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

NHANES III Healthy Eating Index Data File

Series 11, No. 6A

January 2000

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 or through other release mechanisms such as the Internet.

Table 1. Available NHANES III Data

Dataset Name	Release Date	Size in Megabytes	Data Files / Description
NHANES III, 1988-94, Series 11, No. 6A, ASCII Version (this release)	January 2000	2.7	Healthy Eating Index (HEI) Data File and documentation includes number of servings by Food Guide Pyramid food groups and HEI
NHANES III, 1988-94, Series 11, No. 5A, ASCII Version	TBD	54	NHANES III Supplemental Nutrition Survey of Older Americans (SNS), dietary intake data and documentation for a special dietary follow-up study of NHANES III, phase 1 (1988-91) examinees
NHANES III, 1988-94, Series 11, No. 4A, ASCII Version	TBD	0.5	Priority toxicant reference range study data file and documentation
NHANES III, 1988-94, Series 11, No. 3A, ASCII Version	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 release, Series 11 No. 6A, contains information on the number of servings
from the five different Food Guide Pyramid food groups, consumed by
participants. It also includes the Healthy Eating Index, a measure of diet
quality. This release does not replace the previous NHANES III data releases
(Series 11 Nos. 1A, 2A, 3A, 4A and 5A).

Table 2. 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
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

Data File

Topic	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

                                  Data File

Topic	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 Healthy Eating Index Data 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.

Information regarding a bibliography (on disk) of journal articles citing data from all the NHANES and the availability of NHANES III data on CD-ROM 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/nchs/nhanes.htm

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.

GUIDELINES FOR DATA USERS

Please refer to the following important information before analyzing data.

NHANES III Background Documents

o   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.

o   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.

o   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

o   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.

o   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.

o   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.

o   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.

o   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.
o   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.

o   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.

o   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.

o   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.

o   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

o   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.

o   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

o   In publications, please acknowledge NCHS as the original data source.
     For instance, the reference for the NHANES III Healthy Eating Index
     Data File 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 Healthy Eating Index Data
     File (Series 11, No. 6A). Hyattsville, MD.: Centers for Disease
     Control and Prevention, 1999.

o   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
LU             Lutein/zeaxanthin
LY             Lycopene
```

```
CODE           TOPIC


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
TO             Toxoplasmosis antibody
```

```
CODE          TOPIC

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.

NHANES III: The Healthy Eating Index

NCHS would like to recognize U.S. Department of Agriculture Center for Nutrition Policy and Promotion staff for developing the Healthy Eating Index (HEI) and for their invaluable technical assistance in finalizing the third National Health and Nutrition Examination Survey HEI files.

NHANES III: The Healthy Eating Index

GENERAL INFORMATION

Introduction

The role of nutrition and diet in reducing the risk of chronic diseases, such as cardiovascular disease, diabetes mellitus, and certain forms of cancer, has been well documented. Concerns about the influence of dietary practices on the health status of the country have been increasingly emphasized in a number of reports and guidelines, among them the Dietary Guidelines for Americans (1, 2, 3, 4), Healthy People 2000: National Health Promotion and Disease Prevention Objectives (5), Surgeon General's Report on Nutrition and Health (6), and National Research Council's report on Diet and Health (7).

Recommended dietary practices generally include the selection of foods from a variety of food groups, particularly the grain, fruit, and vegetable groups; choice of a diet that is low in fat, saturated fat, and cholesterol; and moderate use of salt and sodium.

Much of the previous research has focused on determining which eating patterns should be recommended to the public for reducing the risk of chronic disease. Some analytical instruments have been developed which evaluate specific dietary components, such as fat and cholesterol. However, few instruments, like the Healthy Eating Index, have been developed to assess the overall quality of a diet. Practical applications of the Healthy Eating Index results include improved targeting of nutrition education and health promotion activities and, as a single summary measure of diet quality, the Healthy Eating Index can be used to monitor changes in food consumption patterns over time. (See the HEI report: http://www.nal.usda.gov/fnic/HEI/HEI.html).

HEI Overview

In an effort to measure how well American diets conform to recommended healthful eating patterns, the U.S. Department of Agriculture developed an Index, called the Healthy Eating Index (8). The Index was designed to provide a measure of the overall quality of an individual's diet.

The Index provides a picture of foods people are eating, the amount of variety in the diet and compliance with specific Dietary Guidelines recommendations. In designing the Index, it was not intended that the dietary components be considered individually as indicators of overall dietary quality. A score on the Index reflects a sum total of the dietary components. The following 10 dietary components were included in the Index based on different aspects of a healthful diet:

0   Components 1-5 measure the degree to which a person's diet conforms to the Food Guide Pyramid serving recommendations for the Grain, Vegetable, Fruit, Milk, and Meat Groups.

0   Component 6 measures total fat consumption as a percentage of total food energy intake.

0   Component 7 measures saturated fat consumption as a percentage of total food energy intake.

0   Component 8 measures total cholesterol intake.

0   Component 9 measures total sodium intake.

0   Component 10 examines the amount of variety in a person's diet over a 1-day period.

The HEI is a practical approach to measuring overall dietary quality in that it examines specific food behaviors which comprise dietary intake. It was first used to assess the quality of diets reported by respondents of the 1989-90 USDA Continuing Survey of Food Intakes by Individuals (CSFII). HEI scores have also been calculated for the 1994-96 CSFII.

In this study, the HEI was applied to the NHANES III dietary intake data (1988-94). The NHANES III data set contains complete food and nutrient intake records (DRPSTAT=1)for 26,350 individuals ages 2 and above. Two of these individuals fasted on the survey day. HEI and component scores were computed for the remaining 26,348 individuals. HEI and component (nutrient, food group, variety) scores for NHANES III respondents were processed using queries developed in Microsoft Access, a relational database system. Query strategies are documented in detail in report by the contractor, TAS-ENVIRON (9).

Data Used To Calculate the Healthy Eating Index

The third National Health and Nutrition Examination Survey was designed to describe the health and nutritional status of the U.S. civilian noninstitutionalized population. The survey included a household interview and a physical examination conducted in a mobile examination center. The examination included a 24-hour dietary recall, body measurements, blood and urine tests, and recording of numerous other variables based on a physical examination (10).

NHANES III was designed as a 6-year survey and was conducted in two 3-year phases. Both phases, as well as the entire 6-year assessment, included nationally representative samples of the U.S. population aged >=2 months living in households. Children aged 2 months to 5 years and persons aged >=60 years were oversampled, thereby expanding the age range in comparison with that in the previous surveys. Blacks and Mexican Americans were also oversampled (10).

Findings

The higher the score on the Healthy Eating Index, the better the diet conformed to the Dietary Guidelines for Americans and the Food Guide Pyramid. The average Healthy Eating Index score was 63.8 out of a possible 100 for the period 1988-1994.

