN) (from biochemistry profile)
BM	Body measurements
BD	Bone densitometry
C1	C-peptide (first venipuncture)
C2	C-peptide (second venipuncture)
CR	C-reactive protein
UD	Cadmium
CN	Central nervous system function evaluation
CL	Chloride (from biochemistry profile)
CO	Cotinine
CE	Creatinine (serum)(from biochemistry profile)
UR	Creatinine (urine)

CODE	TOPIC
DM	Demographic
DE	Dental examination
MQ	Diagnostic interview schedule
DR	Dietary recall (total nutrient intakes)
EO	Eosinophil
EP	Erythrocyte protoporphyrin
FR	Ferritin
FB	Fibrinogen
RB	Folate (RBC)
FO	Folate (serum)
FH	Follicle stimulating hormone (FSH)
FP	Fundus photography
GG	Gamma glutamyl transferase (GGT) (from biochemistry profile)
GU	Gallbladder ultrasonography
GB	Globulin (from biochemistry profile)
G1	Glucose (first venipuncture)
G2	Glucose (second venipuncture)
SG	Glucose (from biochemistry profile)
GH	Glycated hemoglobin
GR	Granulocyte
C3	HCO3 (Bicarbonate)(from biochemistry profile)
HD	HDL cholesterol
HP	Helicobacter pylori antibody
HT	Hematocrit
HG	Hemoglobin
AH	Hepatitis A antibody (HAV)
HB	Hepatitis B core antibody (anti-HBc)
SS	Hepatitis B surface antibody (anti-HBs)
SA	Hepatitis B surface antigen (HBsAg)
HC	Hepatitis C antibody (HCV)
DH	Hepatitis D antibody (HDV)
H1	Herpes 1 antibody
H2	Herpes 2 antibody
HX	Home examination (general)
HO	Homocysteine
HF	Household family questionnaire
HA	Household adult questionnaire
HQ	Household questionnaire variables (composite)
HS	Household screener questionnaire
HY	Household youth questionnaire
HZ	Hypochromia
I1	Insulin (first venipuncture)
I2	Insulin (second venipuncture)
UI	Iodine (urine)
FE	Iron
SF	Iron (from biochemistry profile)
LD	Lactate dehydrogenase (from biochemistry profile)
L1	Latex antibody
LC	LDL cholesterol (calculated)
PB	Lead
LP	Lipoprotein (a)
LH	Luteinizing hormone

CODE	TOPIC
LU	Lutein/zeaxanthin
LY	Lycopene
LM	Lymphocyte
MR	Macrocyte
MC	Mean cell hemoglobin (MCH)
MH	Mean cell hemoglobin concentration (MCHC)
MV	Mean cell volume (MCV)
PV	Mean platelet volume
MA	MEC adult questionnaire
MX	MEC examination (general)
FF	Dietary food frequency (ages 12-16 years)
MP	MEC proxy questionnaire
MY	MEC youth questionnaire
ME	Metamyelocyte
MI	Microcyte
MO	Monocyte
MN	Mononuclear cell
ML	Myelocyte
IC	Normalized calcium (derived from ionized calcium)
OS	Osmolality (from biochemistry profile)
PH	Phlebotomy data collected in MEC (e.g., questions)
PS	Phosphorus (from biochemistry profile)
PF	Physical function evaluation
PE	Physician's examination
PL	Platelet
DW	Platelet distribution width
PK	Poikilocytosis
PO	Polychromatophilia
SK	Potassium (from biochemistry profile)
PR	Promyelocyte
RC	Red blood cell count (RBC)
RW	Red cell distribution width (RDW)
RE	Retinyl esters
RF	Rheumatoid factor antibody
RU	Rubella antibody
WT	Sample weights
SE	Selenium
SI	Sickle cell
NA	Sodium (from biochemistry profile)
SH	Spherocyte
SP	Spirometry
SD	Survey design
TT	Target cell
TE	Tetanus
TH	Thyroid Stimulating Hormone (TSH)
T4	Thyroxine
TB	Total bilirubin (from biochemistry profile)
CA	Total calcium
SC	Total calcium (from biochemistry profile)
TC	Total cholesterol
CH	Total cholesterol (from biochemistry profile)
TI	Total iron binding capacity (TIBC)
TP	Total protein (from biochemistry profile)
TX	Toxic granulation

