**ReadMe file for ASA24-2011 to ASA24-2014**

**HEI-2010 Population Ratio Method**

This ReadMe file describes the methodology implemented in the SAS program, HEI2010\_ASA24\_MC\_PopulationScore.sas. The principles underlying this approach are described in Freedman et al., J Nutr, 2008, 138(9):1725-9 ([see abstract in PubMed](http://www.ncbi.nlm.nih.gov/pubmed/18716176)). The provided SAS program, along with the required macros, calculate mean Healthy Eating Index-2010 (HEI-2010) component and total scores and corresponding standard errors and confidence intervals for a population, subpopulation, or group using 24-hour recall data from ASA24-2011, ASA24-Kids-2013, ASA24-2014 and ASA24-Kids-2014. The specific instructions on how to successfully run the HEI procedure are within the provided SAS program.

This code can be adapted to calculate HEI-2010 scores for other data sources that use complex sampling strategy.

Only participants aged 2 years and older should be included in calculations of HEI-2010 scores because the Dietary Guidelines for Americans were not designed for younger children.

**NOTE:** Another version of this code is available for complex survey datasets, using the analysis of data from the 2007-08 National Health and Nutrition Examination Survey (NHANES) as an example. Also, additional code that calculates HEI-2010 component and total scores at the individual level, for each day or multiple days of 24-hour recalls are available on the [NCI Healthy Eating Index website](https://epi.grants.cancer.gov/hei/sas-code.html). Sample Analysis Files and Data Dictionaries can be downloaded from the ASA24 [Researcher website page](https://epi.grants.cancer.gov/asa24/researcher/sample.html). If any changes are made to the INFMYPHEI analysis file during data cleaning, please recreate an updated TNMYPHEI file prior to running this code. The files should be in CSV format.

This code is an update to that previously available for calculating HEI-2005 component and total scores and standard errors using NHANES data. In contrast to the previous code, this version does not require the use of SUDAAN and does not involve the use of an Excel spreadsheet to calculate confidence intervals for HEI-2010 components and total scores as the previous version did.

Because the HEI-2010 is a multi-dimensional construct involving 12 densities (amounts of food groups 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.

This SAS program includes four main sections:

1. Calculations at the individual participant level to obtain variables needed to calculate HEI-2010 scores.
2. Calculation of weighted means and a variance-covariance matrix and generation of a Monte Carlo dataset, enabling standard errors to be calculated.
3. Allocation of beans and peas (legumes) to Total Proteins and Seafood and Plant Proteins and/or Total Vegetables and Greens and Beans and application of the HEI-2010 scoring algorithm.
4. Calculation of mean HEI-2010 component and total scores and their standard errors and confidence intervals.

***Required datasets:***

* INFMYPHEI output from ASA24
* TNMYPHEI output from ASA24

***Required macros:*** *available on the* [*NCI Healthy Eating Index website*](https://epi.grants.cancer.gov/hei/sas-code.html)*.*

* [hei2010.beanspeas.allocation.macro.sas](https://epi.grants.cancer.gov/hei/sas-code.html)
* [hei2010.score.macro.sas](https://epi.grants.cancer.gov/hei/sas-code.html)

The 12 components of the HEI-2010 calculated by the macros are:

* Total Fruit;
* Whole Fruit;
* Total Vegetables;
* Greens and Beans;
* Whole Grains;
* Dairy;
* Total Protein Foods;
* Seafood and Plant Proteins;
* Fatty Acids;
* Refined Grains;
* Sodium;
* Empty Calories.

Some of these components come directly from the output of the ASA24-2011-14, but others are created as part of the SAS program.

The ASA24 TNMYPHEI data provide the following components used directly in the calculation of HEI-2010: Total Fruit; Whole Fruit; Whole Grains; Refined Grains; and Sodium.

*Additional steps are necessary to create the remaining components that are a combination of variables: Total Vegetables; Greens and Beans; Dairy; Total Protein; Seafood and Plant Proteins; Fatty Acids and Empty Calories.*

**The SAS program carries out 8 steps:**

1. **Reads in ASA24 INFMYPHEI and TNMYPHEI data.**

The data contain values from the Food and Nutrient Database for Dietary Surveys (FNDDS), version 4.1, and the My Pyramid Equivalents Database for USDA Survey Food Codes (MPED), version 2.0.

1. **Moves soy beverages out of Soybean Products (M\_SOY) and into Dairy (D\_TOTAL). This uses the ASA24 INFMYPHEI output file. The adjusted totals are generated for each day of dietary intake and saved in a temporary SAS dataset.**

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.