Mean (weighted) HEI and component scores for the U.S. population (ages 2 years and older) are as follows:

Sample size:	26,348
Fat Score:	6.5
Saturated Fat Score:	6.1
Cholesterol Score:	7.8
Sodium Score:	6.0
Grain Score:	6.7
Fruit Score:	3.8
Vegetable Score:	5.7
Dairy Score:	6.6
Meat Score:	6.8
Variety Score:	7.7
HEI Score:	63.8

Determination of Food Guide Pyramid Servings

Servings used to compute the Index scores were intended to be as consistent as possible with those presented in the Food Guide Pyramid (http://www.nal.usda.gov/fnic/etext/000023.html). However, Pyramid guidance regarding servings is available only for a limited range of foods, and in most cases the servings are commodity-based. Foods reported by survey respondents, on the other hand, often are complex mixtures of foods, and Pyramid serving intakes from these foods are difficult to quantify. Therefore, Pyramid group servings in foods reported by survey respondents were determined by identifying the key underlying commodities contained in these foods and determining appropriate servings conversion factors. Although Pyramid servings assigned to USDA foodcode-commodity combinations were used wherever possible, the USDA mapping scheme was modified to include servings for foods/foodcodes reported only in NHANES III based on similarity to USDA foodcodes.

Serving amounts for breads and grains were determined according to an "equivalent flour" approach. For example, the Food Guide Pyramid designates a slice of bread as one serving. A typical slice of bread contains 17 grams of flour. The number of servings for any bread on the Index was calculated based on the number of grams of flour it contained divided by the number of grams of flour contained in a slice of bread (12). Similar approaches were applied to pastas and cereal grains.

The Food Guide Pyramid counts cup of most cooked vegetables, 1 cup of most raw leafy vegetables, and cup of most raw non-leafy chopped vegetables as single servings. Different vegetables have different densities, resulting in different gram weights per cup or half-cup measures. Different gram/serving size factors were used to calculate the index for most vegetables, based on the weight of a cup or half-cup of the relevant commodities.

Fruits were treated similarly to vegetables. Gram/serving size factors were developed for each fruit based on the weights of the various fruit amounts.

Serving amounts for various kinds of milk and milk products were calculated based on the grams of nonfat milk solids contained in a food divided by the amount of grams of nonfat milk solids contained in 1 cup of milk (the serving specified for milk in the Food Guide Pyramid). To determine a serving for different cheeses, the weight of all milk products in a cheese was totaled and then divided by the conversion factor used in the Food Guide Pyramid for cheese.

Serving sizes of meats are specified in the Food Guide Pyramid in terms of 2-3 ounces of lean meat. The Index calculations assume a serving size of 2.5 ounces for meats. Serving size conversion factors for meats were based on the amount of fat-free meat commodity included in the database for various foods. Conversion factors for converting grams of nuts and peanut butter to serving sizes were developed according to those specified in the Food Guide Pyramid. The gram conversion factors developed were based on the weight of these quantities.

The Food Guide Pyramid indicates that legumes are to be considered contributors to meat group intake, but it was determined in work on USDA HEIs that any legume consumption not "needed" for achieving recommended meat group servings could be applied toward vegetable group servings. Legume consumption by an individual therefore was counted either as meat group consumption or vegetable group consumption, depending on the amount of meat consumed from other sources.

Pyramid serving sizes for food groups other than dairy were determined separately for survey respondents ages 2-3 and for respondents ages 4 and older. It was determined that fruit, grain, legume, meat, and vegetable serving sizes for 2 and 3 year olds would be 2/3 the size of serving sizes for respondents ages 4 and older.

Methodology Difference

The methodology used to determine serving definitions for each of the five major food groups in this report is the same as that used to calculate the 1989-90 HEI (13) using CSFII data but different from that developed by USDA's Agricultural Research Service (14) and used to calculate the 1994-96 HEI. In particular, milk serving definitions in this report were based on grams of nonfat milk solids contained in a food divided by the amount of grams of nonfat milk solids contained in 1 cup of milk. For the 1994-96 HEI, milk serving definitions were based on the Pyramid's underlying criterion for a milk serving, which is that it should provide about the same amount of calcium as 1 cup of skim milk, or 302 milligrams. This approach, while more in line with the advice of nutrition researchers, has implications for lower milk group component scores. This is due to the omission of some foods, previously counted (e.g., butter and cream cheese), based on nonfat milk solids, but no longer given credit towards a milk serving because they do not meet the calcium criterion of the Pyramid. For a complete description and documentation of the Food Guide Pyramid servings, see reference 12.

Allocation of Mixtures to Individual Food Groups

In calculating the Index, it was necessary to assign the foods in "mixtures" to their constituent food groups in the appropriate amounts. This was accomplished through the use of a "recipe file." Pizza, for instance, may make significant contributions to several different food groups, including grain, vegetable,

milk, and meat.

Recipes previously written for USDA foodcodes by a contractor were used in the NHANES recipe file wherever possible. The contractor obtained most NHANES III recipes from the 1994-96 CSFII recipe database, going back to the 1989-91 CSFII and/or the 1987-88 NFCS recipe databases only to obtain recipes for foodcodes not in continued use.

Recipe sources for NHANES III Database

Survey	Number of recipes used in NHANES III recipe database
1994-96 CSFII	3381
1989-91 CSFII	808
1987-88 NFCS	385
NHANES III (new)	187

The approach used was a straight-forward extension of the approach used to estimate serving sizes. Commodity compositions of foods were identified. Once identified, commodities were assigned to appropriate food groups, based on the gram/serving size factor calculated.

Estimation of Food Group Serving Requirements by Age and Gender

Prior to scoring the first five components of the Index, it was necessary to determine the recommended numbers of servings by food group for each of the individuals on the data file. The food energy RDAs (15) for some age/gender combinations were different from the three levels of energy intake (1,600, 2,200, and 2,800 kilocalories) presented in the Food Guide Pyramid (See Table 1). Interpolation techniques were used to estimate the required number of servings for age/gender combinations not addressed in the Food Guide Pyramid. Food servings specified in the Food Guide Pyramid for three food energy RDA levels were used as a basis for predicting comparable food servings at other food energy levels for each food group.

Two issues arose in taking this approach. Children 2-3 years of age have a food energy RDA less than the lowest calorie level in the Food Guide Pyramid. Extrapolation of the Food Guide Pyramid's recommended number of servings to a lower calorie level would result in smaller numbers of servings than the minimums shown. The following statement from the Food Guide Pyramid provided a basis for the technical approach taken to address the issue: "Preschool children need the same variety of foods as older family members do, but may need less than 1,600 calories. For fewer calories, they can eat smaller servings."

The number of servings for children 2-3 years of age was thereby held constant at the minimums shown in the Pyramid, but the serving sizes were reduced proportionately.

