

Calculation of mean Healthy Eating Index-2010 component and total scores and corresponding standard errors and confidence intervals for a population, subpopulation, or group

This file describes the methodology implemented in the SAS program, HEI2010_NHANES0708_MC_PopulationScore.sas. The program demonstrates how to calculate mean Healthy Eating Index-2010 (HEI-2010) component and total scores and corresponding standard errors and confidence intervals for the U.S. population, using 1-day dietary intakes reported by participants in the National Health and Nutrition Examination Survey conducted in 2007-2008 (NHANES 07-08). Output includes mean HEI-2010 component and total scores, along with their standard errors and confidence intervals. The principles underlying this approach are described in Freedman et al., J Nutr, 2008, 138(9):1725-9 ([see abstract in PubMed](#)). This general approach should be used to calculate HEI-2010 component and total scores for any population, subpopulation, or group. Because NHANES dietary data are weighted to equalize the days of the week and are collected year-round, these mean scores can be considered to be estimated scores of the population's long-term or "usual" intake.

This code can be adapted to calculate HEI-2010 scores for other cycles of NHANES or for other data sources that use a complex sampling strategy. This code is an update to that previously available for calculating HEI-2005 component and total scores and standard errors. In contrast to the previous code, this version does not require the use of SUDAAN; rather, SAS survey procedures are used to account for the survey design. This version also does not involve the use of an Excel spreadsheet to calculate confidence intervals for HEI-2010 component and total scores as the previous version did.

Because the HEI-2010 is a multi-dimensional construct involving 12 densities (amounts of food per 1,000 calories and ratios of fatty acids), a simple method for estimating standard errors is not available. In this code, a Monte Carlo simulation step is included for the calculation of standard errors; this step simulates the densities for 10,000 samples to allow stable estimation of the standard errors.

Note: Updated code for HEI-2005 component and total scores is also available.

Note: A version of the HEI-2010 code is also available for non-survey datasets. It uses the analysis of data from the Automated Self-Administered 24-hour Recall (ASA24) system as an example.

This SAS program includes four main sections:

- (I) Calculations at the individual participant level to obtain variables needed to calculate HEI-2010 scores.
- (II) Calculation of weighted means and a variance-covariance matrix and generation of a Monte Carlo dataset, enabling standard errors to be calculated.
- (III) Allocation of beans and peas (legumes) and application of the HEI-2010 scoring algorithm.

- (IV) Calculation of HEI-2010 component and total scores and their standard errors and confidence intervals.

This code was written by Lisa Kahle and Dennis Buckman of Information Management Services, Inc., with assistance from Kevin Dodd of the National Cancer Institute.

Required datasets:

NHANES 07-08 dietary intake and demographics data

MyPyramid Equivalents Database (MPED), version 2.0

CNPP MyPyramid Equivalents Databases for Whole Fruit and Fruit Juice for NHANES 2003-04

CNPP Addendum to the MyPyramid Equivalents Database (MPED), version 2.0B

Required macros:

hei2010.beanspeas.allocation.macro.sas

hei2010.score.macro.sas

Section I. Calculations at the individual participant level to obtain variables needed to calculate HEI-2010 scores.

The HEI-2010 has 12 components:

Total Fruit;

Whole Fruit;

Total Vegetables;

Greens and Beans;

Whole Grains;

Dairy;

Total Protein Foods;

Seafood and Plant Proteins;

Fatty Acids;

Refined Grains;

Sodium; and

Empty Calories.

The MyPyramid Equivalents Database (MPED) and NHANES 07-08 datasets are used to create five of the components: Total Fruit; Whole Grains; Fatty Acids; Refined Grains; and Sodium. The CNPP MyPyramid Equivalents Database for Whole Fruit and Fruit Juice is used in the calculation of the Whole Fruit component. Additional steps are necessary to create the remaining six components: Total Vegetables; Greens and Beans; Dairy; Total Protein Foods; Seafood and Plant Proteins, and Empty Calories.

Step 1. Locate the required datasets and variables and make the necessary exclusions.

(a) MyPyramid Equivalents Database (MPED)

MyPyramid Equivalents Database for USDA Survey Food Codes, Version 2.0 (MyPyrEquivDB_v2) can be extracted from the Agricultural Research Service web site (<http://www.ars.usda.gov/Services/docs.htm?docid=17558>). All foods in the database are uniquely coded with an 8-digit USDA food code.