1. **Merges the TNMYPHEI file with the SAS dataset created in Steps 2, creating a file with adjusted variables for the MPED variables of Total Milk (D\_TOTAL\_ADJ) and Soybean Products (M\_SOY\_ADJ).**
2. **Creates additional required variables: MONOPOLY, ALLMEAT, SEAPLANT, ADDSUGC, SOLDFATC, EXALCCAL and EMPTYCAL10.**

Calculation note for MONOPOLY: MONOPOLY sums up monounsaturated and polyunsaturated fatty acids (MONOPOLY = MFAT + PFAT). To estimate the fatty acid ratio of unsaturated fatty acids to saturated fatty acids, the scoring macro divides this summed value is then divided by saturated fatty acids (MONOPOLY/SFAT (g)).

Calculation note for ALLMEAT and SEAPLANT:ALLMEAT sums together all animal and plant proteins, including meat, poultry, fish, eggs, nuts, seeds, and the soy variable (ALLMEAT = M\_MPF (oz) + M\_EGG (oz) + M\_NUTSD (oz) +M\_SOY\_ADJ (oz)); while SEAPLANT sums together all fish and plant proteins, including fish, soy, nuts, and seeds (SEAPLANT = M\_FISH\_HI (oz)+ M\_FISH\_LOW (oz) + M\_SOY\_ADJ (oz) + M\_NUTSD (oz)).

Calculation note for Empty Calories (EMPTYCAL10): EMPTYCAL10 sumps up three variables, ADDSUGC, SOLFATC, and EXALCCAL described below, to calculate Empty calories. The Empty Calories component of the HEI is calculated as a percentage of calories, so all of the variables that comprise empty calories need to be converted to units of calories.

ADDSUGC: Teaspoons of added sugar (ADD\_ SUGARS) are converted to calories using the conversion factor 1tsp=16kcal.

SOLDFATC: Grams of solid fat (SOLID\_FATS) are converted to calories using the conversion factor 1g=9kcal.

EXALCCAL: In the HEI-2010, energy from alcohol is 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 grams ethanol/2150 calories equals 13 grams ethanol/1000 calories, only amounts greater than 13 grams of ethanol/1000 calories are counted towards Empty Calories.

Note for beans and peas (legumes):In a previous version of the code for calculating population scores, legume allocation occurred as part of this step. In this updated approach, allocation of beans and peas now occurs as part of Step 7, after the Monte Carlo simulation step. This is intended to better reflect usual intake of beans and peas.

# Calculates weighted means and a variance-covariance matrix and generates a Monte Carlo dataset, enabling standard errors to be calculated.

This section of the code performs computations necessary for the calculation of standard errors.

1. Calculates the weighted means and the variance/covariance matrix for the dietary variables of interest.
2. In this step, a Monte Carlo data set with 10,001 rows is generated using the means and variance/covariance matrix from step i.

# Runs the macro to properly allocate intakes of LEGUMES in cup equivalents to Total Proteins and Seafood and Plant Proteins and/or Total Vegetables and Greens and Beans.

This section of the code calls the macro, [hei2010.beanspeas.allocation.macro.sas](https://epi.grants.cancer.gov/hei/sas-code.html).

This step results in 4 additional variables, named LEGUME\_ADDED\_V\_TOTAL, LEGUME\_ADDED\_BEANGRN, LEGUME\_ADDED\_ALLMEAT, and LEGUME\_ADDED\_SEAPLANT are all used in step 8.

Calculation notes for Legumes: Intake of legumes counts toward meeting the standard for the Total Protein Foods component first. Once the Total Protein Foods standard is met, regardless of whether the Seafood and Plant Protein standard is met, any additional amounts of legumes count toward Total Vegetables and Greens and Beans components.

**NOTE:** Units of Beans and peas in MPED are in cup equivalents; therefore, the cup equivalents are first converted to ounce equivalents of meat when they are counted for the Total Proteins and Seafood and Plant Proteins components, and are then converted back to cup equivalents when counted as vegetables.

Calculation note regarding conversion of cup equivalents to ounce equivalents:One-fourth cup of legumes is equal to 1-ounce equivalent of meat. Thus, the number of cup equivalents of legumes is multiplied by 4 to convert to ounce equivalents of meat.

# Runs the HEI-2010 scoring macro which calculates intake density amounts and calculates HEI-2010 component and total scores and their standard errors and confidence intervals.