CODE	TOPIC
TO	Toxoplasmosis antibody
PX	Transferrin saturation
TG	Triglycerides
TR	Triglycerides (from biochemistry profile)
TY	Tympanometry
UA	Uric acid (from biochemistry profile)
UB	Urinary albumin
VU	Vacuolated cells
VR	Varicella antibody
VA	Vitamin A
VB	Vitamin B12
VC	Vitamin C
VD	Vitamin D
VE	Vitamin E
WC	White blood cell count (WBC)
WW	WISC/WRAT cognitive test

GENERAL REFERENCES

Delgado JL, Johnson CL, Roy I, Trevino FM. Hispanic Health and Nutrition Examination Survey: methodological considerations. Amer J Pub Health 80(suppl.):6-10. 1990.

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.

Freeman DH, Freeman JL, Brock DB, Koch GG. Strategies in the multivariate analysis of data from complex surveys II: an application to the United States National Health Interview Survey. Int Stat Rev 40(3):317-30. 1976.

Khare M, Mohadjer LK, Ezzati-Rice TM, Waksberg J. An evaluation of nonresponse bias in NHANES III (1988-91). 1994 Proceedings of the Survey Research Methods section of the American Statistical Association. 1994.

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.

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.

Miller HW. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971-1973. National Center for Health Statistics. Vital Health Stat 1(10a) and (10b). 1973.

National Center for Health Statistics. Plan and initial program of the Health Examination Survey. Vital Health Stat 1(4). 1965.

National Center for Health Statistics. Plan and operation of a health examination survey of U.S. youths 12-17 years of age. Vital Health Stat 1(8). 1969.

National Center for Health Statistics. Plan and operation of the Hispanic Health and Nutrition Examination Survey, 1982-84. Vital Health Stat 1(19). 1985.

National Center for Health Statistics. Plan and operation of the Third National Health and Nutrition Examination Survey, 1988-94. Vital Health Stat 1(32). 1994.

National Center for Health Statistics. Plan, operation, and response results of a program of children's examinations. Vital Health Stat 1(5). 1967.

Shah BV, Barnwell BG, Bieler GS. SUDAAN User's Manual: Software for Analysis of Correlated Data. Research Triangle Park, NC: Research Triangle Institute. Release 6.04. 1995.

Skinner CJ. Aggregated analysis: standard errors and significance tests. In: Skinner CJ, Holt D, Smith TMF, eds. Analysis of complex surveys. New

York: John Wiley and Sons, Inc. 1989.

U.S. Department of Health and Human Services (DHHS). National Center for Health Statistics. NHANES III reference manuals and reports (CD-ROM). Hyattsville, MD: Centers for Disease Control and Prevention, 1996. Available from National Technical Information Service (NTIS), Springfield, VA. Acrobat .PDF format; includes access software: Adobe Systems, Inc. Acrobat Reader 2.1.

Westat, Inc. A User's Guide to WesVarPC. Rockville, MD. Westat, Inc. 1996.

Yetley E, Johnson C. Nutritional applications of the Health and Nutrition Examination Surveys (HANES). Annu Rev Nutr 7:441-63. 1987.

COMBINATION FOODS FILE

Introduction

Multi-component combination foods are foods that are comprised of more than one component food. Any food in the NHANES III Individual Foods File that is comprised of more than one component food is considered to be a multi-component combination food. Fifteen combination food type codes (DRRTCC) were developed for use in NHANES III based on all the multi-component foods reported in NHANES III; the dietary interviewers assigned codes to all multi-component combination foods reported during NHANES III. Some examples of combination type codes include beverage combinations (DRRTCC=7), sandwich combinations (DRRTCC=1), mixed dish combinations (DRRTCC=4) and salad combinations (DRRTCC=3). A cheese sandwich entered by its components: bread, cheese, mustard, and lettuce would have a DRRTCC=1.