Similarly, males 15 - 50 years of age have food energy RDAs slightly higher than the highest calorie level in the Food Guide Pyramid. Simple extrapolation would result in larger numbers of servings than the maximum numbers shown. The text of the Food Guide Pyramid provides no guidance regarding the adjustment of numbers of servings or serving sizes to accommodate higher food energy levels. Rather than exceeding recommended serving sizes, it was decided that food servings would be truncated at the maximums shown in the Food Guide Pyramid. Preliminary analysis indicated that if serving sizes had been slightly increased, the results obtained from the Index would not have been significantly different.

HEI Scoring

For each of the five food group components of the Index, individuals who
consumed the recommended number of servings received a maximum score of 10. A
score of zero was assigned to any food group where no items from that category
were consumed. Intermediate scores were calculated proportionately to the
number of servings consumed. For example, if the recommended level of servings
was eight and an individual consumed four servings, the component score for the
individual was 5 points. A score of 7.5 points was assigned if six servings
were eaten. Table 2 shows the components of the HEI and scoring criteria.

Variety Score Calculation

NHANES III foodcodes for similar foods were grouped for the purpose of assessing
dietary variety. The following principles guided the coding decisions made:

0   Only foods judged nutritionally similar were grouped together.
0   Foods made with separate commodities were generally grouped separately.
0   Foods differing only in fat content were generally grouped together.
0   Vegetables were each given separate codes, but different forms of the same
     vegetable were coded together.
0   Different forms of the same meat were generally coded the same. Organ meats
     and ham were two exceptions.
0   Each type of fish was given a separate code, but different forms of the same
     fish received the same code.
0   Most forms of fluid milk had the same code. Pudding was coded separately
     from milk.
0   Most cheeses had the same code with the exception of cottage cheese.
0   All white bread was given the same code. Sweet rolls and pasta received
     different codes.
0   Whole wheat products were coded differently from products made with refined
     wheat flour.
0   Ready-to-eat cereals made principally from the same grain received the same
     codes; those made from different grains received different codes.

Intakes of fats, sweets, seasonings, and similar foods were not included in
variety score calculations. The list of NCHS foodcodes not included in variety
score calculations appears in the Appendix.

Food components contributing less than a Food Guide Pyramid serving were not
computed in the Index variety score. It is possible that variety scores for
some individuals were slightly underestimated by this approach. Several
servings in a food group consumed in less than serving amounts throughout a
day could exceed the serving threshold when added together. However, the
potential effects of this underestimation on the final results is believed to
be extremely small.

A second simplifying convention made the assumption that food mixtures
containing two or more components from the same food group (e.g., mixed
vegetables) could be reasonably allocated, equally, to two codes representing
the components present in the highest proportions. Note that mixtures with
multiple ingredients in the same pyramid group may overestimate servings, due to
the structure of the pyramid servings database, and the difficulty of separating
out the amount of each ingredient. However, when ingredients are from different
pyramid groups, the program accounts for each appropriately.

A contractor used the foodcode-varcode maps used in calculating variety scores for USDA survey respondents (See Appendix), with mapping records added for NHANES-only foods. The contractor used the variety scoring scale developed for use in scoring CSFII one-day intakes, as shown below:

Design Alternatives: What To Count

Foods often principally fall within one group but contain small amounts of other groups. Bread, for example, is mainly a grain but also contains small amounts of milk and egg products. An approach considered for this study was to exclude the "incidental" food group contributions from the computations. This approach was rejected based on the following rationale:

0   Even relatively small amounts of incidental foods contribute to an
    individual's overall nutrient intake.
0   Suppressing the incidental foods would have often involved arbitrary
    judgements for establishing minimum size cutoffs.

It was decided that all contributions to various food groups would be counted in computing the Index with no minimum size cutoff values imposed. The following examples illustrate some of the implications of this approach:

0   The nutrition value from condiments, such as mayonnaise, was counted in
    computing the Index.
0   The nutrition value of milk used in some sweets, such as a milk chocolate
    bar, was counted in the milk group in computing the Index. If allocated to
    a single food group, the overall food would be assigned to the "Sweets"
    group and not counted in the Index.
0   The fruit juice in a soft drink which is at least 10 percent fruit juice was
    counted in computing the Index. The water and sugar contained in the soft
    drink were not counted.
0   The potato content of potato chips was counted in computing the Index. The
    fat content was not counted in computing the vegetable and variety
    components of the Index but was counted in computing the fat component.

Other Components

In developing the Index, consideration was given to including a component to address food energy intake. Obesity is a significant health problem in this country. Nevertheless, it was decided that the inclusion of physical measures of appropriate body weight, such as a body mass index (BMI) or conformance to standard weight-for-height tables, would be inappropriate since they are influenced by factors, such as physical activity, unrelated to people's eating patterns. Use of a measure based on food energy in relation to the RDA was also rejected for this study, as preliminary tabulations of the data indicated that these measures were not highly correlated with physical measures of obesity.

NOTE TO RESEARCHERS:

In order to analyze these data, you must first link this file with the NHANES
III household interview, exam, and/or lab files described at the beginning of
this document. This is where you will find sociodemographic and
health/nutrition variables, as well as sample weights. Before getting started,
we encourage you to read all of the documentation included along with this file.
You should also read the analytic guidelines and other documentation provided in
the NHANES III Reference Manuals and Reports CD-ROM.

You can order our data and manuals CD-ROMS by calling 1-301-458-4636 or by going
directly to the CDC homepage (http://www.cdc.gov/publications.htm).

Table 1. Recommended number of USDA Food Guide Pyramid servings per day, by
age/gender categories.

Age/gender category	Energy (kcalories)	Grains	Vegetables	Fruits	Milk	Meat(A)
Children 2-3(B)	1300	6.0	3.0	2.0	2.0	2.0
+++(C)	1600	6.0	3.0	2.0	2.0	2.0
Children 4-6	1800	7.0	3.3	2.3	2.0	2.1
Females 51+	1900	7.4	3.5	2.5	2.0	2.2
Children 7-10	2000	7.8	3.7	2.7	2.0	2.3
Females 11-24	2200	9.0	4.0	3.0	3.0	2.4
+++(C)	2200	9.0	4.0	3.0	2.0	2.4
Females 25-50	2200	9.0	4.0	3.0	2.0	2.4
Males 51+	2300	9.1	4.2	3.2	2.0	2.5
Males 11-14	2500	9.9	4.5	3.5	3.0	2.6
+++(C)	2800	11.0	5.0	4.0	2.0	2.8
Males 19-24	2900	11.0	5.0	4.0	3.0	2.8
Males 25-50	2900	11.0	5.0	4.0	2.0	2.8
Males 15-18	3000	11.0	5.0	4.0	3.0	2.8

(A) One serving of meat equals 2.5 ounces of lean meat.

(B) Portion sizes were reduced to two-thirds of adult servings except for milk for children age 2-3.

(C) Recommended number of servings per day at food energy levels specified in the Food Guide Pyramid.