This database contains the number of equivalents per 100 grams of food for the following food groups, which are used to calculate component and total scores for the HEI-2010:

Total Fruit (F_TOTAL) in cup equivalents
Total Vegetables (V_TOTAL) in cup equivalents
Dark Green Vegetables (V_DRKGR) in cup equivalents
Legumes (LEGUMES) in cup equivalents
Whole Grains (G_WHL) in ounce equivalents
Total Milk (D_TOTAL) in cup equivalents
Meat, Poultry, Fish (M_MPF) in ounce equivalents
Eggs (M_EGG) in ounce equivalents
Nut and Seeds (M_NUTSD) in ounce equivalents
Soybean products (M_SOY) in ounce equivalents
Fish and shellfish high in n-3 fatty acids (FISH_HI) in ounce equivalents
Fish and shellfish low in n-3 (FISH_LO) in ounce equivalents
Non-whole/refined Grains (G_NWHL) in ounce equivalents
Discretionary solid fat (DISCFAT_SOL) in grams
Added sugars (ADD_SUG) in teaspoon equivalents

Calculation note for soy beverages (Dairy and Total Protein Foods components): Soy beverages are counted as part of the Dairy component of the HEI-2010. This differs from the MPED, which groups them with other soy products in the Meats and Beans group. The reassignment process is completed in this step. Soy beverages (food codes 11310000, 11320000, 11321000, and 11330000) are converted from soybean products (M_SOY) in ounce equivalents to milk (D_TOTAL) in cup equivalents, based on the weight in grams of 1 cup.

(b) CNPP MyPyramid Equivalents Database for Whole Fruit and Fruit Juice for NHANES 2003-2004 (cnppmyp_v1nhanes0304_wjfrt-sas.zip):

Calculation note for Whole Fruit: This database was created by CNPP from the MyPyramid Equivalents Database for USDA Survey Food Codes, Version 2.0. A SAS-database can be downloaded from the CNPP web site (<http://www.cnpp.usda.gov/HealthyEatingIndexSupportFiles0304.htm>). Foods containing fruit were assigned to either Whole Fruit (WHOLEFRT) or Fruit Juice (FRTJUICE) in cup equivalents per 100 grams of foods. In the case of foods that contain both whole fruit and fruit juice, e.g., fruit canned in fruit juice, the Total Fruit equivalents were assigned to either Whole Fruit or Fruit Juice, whichever was the majority ingredient according to its description or recipe in the Food and Nutrient Database for Dietary Studies (FNDDS), version 2.0, developed by the Agricultural Research Service.

(c) CNPP Addendum to the MyPyramid Equivalents Database (MPED), version 2.0B

This database contains 942 food codes that were reported in NHANES 2005-06 and 2007-08 data cycles for individuals age 2 years and older (excluding breastfed children) who had complete recalls. Combine this database with the MPED, Version 2.0, for complete food equivalent data to the above mentioned data cycles. Corrected data for five food codes found in the Addendum should be used in place of those found in the MPED Version 2.0. These food codes are 11710800, 11710801, 58106210, 58106220, and 58106230. The database can be downloaded from the CNPP website (<http://www.cnpp.usda.gov/OtherProjects.htm>). It includes Whole Fruit and Fruit Juice variables.

(d) NHANES 07-08 datasets for food intake (drxiff_e), nutrient intake (dr1tot_e), and demographics (demo_e)

NHANES 07-08 datasets for food item intakes (drxiff_e), total daily nutrients intake (dr1tot_e), and demographics (demo_e) can be downloaded from the Centers for Disease Control and Prevention, National Center for Health Statistics web site (http://www.nchs.gov/nhanes/search/nhanes01_02.aspx). The NHANES 07-08 food intake dataset, drxiff_e, includes the list of foods eaten by each individual on Day 1 in gram amounts (DR1IGRMS). The NHANES 07-08 nutrient intake dataset, drxtot_e, includes each individual's Day 1 dietary sample weight (WTDRD1) and total nutrient intakes for Day 1. The following information from this dataset is needed to determine intakes of saturated fat, sodium, calories from SoFAAS, and energy:

Dietary day one sample weight (WTDRD1)
Foodcode (DRDIFDCD)
Energy (DR1TKCAL) in kcal
Sodium (DRD1SODI) in grams
Alcohol (DR1IALCO) in grams
Monounsaturated fat (DR1TMFAT) in grams
Polyunsaturated fat (DR1TPFAT) in grams
Saturated fat (DR1TSFAT) in grams

Only participants with complete and reliable dietary data (DR1DRSTZ=1) and age 2 years and older are used to calculate HEI-2010 scores because the Dietary Guidelines for Americans were not designed for younger children. The NHANES 07-08 demographic dataset, demo_e, includes information about individuals, such as age (RIDAGEYR). Other variables needed from this dataset are related to the survey design:

Reliable dietary recall status (DR1DRSTZ)
Age (RIDAGEYR)
Data release number (SDDSRVYR)
Masked Variance Pseudo-PSU (SDMVPSU)
Masked Variance Pseudo-Stratum (SDMVSTRA)

Step 2. Combine the required datasets and make necessary adjustments at the level of individual foods.