In addition to the multi-component type of combination foods, there were single component combination foods reported during the survey. Single component combination foods were entered into the DDC system using generic or brand name product food entry selections in the DDC system. If a respondent ate a sandwich or mixed dish in a restaurant and could not provide any details about the food, the DDC system's generic combination food selections were used. A cheese sandwich or beef taco could be entered as such using the appropriate selections in the DDC system; generic combination food options were available for many combination foods. A second example would be for a sandwich eaten in a "fast food restaurant"; if the restaurant was one of the commercial establishments in the DDC system, the interviewers could select the restaurant menu item directly as long as the respondent ate the sandwich as prepared.

The NHANES III Individual Foods File lists the food records and component food records for single and multi-component combination foods. Multi-component combination foods have a Combination Food Flag (DRRCFF) value equal to 1. Single component combination foods have DRRCFF=0. Some multi-component foods are comprised of a single component combination food plus additional food components and have a DRRCFF=1 because the food is a multi-component food.

The CFFSE provides a single, summary data record for multi-component combination foods reported in the survey; descriptive information and nutrient values for multi-component combination foods are included in the CFFSE records. The CFFSE was produced by combining component food data in the Individual Foods File into a single combination food record. In the sandwich combination example described earlier, the CFFSE data for the "sandwich" were derived by summing food gram weights and nutrient values for the bread, cheese, mustard, and lettuce components of the sandwich. A single food record for the "cheese sandwich" would appear in the CFFSE for this combination food. Food energy and nutrient values for CFFSE records have distinctive variable names to distinguish them from IFFSE food energy and nutrient variables.

Two criteria were required for inclusion in the CFFSE. First, the final dietary recall status code (DRRSTAT) for the dietary recall had to be equal to 1, 2, or 5. Second, a multi-component food had to have been reported during the dietary interview. Only the foods that were multi-component foods in the Individual Foods File are included in the CFFSE. Some NHANES III examinees do not have any

records in the CFFSE because their recall did not include multi-component foods.

The single component combination foods have single component record foods in the Individual Foods File; the Combination Food Flag (DRRCFF) value for single component combination foods is equal to zero. When respondents were unable to describe the components of a combination food such as a sandwich, the DDC system's generic food entries were selected during the dietary interview. Many brand name combination foods are single component combination foods. When the generic and brand name combination food entries were selected, a standard foods database recipe was used to assign the food code and nutrient values to the combination food.

Single component generic and brand name combination foods were not included in the CFFSE unless component foods were added to or were omitted from the base food. For example, if a respondent added avocado slices to a base food item such as a "fast food restaurant" sandwich, the resulting food was a combination food that consisted of avocado and the restaurant sandwich. If major sandwich components (salad dressing, cheese, etc.) were removed or omitted from a "fast food restaurant" sandwich, the component foods that were eaten were entered individually into the DDC system, thus creating a multi-component combination food. The Individual Foods File and CFFSE data may be combined to identify all reports of specific combination foods. The DRRFCODE in the Individual Foods File, and the DRRCCODE in the CFFSE can be used to accomplish this.

The CFFSE data include food gram weights, nutrient data, and descriptive information about the combination foods. A combination food code (DRRCCODE) was assigned to many combination foods. The DRRCCODE was either a USDA food code or a special food code that was developed by NCHS. NCHS created a small number of special food codes because many combination foods reported during NHANES III could not be coded with existing USDA food codes. For example, there were numerous reports of tortilla combinations containing eggs and meat. NCHS also added a small number of food codes for unusual food combinations. All combination food codes (DRRCCODE) match those in the primary file combination
food code variable, DRPCCODE. The look-up table called "CODEBOOK" can be used to
obtain text descriptions for DRPCCODE.