Table 2. Components of the Healthy Eating Index and scoring system

	Score ranges(A)	Criteria for maximum score of 10	Criteria for minimum score of 0
Grain consumption	0 to 10	6-11 servings(B)	0 servings
Vegetable consumption	0 to 10	3-5 servings(B)	0 servings
Fruit consumption	0 to 10	2-4 servings(B)	0 servings
Milk consumption	0 to 10	2-3 servings(B)	0 servings
Meat consumption	0 to 10	2-3 servings(B)	0 servings
Total fat intake	0 to 10	30% or less energy from fat	45% or more energy from fat
Saturated fat intake	0 to 10	>10% energy from saturated fat	15% or more energy from saturated fat
Cholesterol intake	0 to 10	300 mg or less	450 mg or more
Sodium intake	0 to 10	2400 mg or less	4800 mg or more
Food variety	0 to 10	8 or more different items in a day	3 or fewer different items in a day

(A) People with consumption or intakes between the maximum and minimum ranges or amounts were assigned scores proportionately.

(B) Number of serving depends on Recommended Energy Allowance See Table 1. All amounts are on a per day basis.

Table 3. Counts of component scores for the NHANES III population are as follows:

HEI	Fat	Saturated fat	Cholesterol	Sodium	Dairy
0	2326	4726	4558	4061	645
>0 & <1	771	657	278	637	2423
>=1 & <2	1053	1225	306	703	1905
>=2 & <3	1390	1112	293	846	1703
>=3 & <4	1387	1244	322	957	1709
>=4 & <5	1647	1288	365	1127	1605
>=5 & <6	1812	1110	390	1288	1693
<=6 & <7	1834	1762	460	1442	1541
<=7 & <8	1907	1369	445	1627	1404
<=8 & <9	1792	1062	548	1685	1229
<=9 & <10	1700	1672	541	1880	1165
10	8727	9121	17842	10095	9326
ALL	26348	26348	26348	26348	26348

Table 3 Continued. Counts of component scores for the NHANES III population are as follows:

HEI	Fruit	Grain	Meat	Vegetables	Variety
0	6688	181	222	1753	2862
>0 & <1	3969	377	928	2074	0
>=1 & <2	1326	918	1153	2215	0
>=2 & <3	1335	1617	1556	2314	2785
>=3 & <4	1467	2211	1826	2286	0
>=4 & <5	1427	2696	2099	2298	3668
>=5 & <6	1201	2928	2136	1939	0
<=6 & <7	1134	2768	2041	1818	4002
<=7 & <8	928	2480	2049	1494	0
<=8 & <9	895	2165	1815	1353	3767
<=9 & <10	731	1896	1622	1136	0
10	5247	6111	8901	5668	9264
ALL	26348	26348	26348	26348	26348

APPENDIX

GENERIC NAMES FOR FOOD CODES TO WHICH
ALL USDA CODES USED IN THE VARIETY INDEX WERE MAPPED

MILK GROUP:

     1111600 MILK, GOAT'S
     1110000 MILK AND MILK PRODUCTS, PLAIN
     1151100 MILK AND MILK PRODUCTS, SWEETENED AND/OR FLAVORED
(INCLUDE
             CHOCOLATE MILK, NS AS TO MILK OR DRINK)
     1320011 PUDDING
     1321030 CUSTARD
     1321050 PUDDING, TAPIOCA
     1325000 MOUSSE
     1321011 PUDDING, BREAD
     1321041 PUDDING, RICE
     1321061 PUDDING, COCONUT
     1321071 PUDDING, INDIAN
     1321075 PUDDING, PUMPKIN
     1401000 CHEESE
     1420010 CHEESE, COTTAGE

MEAT POULTRY, FISH, DRY BEANS, EGGS AND NUTS:

                    MEAT AND POULTRY

     2100010 BEEF
     2200010 PORK
     2230012 HAM
     2300010 LAMB
     2315010 GOAT
     2320010 VEAL
     2331000 RABBIT
     2332100 VENISON
     2332310 MOOSE
     2332350 BEAR
     2332410 CARIBOU
     2333110 GROUND HOG
     2333210 OPOSSUM
     2333310 SQUIRREL
     2333410 BEAVER
     2333510 RACCOON
     2334010 ARMADILLO
     2334510 WILD PIG
     2410000 CHICKEN
     2420100 TURKEY
     2430010 DUCK
     2431101 GOOSE
     2440000 CORNISH GAME HEN

MEAT, POULTRY, FISH, DRY BEANS, EGGS AND NUTS (cont'd):

MEAT AND POULTRY (cont'd)

```
2440210 DOVE
2440310 QUAIL
2440410 PHEASANT (INCLUDE GROUSE)
2510001 LIVER
2512000 HEART
2513000 KIDNEY
2514011 SWEETBREADS
2515000 BRAINS
2516000 TONGUE
2517011 TRIPE
2517041 GIBLETS
2521011 FRANKFURTER, WIENER, OR HOT DOG
2522001 COLD CUT
2522010 SAUSAGE
```

FISH, SHELLFISH, AND OTHER SEAFOOD

```
2610010 FISH
2610111 ANCHOVY
2610311 BARRACUDA
2610511 CARP (INCLUDE BREAM, BUFFALO FISH, CHUB, SUCKER)
2610711 CATFISH (INCLUDE BULLHEAD)
2610911 COD
2611111 CROAKER (INCLUDE ANGELFISH, BUTTERFLY FISH, DRUMFISH,
        GOATFISH, KINGFISH, SEA TROUT, FRESHWATER SHEEPSHEAD,
        SPADEFISH, SPOT, SURGEONFISH, WEAKFISH, WEKE)
2611311 EEL
2611500 FLOUNDER (INCLUDE DAB, FLUKE, HALIBUT, SOLE, TURBOT)
2611711 HADDOCK (INCLUDE BURBOT, CUSK, HAKE, LING, MONKFISH,
        POLLOCK, SCROD)
2611910 HERRING (INCLUDE ALEWIFE, MILKFISH, SHAD)
2612110 MACKEREL (INCLUDE ENENUI, GARFISH, ONO, NEEDLEFISH, WAHOO)
2612310 MULLET
2612510 OCEAN PERCH (INCLUDE BOCACCIO, MENPACHI, ORANGE ROUGHY,
        REDFISH, ROCKFISH)
2612711 PERCH (INCLUDE FRESHWATER BASS, BLUEGILL, CRAPPIE, SUNFISH,
        WALLEYE)
2612911 PIKE (INCLUDE MUSKELLUNGE, PICKEREL)
2613110 POMPANO (INCLUDE AKULE, BLACKFISH, BLUEFISH, BUTTERFISH,
        DOLPHINFISH, JACK, MAHIMAHI, PAPLO, PARROT FISH, SABLEFISH,
        SCAD, TILEFISH, ULVA, YELLOWTAIL)
2613310 PORGY (INCLUDE SCUP, SEA BREAM, MARINE SHEEPSHEAD, SNAPPER)
2613511 RAY (INCLUDE SKATE)
```

MEAT, POULTRY, FISH, DRY BEANS, EGGS AND NUTS (cont'd):

