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

**HEI-2005 Population Ratio Method**

This ReadMe file describes the methodology implemented in the SAS program, HEI2005\_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-2005 (HEI-2005) 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-2005 scores for other data sources that use complex sampling strategy.

Only participants aged 2 years and older should be used to calculate HEI-2005 scores because the Dietary Guidelines for Americans were not designed for younger children.

**NOTE:** A version of this code is available for complex survey datasets, using the analysis of data from the 2001-02 National Health and Nutrition Examination Survey (NHANES) as an example. Also, additional code that calculates HEI-2005 component and total scores at the individual level, for each day or multiple days of 24-hour recalls or food records 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-2005 components and total scores as the previous version did.

Because the HEI-2005 is a multi-dimensional construct involving 12 densities, 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-2005 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 legumes to Meat and Beans and/or Total Vegetables and Dark Green and Orange Vegetables and Legumes and application of the HEI-2005 scoring algorithm.
4. Calculation of mean HEI-2005 component and total scores and their 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)*.*

* [hei2005.beanspeas.allocation.macro.sas](https://epi.grants.cancer.gov/hei/sas-code.html)
* [hei2005.score.macro.sas](https://epi.grants.cancer.gov/hei/sas-code.html)

The 12 components of the HEI-2005 calculated by the macros are

* Total Fruit;
* Whole Fruit;
* Total Vegetables;
* Dark Green and Orange Vegetables and Legumes;
* Total Grains;
* Whole Grains;
* Milk;
* Meat and Beans;
* Oils;
* Saturated Fat;
* Sodium;
* Calories from Solid Fats, Alcoholic beverages, and Added Sugars (SoFAAS).

Some of these components come directly from the output of the ASA24-2011 to ASA24-2014, 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-2005: Total Fruit; Total Grains; Whole Grains; Oils; Saturated Fat; and Sodium.

*Additional steps are necessary to create the remaining components that are a combination of variables: Total Vegetables; Dark Green and Orange Vegetables and Legumes; Whole Fruit; Dairy; Meat and Beans; and Calories from SoFAAS.*

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

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

The INFMYPHEI and TNMYPHEI files contain the values from the Food and Nutrient Database for Dietary Surveys (FNDDS), version 4.1, the MyPyramid Equivalents Database for USDA Survey Food Codes (MPED), version 2.0, and the CNPP MyPyramid Equivalents Database for Whole Fruit and Fruit Juice needed to calculate HEI-2005 component and total scores.

Calculation note for Whole Fruit:The MyPyramid Equivalents Database for Whole Fruit and Fruit Juice database was created by CNPP. 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 FNDDS.

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 (Milk and Meat and Beans components): Soy beverages are counted as part of the Milk component of the HEI-2005. 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 from Steps 2 and 3, 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: ALLMEAT, V\_DOL, ALC, CARB, SoFAAS.**

Calculation note for ALLMEAT: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)).

Calculation note for V\_DOL:V\_DOL sums together Dark Green and Orange Vegetables (V\_DOL = V\_ORANGE (cups) + V\_DRKGR (cups)).

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

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

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

EXALCCAL: Alcoholic beverage intake is defined as beer, wine, or distilled spirits consumed as a beverage. To calculate the calories from alcoholic beverages, beer, wine, and distilled spirits reported or coded separately are used. Cooking wine and alcoholic beverages otherwise used as ingredients are excluded. Calories from alcoholic beverages are calculated from the amounts of ethanol and carbohydrate contained in the beverages. To prevent double-counting calories from added sugars, any calories from added sugars are subtracted from the calories from alcoholic beverages because they are accounted for in the Added Sugars part of the Calories from SoFAAS component.

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.

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 the intakes of LEGUMES in cup equivalents to either Total Protein Foods (ALLMEAT) or Total Vegetables and Dark Green and Orange Vegetables.

This section of the code calls the macro, [hei2005.beanspeas.allocation.macro.sas](https://epi.grants.cancer.gov/hei/sas-code.html).

This step results in 3 additional variables, named LEGUME\_ADDED\_V\_TOTAL, LEGUME\_ADDED\_V\_DOL, and LEGUME\_ADDED\_ALLMEAT are all used in step 7.

Calculation notes for Legumes: Intake of legumes counts toward meeting the standard for the Meat and Beans component first. Once the Meat and Beans standard is met, any additional amount of legumes counts toward the Total Vegetables and the Dark Green and Orange Vegetables and Legumes components.

**NOTE:** Legume amounts in the MPED are in cup equivalents; therefore, the cup equivalents are first converted to ounce equivalents of meat when they are counted for the Meat and Beans component, 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-2005 scoring macro which calculates intake density amounts and calculates HEI-2005 component and total scores and their standard errors and confidence intervals.

This section of the code calls the macro, [hei2005.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-2005 scoring macro which calculates intake density amounts and HEI scores.
2. Univariate and means procedures are used to compute one HEI-2005 total score and one set of HEI-2005 component scores and their standard errors and confidence intervals for the group, subgroup, or population.

Below are the HEI-2005 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–2005**[**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** | 5 | ≥0.8 cup equiv., per 1,000 kcal | No fruit |
| **Whole Fruit** | 5 | ≥0.4 cup equiv., per 1,000 kcal | No whole fruit |
| **Total Vegetables** | 5 | ≥1.1 cup equiv., per 1,000 kcal | No vegetables |
| **Dark Green, Orange Vegetables and Legumes2** | 5 | ≥0.4 cup equiv., per 1,000 kcal | No dark green/orange vegetables or legumes |
| **Total Grains** | 5 | ≥3.0 oz equiv., per 1,000 kcal | No grains |
| [**Whole Grains**](https://epi.grants.cancer.gov/hei/comparing.html#f3) | 5 | ≥1.5 oz equiv., per 1,000 kcal | No whole grains |
| **Milk3** | 10 | ≥1.3 c equiv., per 1,000 kcal | No dairy |
| **Meat & Beans** | 10 | ≥2.5 oz equiv., per 1,000 kcal | No meat or beans |
| **Oils4** | 10 | ≥12 grams, per 1,000 kcal | No Oil |
| **Moderation:** | | | |
| **Saturated Fats** | 10 | ≤7% of energy | ≥15% of energy |
| **Sodium** | 10 | ≤0.7 gram, per 1,000 kcal | ≥2.0 grams per 1,000 kcal |
| **Calories from Solid Fats, Alcoholic beverages, and Added Sugar (S0FAAS)** | 20 | ≤20% of energy, per 1,000 kcal | ≥ 50% of energy |

1: Intakes between the minimum and maximum levels are scored proportionately, except for Saturated Fat and Sodium

2: Legumes counted as vegetables only after Meat and Beans standard is met.

3: Includes all milk products, such as fluid milk, yogurt, and cheese, and soy beverages.

4: Includes nonhydrogenated vegetable oils and oils in fish, nuts, and seeds.

5: Saturated Fat and Sodium get a score of 8 for the intake levels that reflect the 2005 Dietary Guidelines, <10% of calories from saturated fat and 1.1 grams of sodium/1,000 kcal, respectively.

1. **Displays and saves the** **results** **in the specified output folder** **(see notes in the provided SAS program).**
2. 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.
3. Print

This step is included as a data check. The min and max can be compared to the bounds of HEI-2005 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).