Seven of the fifteen combination food types (DRRTCC) were assigned combination food codes (DRRCCODE): sandwiches; mixed dishes; mixed dishes with additions; meat, fish, and poultry with additions; soups and stews; other foods eaten together; and tortilla combinations. The DRRCCODE was assigned by NCHS to provide a means of describing the combination foods reported during NHANES III. Many of the combination foods reported during the survey were difficult to code due to the fact that unusual combinations were reported, equal amounts of multiple components were reported as components of a combination food, and numerous variations of foods were reported.

Eight combination food types have generic combination food descriptions that are based on their combination food type label (DRRTCC). Beverage combinations, cereal with additions, vegetables with additions, desserts, vegetable combinations, bread with additions, frozen meals, and salads have generic combination type descriptions that are associated with DRRTCC. Many of the combination foods reported in these categories were comprised of equal amounts of component foods. There were no USDA food codes for many of these combination foods, for example, ready-to-eat breakfast cereals eaten

with sugar and fresh fruit, or unusual vegetable combinations, for example, do not have USDA food codes. Additionally, many combination foods were foods eaten in combination with other foods instead of being components of a single "food" such as a sandwich, taco or stew.

It is suggested that analysts check the Individual Foods File first to review the frequency of occurrence of single component and multi-component combination foods. The food code, brand name code, and food identification code (DRRFCODE, DRRCOMM, and DRRFID) variables are useful to describe single component combination foods as well as specific components of multi-component combination foods. A particular component food may appear in several combination type codes. The CFFSE combination type code (DRRTCC) and combination food code (DRRCCODE) variables are useful for planning analyses involving multi-component combination foods. Detailed documentation is provided to describe DRRCCODE.

Notes to Analysts

There are two exceptions to the basic combination food coding procedures described earlier. The approach NCHS used to handle these data entry exceptions is described.

Exception #1: There are single component foods in the CFFSE that were prepared with more than one type of ingredient. In such cases, the only way to capture specific ingredient information reported by the respondent using the DDC system was to enter the food as a multi-component or combination food. One example was homemade cookies prepared with equal amounts of butter and margarine. The dietary interviewers were instructed to enter half of the total amount of cookies eaten using margarine as an ingredient (component 1) and half of the amount eaten using butter as an ingredient (component 2). To describe this particular "combination," the combination type code (DRRTCC) "Other foods eaten together" was used.

Exception #2- Modified Recipe Foods: The second example of single component foods appearing as combination foods were those with modified recipes -- e.g., homemade cornbread prepared without fat. To report this modified cornbread, NCHS and NCC developed a fat-free cornbread recipe. The ingredients were entered into the DDC system as separate food components resulting in a "combination food" as follows: corn meal (component 1), egg (component 2), milk (component 3), and so forth. Modified recipe foods were coded using DRRTCC=8 or "other foods eaten together".

In summary, combination foods were reported either as single component foods or multi-component foods in the Individual Foods File. The CFFSE provides a single record summary of the multi-component foods reported in the survey. The CFFSE provides total gram weight and nutrient data for many combination foods
reported during the survey. The CFFSE is sorted by sequence number (SEQN), meal number (DRRMN), and food number(DRRFN).

NHANES III Combination Foods Data File from the Dietary Recall Index

Description	Variable Name	Positions
GENERAL INFORMATION		
Respondent identification number	SEQN	1-5
SECOND EXAM DATA		
Recall Status Code	DRRSTAT	6
Meal number	DRRMN	7-8
Food number	DRRFN	9-10
Type of combination code	DRRTCC	11-12
USDA code. Combination food code	DRRCCODE	13-19
SECOND EXAM DATA		
USDA SURVEY NUTRIENT DATABASE		
CFF Total grams of food/beverage	DRRCGW	20-23
CFF Alcohol (gm)	DRRCALCO	24-26
CFF Calcium (mg)	DRRCCALC	27-30
CFF Carbohydrate (gm)	DRRCCARB	31-35
CFF Carotenes (RE)	DRRCCARO	36-39
CFF Cholesterol (mg)	DRRCCHOL	40-43
CFF Copper (mg)	DRRCCOPP	44-46
CFF Total dietary fiber (gm)	DRRCFIBE	47-50
CFF Folacin (micrograms)	DRRCFOLA	51-54
CFF Iron (mg)	DRRCIRON	55-58
CFF Food energy (Kcal)	DRRCKCAL	59-62
CFF Magnesium (mg)	DRRCMAGN	63-65
CFF TOT Monounsaturated fatty acids (gm)	DRRCMFAT	66-70
CFF Niacin (mg)	DRRCNIAC	71-74
CFF TOT Polyunsaturated fatty acids (gm)	DRRCPFAT	75-78
CFF Phosphorus (mg)	DRRCPHOS	79-82
CFF Potassium (mg)	DRRCPOTA	83-86
CFF Protein (gm)	DRRCPROT	87-91
CFF Total saturated fatty acids (gm)	DRRCSFAT	92-95
CFF Sodium (mg)	DRRCSODI	96-100
CFF Total fat (gm)	DRRCTFAT	101-105