FISH, SHELLFISH, AND OTHER SEAFOOD (cont'd)

```
2613710 SALMON (INCLUDE SALTWATER TROUT)
2613911 SARDINES
2614111 SEA BASS (INCLUDE GROUPER  STRIPED BASS, WREAKFISH)
2614311 SHARK (INCLUDE DOGFISH, GRAYFISH)
2614511 SMELT (INCLUDE CAPELIN)
2614711 STURGEON
2614911 SWORDFISH (INCLUDE MARLIN)
2615111 TROUT (INCLUDE CISCO, LAKE HERRING, STEELHEAD, WHITEFISH)
2615310 TUNA (INCLUDE AHI, AKU, BONITO)
2615711 WRITING
2620311 FROG LEGS
2620418 JELLYFISH
2620511 OCTOPUS
2620711 ROE, COD AND SHAD
2621210 SEA URCHIN (ROE)
2621310 SQUID (INCLUDE CUTTLEFISH)
2621512 TURTLE (TERRAPIN)
2630111 ABALONE
2630310 CLAMS
2630511 CRAB (INCLUDE WHITE AND KING CRAB MEAT)
2630914 CRAYFISH
2631111 LOBSTER (INCLUDE SPINY LOBSTER; ROCK LOBSTER)
2631310 MUSSELS
2631510 OYSTERS
2631711 SCALLOPS
2631911 SHRIMP (INCLUDE PRAWN)
2632111 SNAILS
2840000 GELATIN
```

EGGS

```
3110101 EGG
3300010 EGG SUBSTITUTE
```

DRY BEANS, NUTS, SEEDS

```
4110100 BEANS, DRY
4110700 SOYBEANS
4130100 DRIED PEAS, LENTILS COWPEAS, DRY (INCLUDE BLACKEYE PEAS,
        FIELD PEAS)
4141001 SOYBEAN DERIVED PRODUCTS:  SOY NUTS (INCLUDE PERNUTS)
4142001 BEAN CURD (TOFU)
```

MEAT, POULTRY, FISH, DRY BEANS, EGGS AND NUTS (cont'd):

DRY BEANS, NUTS, AND SEEDS (cont'd)

```
4143000 PROTEIN SUPPLEMENT, MEAL REPLACEMENT SOY-BASED PROTEIN
        POWDER
4146001 HI-PROTEIN WAFERS(INCLUDE HI-PROTEIN OATMEAL WAFER)
4181020 HEAT SUBSTITUTES, MAINLY VEGETABLE PROTEIN
        BACON STRIPS, MEATLESS(INCLUDE MORNING STAR BREAKFAST
        STRIPS, STRIPPLE)
4148000 TOFU (INCLUDE TOFUTTI)
4210010 ALMONDS
4210200 BRAZIL NUTS
4210300 BUTTER NUTS
4210400 CASHEW NUTS
4210500 CHESTNUTS
4210600 COCONUT MEAT
4210700 FILBERTS, HAZEL NUTS
4210750 GINKGO NUTS
4210800 HICKORY NUTS
4210900 MACADAMIA NUTS
4211000 MIXED NUTS
4211100 PEANUTS
4211200 PECANS
4211300 PINE NUTS (PIGNOLIAS)
4211400 PISTACHIO NUTS
4211600 WALNUTS
4310010 BREADNUTS
4310100 PUMPKIN AND/OR SQUASH SEEDS
4310200 SUNFLOWER SEEDS, HULLED
4310300 SESAME SEEDS
4310400 FLAX SEEDS
4310700 MIXED SEEDS
4410100 CAROB POWDER OR FLOUR
```

BREADS, CEREALS, RICE, AND PASTA:

```
5100010 BREADS, ROLLS, WHITE FLOUR BASED,
5116000 BREADS, ROLLS, SWEET OR WITH ADDED INGREDIENTS
5120101 BREAD, ROLLS, WHOLE WHEAT, 100%
5120111 BREAD, WHOLE WHEAT, 100%, WITH RAISINS
5130011 BREAD, WHOLE WHEAT
5130022 BREAD, WHOLE WHEAT, NS AS TO 100%, WITH RAISINS
5130201 BREAD, BRAN (INCLUDE GRANOLA, BRANOLA, HONEY BRAN)
5130252 MUFFIN, ENGLISH, BRAN, WITH RAISINS
5140101 BREAD, ROLLS, RYE/PUMPERNICKEL (INCLUDE CORN RYE)
5150101 BREAD, ROLLS OATMEAL
5160101 BREAD, ROLLS, MULTI-GRAIN
```

BREADS, CEREALS, RICE, AND PASTA (cont'd):

```
5160121 BREAD, MULTI-GRAIN, WITH RAISINS
5180602 BREAD, RICE
5220100 CORNBREAD CORNMEAL-BASED(INCLUDE JALAPENO CORNBREAD)
5510500 PANCAKES, BUCKWHEAT
5310010 CAKE, WHEAT FLOUR
5311540 CAKE, OATMEAL
5312410 CAKE, ZUCCHINI
5320100 COOKIE
5330572 PIE
5334700 PIE, PUMPKIN
5440300 POPCORN
5610100 MACARONI(INCLUDE LASAGNA NOODLES, ORZO, ZITI, ROTINI,
        SHELLS, WAGON WHEELS, CART WHEELS, MANICOTTI, RIGATONI,
        MOSTACCIOLI, CAVATONI, RICCI, BOWS, TWIRLS, SPIRALS)
5610200 MACARONI, WHOLE WHEAT (INCLUDE LASAGNA NOODLES, ORZO, ZITI,
        ROTINI, SHELLS, WAGON MANICOTTI, RIGATONI, MOSTACCIOLI,
        CAVATONI RICCI, BOWS, TWIRLS, SPIRALS)
5614000 PASTA, CORN-BASED (INCLUDE MACARONI, NOODLES, SPAGHETTI)
5611700 RICE NOODLES
5620040 BARLEY (INCLUDE EGG BARLEY)
5620050 BUCKWHEAT GROATS (INCLUDE KASHA)
5620100 GRITS
5620200 MILLET
5620300 OATMEAL
5620500 RICE (INCLUDE YELLOW RICE)
5620511 RICE, BROWN
5620521 RICE, WILD
5620700 WHEAT, CREAM OF (INCLUDE FARINA)
5620711 BULGUR (INCLUDE WHEAT PILAF)
5620720 WHOLE WHEAT CEREAL(INCLUDE WHEATENA, RALSTON, ZOOM,
ROMAN
        MEAL, BRANOLA, HOME GROUND CEREAL, ROMAN MEAL WITH OATS)
5620800 MULTI-GRAIN CEREAL (INCLUDE SEVEN-GRAIN CEREALS)
```