This section of the code calls the macro, [hei2010.score.macro.sas](https://epi.grants.cancer.gov/hei/sas-code.html)

1. This step uses the Monte Carlo dataset with the beans and peas allocated from step 4 and calls the HEI-2010 scoring macro which calculates intake density amounts and HEI scores.
2. Univariate and means procedures are used to compute one HEI-2010 total score and one set of HEI-2010 component scores and their standard errors and confidence intervals for the group, subgroup, or population.

Below are the HEI-2010 Component Scoring Standards. For more information on HEI components, see [Comparing Versions of the HEI](https://epi.grants.cancer.gov/hei/comparing.html) on the NCI website.

**HEI–2010**[**1**](http://epi.grants.cancer.gov/hei/developing.html#f1) **Components & Scoring Standards**

|  |  |  |  |
| --- | --- | --- | --- |
| **Component** | **Max points** | **Standard for maximum score** | **Standard for minimum score of zero** |
| **Adequacy:** |  |  |  |
| **Total Fruit**[**2**](http://epi.grants.cancer.gov/hei/developing.html#f2) | 5 | ≥0.8 cup equiv. per 1,000 kcal | No Fruit |
| **Whole Fruit**[**3**](http://epi.grants.cancer.gov/hei/developing.html#f3) | 5 | ≥0.4 cup equiv. per 1,000 kcal | No Whole Fruit |
| **Total Vegetables**[**4**](http://epi.grants.cancer.gov/hei/developing.html#f4) | 5 | ≥1.1 cup equiv. per 1,000 kcal | No Vegetables |
| **Greens and Beans**[**4**](http://epi.grants.cancer.gov/hei/developing.html#f4) | 5 | ≥0.2 cup equiv. per 1,000 kcal | No Greens and Beans |
| **Whole Grains** | 10 | ≥1.5 oz equiv. per 1,000 kcal | No Whole Grains |
| **Dairy**[**5**](http://epi.grants.cancer.gov/hei/developing.html#f5) | 10 | ≥1.3 cup equiv. per 1,000 kcal | No Dairy |
| **Total Protein Foods**[**6**](http://epi.grants.cancer.gov/hei/developing.html#f6) | 5 | ≥2.5 oz equiv. per 1,000 kcal | No Protein Foods |
| **Seafood and Plant Proteins**[**6**](http://epi.grants.cancer.gov/hei/developing.html#f6)**,**[**7**](http://epi.grants.cancer.gov/hei/developing.html#f7) | 5 | ≥0.8 oz equiv. per 1,000 kcal | No Seafood or Plant Proteins |
| **Fatty Acids**[**8**](http://epi.grants.cancer.gov/hei/developing.html#f8) | 10 | (PUFAs + MUFAs)/SFAs ≥2.5 | (PUFAs + MUFAs)/SFAs ≤1.2 |
| **Moderation:** |  |  |  |
| **Refined Grains** | 10 | ≤1.8 oz equiv. per 1,000 kcal | ≥4.3 oz equiv. per 1,000 kcal |
| **Sodium** | 10 | ≤1.1 gram per 1,000 kcal | ≥2.0 grams per 1,000 kcal |
| **Empty Calories**[**9**](http://epi.grants.cancer.gov/hei/developing.html#f9) | 20 | ≤19% of energy | ≥50% of energy |

**1:** Intakes between the minimum and maximum standards are scored proportionately.

**2:** Includes fruit juice.

**3:** Includes all forms except juice.

**4:** Includes any beans and peas not counted as Total Protein Foods.

**5:** Includes all milk products, such as fluid milk, yogurt, and cheese, and fortified soy beverages.

**6:** Beans and peas are included here (and not with vegetables) when the Total Protein Foods standard is otherwise not met.

**7:** Includes seafood, nuts, seeds, soy products (other than beverages) as well as beans and peas counted as Total Protein Foods.

**8:** Ratio of poly- and monounsaturated fatty acids to saturated fatty acids.

**9:** Calories from solid fats, alcohol, and added sugars; threshold for counting alcohol is >13 grams/1000 kcal.

1. **Displays and saves the** **results** **in the specified output folder** **(see notes in the provided SAS program).**
   1. The program saves total score and set of component scores for the population/group of interest, together with minimum and maximum values, standard errors and confidence intervals. An option is provided to export the results into a CSV file that can be opened in Excel.
   2. Print

This step is included as a data check. The min and max can be compared to the bounds of HEI-2010 scores – if any scores <0 or >100, this is a red flag.

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. Please send any comments and questions regarding this code to [RFAB@mail.nih.gov](mailto:RFAB@mail.nih.gov).