NHANES III Combination Foods Data File from the Dietary Recall Index

Description	Variable Name	Positions
CFF Vitamin A (IU)	DRRCVAIU	106-110
CFF Vitamin A (RE)	DRRCVARE	111-115
CFF Thiamin (mg)	DRRCVB1	116-119
CFF Riboflavin (mg)	DRRCVB2	120-123
CFF Vitamin B6 (mg)	DRRCVB6	124-127
CFF Vitamin B12 (micrograms)	DRRCVB12	128-133
CFF Ascorbic acid (mg)	DRRCVC	134-136
CFF Vitamin E (alpha tocopherol equival)	DRRCVE	137-140
CFF Water (gm)	DRRCWATE	141-144
CFF Zinc (mg)	DRRCZINC	145-148

NHANES III Combination Foods Data File from the Dietary Recall

N=5256 DATASET=CFFSE DOCUMENTATION DATE=06/18/99

GENERAL INFORMATION

Positions SAS name	Counts	Item description and code	Notes
1-5 SEQN	5256	Sample person identification number 00009-53594	

NHANES III Combination Foods Data File from the Dietary Recall

SECOND EXAM DATA

GENERAL INFORMATION

Positions SAS name	Counts	Item description and code		Notes
6		Recall status code		
DRRSTAT	5242	1	Reliable and complete	
	14	2	Reliable, but incomplete	
7-8		Meal number		See note
DRRMN	5256	01-13		
9-10		Food number		See note
DRRFN	5256	01-12		
11-12		Type of combination code		See note
DRRTCC	957	01	Sandwich	
	1	02	Frozen meal	
	296	03	Salad	
	146	04	Mixed dish	
	134	05	Mixed dish with additions	
	38	06	Soup or stew	
	1142	07	Beverage	
	300	08	Other foods eaten together	
	626	09	Bread with additions	
	529	10	Cereal with additions	
	317	11	Meat, fish, or poultry with additions	
	424	12	Vegetable with additions	
	46	13	Vegetable combination	
	110	14	Dessert	
	190	15	Tortilla with additions	
13-19		USDA Combination food code		See note
DRRCCODE	2074	1141130-9170001		
	8	9999999	Don't know	
	3174	Blank		

NHANES III Combination Foods Data File from the Dietary Recall

SECOND EXAM DATA

USDA SURVEY NUTRIENT DATABASE

Positions SAS name	Counts	Item description and code	Notes
20-23 DRRCGW	5256	CFF Total grams of food/bevarage consumed 0004-5415	See note
24-26 DRRCALCO	5256	CFF Alcohol (gm) 000-159	
27-30 DRRCCALC	5256	CFF Calcium (mg) 0000-2074	
31-35 DRRCCARB	5256	CFF Carbohydrate (gm) 00000-300.7	
36-39 DRRCCARO	5256	CFF Carotenes (RE) 0000-5908	
40-43 DRRCCHOL	5256	CFF Cholesterol (mg) 0000-1401	
44-46 DRRCCOPP	5256	CFF Copper (mg) 000-6.6	
47-50 DRRCFIBE	5256	CFF Total dietary fiber (gm) 0000-72.7	
51-54 DRRCFOLA	5256	CFF Folacin (mcg) 0000-1231	
55-58 DRRCIRON	5256	CFF Iron (mg) 0000-62.5	
59-62 DRRCKCAL	5256	CFF Food energy (kcal) 0003-4056	