FRUIT:

```
6110050 CALAMONDIN
6110101 GRAPEFRUIT (INCLUDE CHIRONJA)
6111001 KUMQUAT
6111301 LEMON
6111601 LIME
6111901 ORANGE
6112500 TANGELO
6112501 TANGERINE (INCLUDE MANDARIN ORANGE, FRESH; SATSUMA)
6120050 ACEROLA JUICE
```

FRUIT (cont'd):

```
6120100  GRAPEFRUIT JUICE
6120400  LEMON M CE
6120700  LIME JUICE
6121000  ORANGE JUICE
6121300  TANGERINE JUICE
6210100  FRUIT, DRIED
6310301  APRICOT
6310701  BANANA
6311701  CURRANTS
6211010  DATE
6311901  FIG
6312651  LYCHEE
6313301  PAPAYA
6313501  PEACH
6314101  PINEAPPLE
6314301  PLUM
6212210  PRUNE
6212510  RAISINS
6212600  TAMARIND PULP, DRIED
6310100  APPLE
6310301  APRICOT
6310501  AVOCADO
6310901  CANTALOUPE (MUSKMELON)
6310975  CARAMBOLA (STARFRUIT)
6311101  CHERRIES, MARASCHINO
6311200  ACEROLA
6311301  CHERRIES
6312101  GENIP
6312300  GRAPES
6312501  GUAVA
6312521  JACKFRUIT
6312601  JUNEBERRY
6312650  KIWI FRUIT
6312701  HONEYDEW MELON
6312901  MANGO
6313101  NECTARINE
6313401  PASSION FRUIT
6313701  PEAR, RAW
6314501  POMEGRANATE
6314601  QUINCE
6314701  RHUBARB
6314790  SAPODILLA
6314801  SUGAR APPLE
6314850  SOURSOP (ANNONA MURICATA)
6212600  TAMARIND PULP, DRIED, SWEETENED ('PULPITAS')
6314901  WATERMELON
6320010  BERRIES
```

FRUIT (cont'd):

```
6320101 BLACKBERRIES (INCLUDE DEWBERRIES AND YOUNGBERRIES)
6320301 BLUEBERRIES
6320700 CRANBERRIES
6321101 ELDERBERRIES
6321301 GOOSEBERRIES
6321501 LOGANBERRIES
6321701 MULBERRIES
6321900 RASPBERRIES
6322302 STRAWBERRIES
6331100 FRUIT COCKTAIL OR MIX (EXCLUDING CITRUS FRUITS)
6212210 PRUNES
6340902 CHUTNEY
6410010 FRUIT JUICE
6410101 APPLE JUICE
6410460 BLACKBERRY JUICE
6410540 CRANBERRY JUICE
6411601 GRAPE JUICE
6412001 PAPAYA JUICE
6412100 PASSION FRUIT JUICE(INCLUDE YELLOW, LILIKOI)
6412203 PEACH JUICE
6412401 PINEAPPLE JUICE
6413201 PRUNE JUICE
6413250 STRAWBERRY JUICE
6413310 WATERMELON JUICE
6420101 APRICOT NECTAR
6420150 BANANA NECTAR
6420201 CANTALOUPE NECTAR
6420302 GUAVA NECTAR
6420401 MANGO NECTAR
6420501 PEACH NECTAR
6421001 PAPAYA NECTAR
6421301 PASSION FRUIT NECTAR
6421501 PEAR NECTAR
6422101 SOURSOP (GUANABANA) NECTAR
```

VEGETABLES:

```
7100010 WHITE POTATO
7190010 PLANTAIN
7193009 CASSAVA, YUCA BLANCA
7194501 YAM, PUERTO RICAN
7195001 TANNIER
7196201 DASHEEN
7210121 BEET GREENS
7210420 CHARD
7210720 COLLARDS
```

VEGETABLES (cont'd):

```
7211022 CRESS
7211310 DANDELION GREENS
7211600 ENDIVE, CHICORY, ESCAROLE, OR ROMAINE
7211921 KALE
7212220 MUSTARD GREENS
7212302 POKE GREENS
7212510 SPINACH
7212820 TURNIP GREENS
7213010 WATERCRESS
7220110 BROCCOLI
7310101 CARROTS
7320100 PUMPKIN
7330100 SQUASH, WINTER TYPE
7340100 SWEET POTATO
7410100 TOMATOES
7430110 TOMATO JUICE AND PLAIN SAUCES
7510050 ALFALFA SPROUTS
7510075 ARTICHOKE, JERUSALEM
7510080 ASPARAGUS
7510100 BEAN SPROUTS
7510180 BEANS, STRING, GREEN
7510250 BEETS
7510300 CABBAGE
7510400 CABBAGE, CHINESE
7510500 CABBAGE, RED
7510700 CAULIFLOWER
7510900 CELERY
7510950 CHIVES
7510960 CORN
5711100 CUCUMBER
7511150 GARLIC
7511180 JICAMA
7511200 OHLRABI
7511300 LETTUCE
7511306 LETTUCE, BOSTON
7511500 MUSHROOMS
7511701 ONIONS, YOUNG GREEN
7511702 ONIONS, MATURE
7511900 PARSLEY
7512000 PEAS, GREEN
7512100 PEPPER, HOT CHILI
7512210 PEPPER, SWEET GREEN
7512220 PEPPER, SWEET RED
7512500 RADISH
7512775 SNOW PEAS
7512800 SQUASH, SUMMER, YELLOW
7512900 TURNIP
```

VEGETABLES (cont'd):

```
7514300 LETTUCE SALAD
7520100 ARTICHOKE, GLOBE
7520300 BAMBOO SHOOTS
7520400 BEANS, LIMA
7520600 BEANS, STRING, YELLOW
7520829 BITTER MELON
7520900 BRUSSELS SPROUTS
7521311 CACTUS
7521700 EGGPLANT
7521730 FLOWERS OR BLOSSOMS OF SESBANIA,, SQUASH, OR LILY
7522000 OKRA
7522201 PARSNIPS
7522800 RUTABAGA
7523200 SEAWEED
7523320 SQUASH, SPAGHETTI
7523500 WATER CHESTNUT
7531100 MIXED VEGETABLES
7534000 VEGETABLE COMBINATION
7550309 HORSERADISH
7551000 OLIVES
```

Prepared for Mathematica by TAS, Inc.

REFERENCES

1. U.S. Department of Agriculture and U.S. Department of Health and Human Services. 1980. Nutrition and Your Health: Dietary Guidelines for Americans. (1st ed.). U.S. Department of Agriculture. Home and Garden Bulletin No. 232.

2. U.S. Department of Agriculture and U.S. Department of Health and Human Services. 1985. Nutrition and Your Health: Dietary Guidelines for Americans. (2nd ed.). U.S. Department of Agriculture. Home and Garden Bulletin No. 232.

3. U.S. Department of Agriculture and U.S. Department of Health and Human Services. 1990. Nutrition and Your Health: Dietary Guidelines for Americans. (3rd ed.). U.S. Department of Agriculture. Home and Garden Bulletin No. 232.