MyPyramid Equivalents Database, version 2.0, CNPP MyPyramid Equivalents Database for Whole Fruit and Fruit Juice for NHANES 2003-04, and the CNPP Addendum 2.0B are first combined and further merged with NHANES 07-08 individual food intake data to calculate individuals' 1-day food-group intake, expressed in MyPyramid equivalents (i.e., cup-, ounce-, and teaspoon-equivalents).

Calculation note for Fatty Acids: Monounsaturated fatty acids and polyunsaturated fatty acids are summed together in Section I. The Fatty Acids component scores are determined by the ratio of unsaturated fatty acids to saturated fatty acids ((monounsaturated fatty acids + polyunsaturated fatty acids)/saturated fatty acids).

Calculation note for Empty Calories: An individual's intake of Empty Calories is calculated using the discretionary solid fats and added sugars data from MPED and the alcohol intake data from NHANES 07-08. Calories from alcohol are calculated using DR11ALCO. In the HEI-2010, calories from alcohol are considered to be empty calories, but only when alcohol is consumed beyond moderate amounts. The least restrictive of the two levels defined as moderate drinking in the Dietary Guidelines, 2 drinks per day (converted to 28 grams of ethanol), was used to set the threshold for counting alcohol as empty calories. A value of 2150 calories was used to energy-adjust the alcohol threshold, based on the estimated median energy intake of adults. Because 28 g ethanol/2150 calories equals 13 g ethanol/1000 calories, up to 13 g of ethanol/1000 calories are not counted towards Empty Calories.

Note for beans and peas (legumes): In a previous version of the code for calculating population HEI scores, legume allocation occurred as part of this step. In this updated approach, allocation of beans and peas (legumes) now occurs as part of Section III, after the Monte Carlo simulation step. This is intended to better reflect usual intake of beans and peas.

Step 3. Calculate NHANES 07-08 individual food and nutrient intakes for each individual.

In this step, computations are performed to create HEI-2010 components that are a combination of food groups from MPED. Note that beans and peas (legumes) are allocated to the Total Protein Foods and the Seafood and Plant Proteins components and/or to the Total Vegetables and the Greens and Beans components in Section III.

Section II. Calculation of weighted means and a variance-covariance matrix and generation of a Monte Carlo dataset, enabling standard errors to be calculated.

This section of the code performs computations necessary for the calculation of standard errors, accounting for the complex sampling design of NHANES 07-08. SAS survey procedures are used.

Step 1. Calculate the weighted means and the variance/covariance matrix for the dietary variables of interest.

Step 2. In this step, a Monte Carlo data set with 10,001 rows is generated using the weighted means and variance/covariance matrix from step 1.

Section III. Allocation of beans and peas (legumes) and application of the HEI-2010 scoring algorithm.

This section of the code calls two macros, `hei2010.beanspeas.allocation.macro.sas` and `hei2010.score.macro.sas`.

Step 1. Allocate beans and peas (legumes) using the Monte Carlo data set from Section II and the beans and peas allocation macro.

Intake of beans and peas counts toward meeting the standard for the Total Proteins Foods (and Seafood and Plants Proteins) components first. Once the Total Proteins Foods standard is met, any additional amount of beans and peas count toward the Total Vegetables and the Greens and Beans components. Beans and peas amounts in the MPED are in cup equivalents; therefore, the cup equivalents are first converted to ounce equivalents of protein food when they are counted for the Total Proteins Foods component, and are then converted back to cup equivalents when counted as vegetables. One-fourth cup of beans and peas is equal to 1 ounce equivalent of protein food. Thus, the number of cup equivalents of beans and peas is multiplied by 4 to convert to ounce equivalents of protein food.

Step 2. Apply the HEI-2010 scoring macro.

Using the Monte Carlo data set with the beans and peas allocated from step 1, the HEI-2010 scoring macro is called to calculate densities for each HEI-2010 component and then the scoring algorithm is applied.

Section IV. Calculation of HEI-2010 component and total scores and their standard errors and confidence intervals.

Step 1. This step uses univariate and means procedures to compute total and component scores and their standard errors and confidence intervals.

Steps 2-3. These steps prepare and display the results.

Step 4. This step provides an option to export the results into a CSV file that can be opened in Excel.