NHANES III Combination Foods Data File from the Dietary Recall

SECOND EXAM DATA

USDA SURVEY NUTRIENT DATABASE

Positions SAS name	Counts	Item description and code	Notes
63-65 DRRCMAGN	5256	CFF Magnesium (mg) 000-486	
66-70 DRRCMFAT	5256	CFF Total monounsaturated fatty acids (gm) 00000-104.6	
71-74 DRRCNIAC	5256	CFF Niacin (mg) 0000-60.3	
75-78 DRRCPFAT	5256	CFF Total polyunsaturated fatty acids (gm) 0000-59.7	
79-82 DRRCPHOS	5256	CFF Phosphorus (mg) 0000-3124	
83-86 DRRCPOTA	5256	CFF Potassium (mg) 0000-3953	
87-91 DRRCPROT	5256	CFF Protein (gm) 00000-00250	
92-95 DRRCSFAT	5256	CFF Total saturated fatty acids (gm) 0000-91.2	
96-100 DRRCSODI	5256	CFF Sodium (mg) 00000-13972	
101-105 DRRCTFAT	5256	CFF Total fat (gm) 00000-237.4	
106-110 DRRCVAIU	5256	CFF Vitamin A (IU) 00000-83336	

NHANES III Combination Foods Data File from the Dietary Recall

SECOND EXAM DATA

USDA SURVEY NUTRIENT DATABASE

Positions SAS name	Counts	Item description and code	Notes
111-115 DRRCVARE	5256	CFF Vitamin A (RE) 00000-24763	
116-119 DRRCVB1	5256	CFF Thiamin (mg) 0000-5.87	
120-123 DRRCVB2	5256	CFF Riboflavin (mg) 0000-4.94	
124-127 DRRCVB6	5256	CFF Vitamin B6 (mg) 0000-6.63	
128-133 DRRCVB12	5256	CFF Vitamin B12 (mcg) 000000-146.28	
134-136 DRRCVC	5256	CFF Ascorbic acid (mg) 000-528	
137-140 DRRCVE	5256	CFF Vitamin E (alpha tocopherol equivalents) 0000-70.2	
141-144 DRRCWATE	5256	CFF Water (gm) 0000-5362	
145-148 DRRCZINC	5256	CFF Zinc (mg) 0000-35.9	

NOTES

DRRMN: Meal number

Meal numbers in recalls that were complete and reliable (DRRSTAT=1) always begin with meal number=1; meal numbers increase by one for each consecutive meal or snack reported during the dietary interview. If a recall was coded reliable, but incomplete (DRRSTAT=2), the meal numbers may not be consecutive; information is reported for meals that were reported during the dietary interview. Meal numbers are not sorted by the time of day.

DRRFN: Food number

Every food has a food number. Foods are numbered within meals. If the recall was coded complete and reliable (DRRSTAT=1), the first food in each meal has a food number=1, and the other foods reported in the same meal are numbered consecutively. If the recall was coded reliable, but incomplete (DRRSTAT=2), the food numbers may not be consecutive; information is reported for all foods that were reported by the respondent.

DRRTCC: Combination type code

Fifteen combination food types were used to code the multi-component combination foods reported during the survey. Each DRRTCC has a generic description associated with it as follows:

*01Sandwich
02Frozen meal
03Salad
*04Mixed dish
*05Mixed dish with additions
*06Soup/Stew
07Beverage
*08Other foods eaten together
09Bread with additions
10Cereal with additions
*11Meat, fish, or poultry with additions
12Vegetables with additions
13Vegetable combination
14Dessert
*15Tortilla with additions

* Assigned combination food codes (DRRCCODE)

DRRCCODE: Combination food code

NCHS assigned combination food codes to 7 of the 15 combination food types. DRRCCODE matches the primary file variable, DRPCCODE, which is linked to text descriptions in a look-up table called "CODEBOOK".