4. U.S. Department of Agriculture and U.S. Department of Health and Human Services. 1995. Nutrition and Your Health: Dietary Guidelines for Americans. (4 th ed.). U.S. Department of Agriculture. Home and Garden Bulletin No. 232.

5. U.S. Department of Health and Human Services. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. DHHS Publication No. (PHS) 91-50212. U.S. Government Printing Office, Washington, DC. 1991.

6. U.S. Department of Health and Human Services. 1988. The Surgeon General's Report on Nutrition and Health. DHHS (PHS) 88-50210, Washington, DC.

7. National Research Council, Committee on Diet and Health, Food and Nutrition Board. 1989. Diet and Health: Implications for Reducing Chronic Disease Risk. National Academy Press, Washington, DC.

8. Kennedy, ET, Ohls, J, Carlson, S, and Fleming, K 1995. The Healthy Eating Index: Design and applications. J Am Diet Assoc 95(10):1103-1108.

9. Douglass JS, Waylett DK, Doyle E, Fleming KH, Heimbach JT. Healthy Eating Index Scores for the Third National Health and Nutrition Examination Survey. TAS-ENVIRON Final Report prepared for the National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. April 1999.

10. Briefel RR, Sempos CT, McDowell MA, Chien S, Alaimo K. Dietary methods research in the third National Health and Nutrition Examination Survey: under reporting of energy intake. Am J Clin Nutr 1997;65(suppl): 1203S-9S.

11. U.S. Department of Agriculture, Human Nutrition Information Service. 1992. The Food Guide Pyramid. Home and Garden Bulletin Number 252.

12. Variyam, JN, Blaylock, J, Smallwood, D, and Basiotis, PP 1998. USDA's Healthy Eating Index and Nutrition Information. U.S. Department of Agriculture, Economic Research Service. Technical Bulletin No. 1866. 19

13. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion. 1995. The Healthy Eating Index. CNPP-1.

14. U.S. Department of Agriculture, Agricultural Research Service. 1998, February. 1994-96 Continuing Survey of Food Intakes by Individuals and 1994-96 Diet and Health Knowledge Survey and related materials [CD-ROM].

15. National Research Council, Subcommittee on the Tenth Edition of the RDAs, Food and Nutrition Board. 1989. Recommended Dietary Allowances. (10 th ed.). National Academy Press, Washington, DC.

NHANES III Healthy Eating Index Data File Index

```
---------------------------------------------------------------------
                                    Variable
    Description                     Name            Positions
---------------------------------------------------------------------


    GENERAL INFORMATION

    Sequence Number ....................    SEQN            1-5
    Grains, number of servings.........    GRAINNS         6-11
    Fruits, number of servings.........    FRUITNS         12-17
    Vegetables, number of servings......   VEGETNS         18-23
    Dairy, number of servings...........   DAIRYNS         24-29
    Meats, number of servings...........   MEATNS          30-35
    Legumes, number of servings........    LEGUMENS        36-41
    Grains, HEI score (0-10) ...........   GRAINHEI        42-46
    Fruits, HEI score (0-10) ...........   FRUITHEI        47-51
    Vegetables, HEI score (0-10) .......   VEGETHEI        52-56
    Dairy, HEI score (0-10) ............   DAIRYHEI        57-61
    Meats, HEI score (0-10) ............   MEATHEI         62-66
    Fats, HEI score (0-10) .............   FATHEI          67-71
    Saturated fat, HEI score (0-10) ....   SFATHEI         72-76
    Cholesterol, HEI score (0-10)......    CHOLEHEI        77-81
    Sodium, HEI score (0-10)...........    SODIUHEI        82-86
    Variety, HEI score (0-10)..........    VARIEHEI        87-91
    Healthy Eating Index score (0-100) .   HEISCORE        92-96
```

```
                NHANES III Healthy Eating Index Data File
     ----------------------------------------------------------------------
         File name=HEI      Version 1.0                    N=26350
     ----------------------------------------------------------------------
                          GENERAL INFORMATION
     ----------------------------------------------------------------------
     Positions                       Item description
     SAS name        Counts          and code                    Notes
     ----------------------------------------------------------------------
         1-5                         Sequence Number
       SEQN         26350            00003-53617

         6-11                        Grains, number of servings
       GRAINNS      181              00.000
                    26167            00.013-68.198
                    2                888888  Blank but applicable

         12-17                       Fruits, number of servings
       FRUITNS      6731             00.000
                    19617            00.001-59.276
                    2                888888  Blank but applicable

         18-23                       Vegetables, number of servings
       VEGETNS      1757             00.000
                    24591            00.001-36.981
                    2                888888  Blank but applicable

         24-29                       Dairy, number of servings
       DAIRYNS      645              00.000
                    25703            00.001-30.733
                    2                888888  Blank but applicable

         30-35                       Meats, number of servings
       MEATNS       231              00.000
                    26117            00.001-24.175
                    2                888888  Blank but applicable

         36-41                       Legumes, number of servings
       LEGUMENS     19700            00.000
                    6648             00.001-07.645
                    2                888888  Blank but applicable

         42-46                       Grains, HEI Score
       GRAINHEI     189              00000 None/never
                    26159            000.1-00010
                    2                88888 Blank but applicable

         47-51                       Fruits, HEI Score
       FRUITHEI     7209             00000 None/never
                    19139            000.1-00010
                    2                88888 Blank but applicable
```

```
                 NHANES III Healthy Eating Index Data File
-----------------------------------------------------------------------
Positions                        Item description
SAS name        Counts           and code                    Notes
-----------------------------------------------------------------------


    52-56                        Vegetables, HEI Score
  VEGETHEI      1875             00000 None/never
                24473            000.1-00010
                2                88888 Blank but applicable


    57-61                        Dairy, HEI Score
  DAIRYHEI      735              00000 None/never
                25613            000.1-00010
                2                88888 Blank but applicable


    62-66                        Meats, HEI Score
  MEATHEI       325              00000 None/never
                26023            000.1-00010
                2                88888 Blank but applicable


    67-71                        Fats, HEI Score
  FATHEI        2328             00000 None/never
                24020            000.1-00010
                2                88888 Blank but applicable


    72-76                        Saturated Fat, HEI Score
  SFATHEI       4726             00000 None/never
                21622            000.2-00010
                2                88888 Blank but applicable


    77-81                        Cholesterol, HEI Score
  CHOLEHEI      4558             00000 None/never
                21790            000.1-00010
                2                88888 Blank but applicable


    82-86                        Sodium, HEI Score
  SODIUHEI      4078             00000 None/never
                22270            000.1-00010
                2                88888 Blank but applicable


    87-91                        Variety, HEI Score
  VARIEHEI      1633             000.0
                1781             002.0
                2650             004.0
                3210             006.0
                3514             008.0
                13560            010.0
                2                88888 Blank but applicable


    92-96                        Healthy Eating Index Score
  HEISCORE      26348            00010-00100
                2                88888 Blank but applicable
```

NOTES

SEQN: Sequence Number

GRAINSNS: Grains, number of servings

This represents the number of servings consumed from the Food Guide
Pyramid grain group. The group includes bread, cereal, rice, and
pasta. Grain recommendations range from 6-11 servings depending on
caloric need. In turn, caloric need is dependent on age, sex, size,
and level of activity. See Table 1 for recommended number of USDA
Food Guide Pyramid servings of grains per day by sex and age.

FRUITNS: Fruits, number of servings

This represents the number of servings consumed from the Food Guide
Pyramid fruit group. Fruit recommendations range from 2-3 servings
depending on caloric need. In turn, caloric need is dependent on
age, sex, size, and level of activity. See Table 1 for recommended
number of USDA Food Guide Pyramid servings of fruits and fruit juices
per day by sex and age.

VEGETANS: Vegetables, number of servings

This represents the number of servings consumed from the Food Guide
Pyramid vegetable group. Vegetable recommendations range from 3-5
servings depending on caloric need. In turn, caloric need is
dependent on age, sex, size, and level of activity. See Table 1 for
recommended number of USDA Food Guide Pyramid servings of vegetables
and vegetable juice per day by sex and age.

DAIRYNS: Dairy, number of servings

This represents the number of servings consumed from the Food Guide
Pyramid dairy group. The group includes milk, yogurt, and cheese.
Dairy recommendations range from 2-3 servings depending on caloric
need. In turn, caloric need is dependent on age, sex, size, and
level of activity. See Table 1 for recommended number of USDA Food
Guide Pyramid servings of dairy servings per day by sex and age.

MEATNS: Meats, number of servings

This represents the number of servings consumed from the Food Guide
Pyramid meat group. The group includes meat, poultry, fish, dried
beans, eggs, and nuts. Meat food group recommendations range from
2-3 servings depending on caloric need. In turn, caloric need is
dependent on age, sex, size, and level of activity. See Table 1 for
recommended number of USDA Food Guide Pyramid servings of meat,
poultry, fish, nuts, dried peas and beans per day by sex and age.

LEGUMENS: Legumes, number of servings

This represents the number of servings of legumes consumed. Legumes
include dried peas and beans. The Food Guide Pyramid indicates that
legumes are to be considered contributors to meat group intake, but
it was decided in work on USDA HEIs that any legume consumption not

"needed" for achieving recommended meat group servings could be
applied toward vegetable group servings. Legume consumption by an
individual therefore was counted either as meat group consumption or
vegetable group consumption, depending on the amount of meat consumed
from other sources.

GRAINHEI: Grains, HEI score

This is a score between 0 and 10 (with 10 being the best score)
indicating to what extent a person's diet conforms to the Food Guide
Pyramid serving recommendations from the grain group.

FRUITHEI: Fruits, HEI score

This is a score between 0 and 10 (with 10 being the best score)
indicating to what extent a person's diet conforms to the Food Guide
Pyramid serving recommendations from the fruit group.

VEGETHEI: Vegetables, HEI score

This is a score between 0 and 10 (with 10 being the best score)
indicating to what extent a person's diet conforms to the Food Guide
Pyramid serving recommendations from the vegetable group.

DAIRYHEI: Dairy, HEI score

This is a score between 0 and 10 (with 10 being the best score)
indicating to what extent a person's diet conforms to the Food Guide
Pyramid serving recommendations from the dairy group.

MEATHEI: Meats, HEI score

This is a score between 0 and 10 (with 10 being the best score)
indicating to what extent a person's diet conforms to the Food Guide
Pyramid serving recommendations from the meat group.

FATHEI: Fats, HEI score

This component measures total fat consumption as a percentage of
total food energy intake (or calories). Fat intakes less than or
equal to 30 percent of the total calories were assigned a score of 10
points. The score declined to zero when the proportion of fat to
total calories reached 45 percent. Intakes between 30 percent and 45
percent were scored proportionately (i.e., intake of 31.5% calories
from fat yielded a score of 9, a score of 8 was received for
consuming 33% calories from fat, etc).

SFATHEI: Saturated fat, HEI score

This component measures saturated fat consumption as a percentage of
total food energy intake (or calories). A score of 10 points was
assigned to saturated fat intakes at less than 10 percent of total
calories. Zero points were assigned when the saturated fat intake
reached a level of 15 percent of the total calories. Scores between
the two cutoff values were calculated proportionately (i.e., people
who consumed 14% of total calories from saturated fat received a
score of 2, those who consumed 12.5% of total calories from saturated

fat received a score of 5, etc.).

CHOLEHEI: Cholesterol, HEI score

This component measures total cholesterol intake and was based on milligrams consumed. A maximum point value for cholesterol was assigned when intake was at a level of 300 milligrams or less. Zero points were assigned when intake reached a level of 450 milligrams or more. Values between the two cutoff points were scored proportionately (i.e., people who consumed 420 milligrams of cholesterol obtained a score of 2, those who consumed 390 milligrams obtained a score of 4, etc.).

SODIUHEI: Sodium, HEI score

This component measures total sodium intake and was based on milligrams consumed. A maximum score for sodium was assigned at an intake level of 2,400 milligrams or less. Zero points were assigned at a level of 4,800 milligrams or more. Scores between the two levels of intake were scored proportionately (i.e., people who consumed 3,600 milligrams of sodium obtained a score of 5, those who consumed 3120 milligrams of sodium obtained a score of 7, etc.).

VARIEHEI: Variety, HEI score

Dietary variety was assessed by totaling the number of "different" foods eaten by an individual in amounts sufficient to contribute at least one-half of a serving in a food group. Similar foods were grouped together and counted only once in measuring variety Food mixtures were broken down into their component ingredients and assigned to the relevant food groups. Index scores for variety were calculated in a manner analogous to the method used for the other Index components. A maximum score of 10 was given if 8 or more different food items were consumed over a 1-day period. A score of zero was given if three or less different items were eaten. Intermediate intakes were calculated proportionately (i.e., a score of 2 was obtained for consuming 4 different items, a score of 4 was assigned to those who consumed 6 different items, etc.).

HEISCORE: Healthy Eating Index score

An index that provides a measure of overall quality of an individual's diet. It provides a picture of foods people are eating, the amount of variety in the diet and compliance with specific Dietary Guidelines recommendations. A score on the Index reflects the sum total of ten diet components (consumption of foods from the grain, fruit, vegetables, dairy and meat food groups; intake of dietary fats, saturated fats, cholesterol, and sodium; and a variety score). For each of the five food group components of the Index, individuals who consumed the recommended number of servings received a maximum score of 10. A score of zero was assigned to any food group where no items from that category were consumed. Intermediate scores were calculated proportionately to the number of servings consumed. For example, if the recommended level of servings was eight and an individual consumed four servings, the component score for the individual was 5 points. A score of 7.5 points was assigned if six servings were eaten. A maximum score of 100 was possible for

the Healthy Eating Index.