# The Five-Factor Screener in the 2005 National Health Interview Survey Cancer Control Supplement

# (NHIS 2005)

November 2019

# THE FIVE FACTOR SCREENER: NHIS 2005

This information was originally published online on the Epidemiology and Genomic Research Program, National Cancer Institute's website as a reference for the Five-Factor Screener in the 2005 National Health Interview Survey Cancer Control Supplement (NHIS 2005). The information contained in this document informs analyses of data from this screener. This information is archived and provided for reference purposes only.

This publication may be viewed and downloaded from
https://epi.grants.cancer.gov/diet/screeners/files.

Suggested citation for information contained in this report: The Five-Factor Screener in the 2005 National Health Interview Survey Cancer Control Supplement. Epidemiology and Genomics Research Program. National Cancer Institute.
https://epi.grants.cancer.gov/diet/screeners/files. Updated November 20, 2019.

# Acknowledgements

The Epidemiology and Genomics Research Program, Risk Factors Assessment branch, National Cancer Institute (EGRP, RFAB, NCI) acknowledge the work of Dr. Frances E. Thompson, Dr. Amy F. Subar, Douglas Midthune, Dr. Kevin Dodd, Dr. Victor Kipnis, and Lisa Kahle, whose work formed the basis of this document.

The EGRP, NCI would also like to acknowledge the important role of Dr. Edwina Wambogo and Jennifer Lerman who turned the webpages associated with this screener into this document.

If you have a comment or question about this document, please contact the EGRP, RFAB, NCI at RFAB@mail.nih.gov.

# CONTENTS

# TABLES

# THE FIVE FACTOR SCREENER: NHIS 2005

# 1. Overview

The Five-Factor Screener may be useful to assess approximate intakes of fruits and vegetables, fiber, added sugar, calcium, and dairy foods. A single question about red meat is also asked. The screener asks respondents to report how frequently they consume foods in 18 categories. The screener also asks one question about the type of cereal consumed. No portion size questions are asked. This screener does not attempt to assess total diet.

The questions for the Five-Factor Screener were in-person interviewer-administered in the 2005 National Health Interview Survey (NHIS). PDF files of the cancer part of 2005 NHIS questionnaire, which includes the food questions (NAC.010 - NAC.138), are available in English and Spanish.
You can view or print the Five-Factor Screener from the National Cancer Institute's (NCI) Register of Validated Short Dietary Assessment Instruments.

The process of scoring the individual response data is described in Scoring Procedures. A description and guidelines for the appropriate uses of the screener-estimated dietary intakes is found in Uses of Screener Estimates. Validation data for the NHIS 2005 screener are presented in Validation Results. Finally, the various dietary intake variables are found in Computed Variables.

# 2. Scoring Procedures

## How Analytical Scoring Procedures Were Developed

Scoring procedures were developed to convert a respondent's screener responses to estimates of individual dietary intake for fruits and vegetables (servings), fiber (gm), added sugar (tsp), calcium (mg), and dairy (servings) using USDA's 1994-96 Continuing Survey of Food Intakes of Individuals (CSFII 94-96) dietary recall data. The following equations were estimated in the CSFII 94-96:

For **fiber** and **calcium**:

$$E\,(\text{Dietary Factor}) = b_0 + b_1(N_{FG1}P_1 + b_2N_{FG2}P_2 + \ldots + b_{20}N_{FG20}P_{20})$$

E (Dietary Factor) indicates the expected values for fiber and calcium and assumes a normal distribution. In the CSFII 94-96 dataset fiber was positively skewed and required a cube-root transformation to approximate normality. Calcium required a quarter-root transformation. $N_{FGk}$ is the usual number of times per day an individual consumed food group k; $P_k$ is the median portion size of group k; and k indexes the 20 food groups. These 20 food groups were formed to reflect the same food groups on the screener. We calculated weighted least-squares estimates of the regression coefficients $b_k$, k = 0, …, 20 on CSFII 94-96 adults aged 18 and above, stratifying by gender and excluding extreme exposure values. We first included all 20 food groups in the regression model. After examining the results,

we dropped food groups that failed to attain statistical significance at the $\alpha = 0.25$ level to form more parsimonious final models. In the fiber model, salad and soda were dropped for women. In the calcium model, all foods attained statistical significance. Because of the complex survey design, the analysis was performed using SUDAAN (RTI Inc., Research Triangle Park, NC).

For servings of **fruits and vegetables**:

$$E([\text{Fruits and Veg}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG9}P_9]^{1/2})$$

Servings of fruits and vegetables was square-root-transformed to approximate normality; $N_{FGk}$ is the usual number of times per day an individual consumed food group k; $P_k$ is the median portion size of group k; and k indexes the 9 fruit and vegetable food groups. We calculated weighted least-squares estimates of the regression coefficients $b_0$ and $b_1$ on the adults in the CSFII 94-96 sample, stratifying by gender and excluding extreme exposure values.

For teaspoons of **added sugar**:

$$E([\text{Added Sugar}]^{.33}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG4}P_4]^{.33})$$

Teaspoons of added sugar was cube-root-transformed to approximate normality; $N_{FGk}$ is the usual number of times per day an individual consumed food group k; $P_k$ is the median portion size of group k; and k indexes the 4 added sugar food groups. We calculated weighted least-squares estimates of the regression coefficients $b_0$ and $b_1$ on the adults in the CSFII 94-96 sample, stratifying by gender and excluding extreme exposure values.

For servings of **dairy**:

$$E([\text{Dairy}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2]^{1/2})$$

Servings of dairy was square-root-transformed to approximate normality; $N_{FGk}$ is the usual number of times per day an individual consumed food group k; $P_k$ is the median portion size of group k; and k indexes the 2 dairy food groups. We calculated weighted least-squares estimates of the regression coefficients $b_0$ and $b_1$ on the adults in the CSFII 94-96 sample, stratifying by gender and excluding extreme exposure values.

## Scoring Procedures

We performed the following steps with the NHIS 2005 Cancer Control Supplement dietary data to estimate the individual's intake of servings of fruits and vegetables, fiber, added sugar, calcium, and servings of dairy.

1. Estimation of $N_{FGk}$: All reported frequencies were standardized to a common unit of time by converting them to daily frequencies.

**Table 2- 1 Reported frequencies converted to daily frequencies**

| Frequency Response | NFGk: Daily Frequency |
|---|---|
| Never | 0 |
| 1-3 times per month | 0.067 |
| 1-2 times per week | 0.214 |
| 3-4 times per week | 0.5 |
| 5-6 times per week | 0.786 |
| 1 time per day | 1 |
| 2 times per day | 2 |
| 3 times per day | 3 |
| 4 times per day | 4 |
| 5 or more times per day | 5 |

2. Estimation of $P_k$: The median age- and gender-specific portion sizes for each food were estimated from CSFII 94-96. For fiber and calcium variables, the units were in grams (Table 2-2 & Table 2-3); for fruit and vegetable servings variables, the units were in Pyramid servings of fruits and vegetables (Table 2-4 & Table 2-5); for dairy servings, the unit was Pyramid servings of dairy (Table 2-6); and for teaspoons of added sugar, the unit was Pyramid teaspoons of added sugar (Table 2-7).

   For fruit and vegetables, a Pyramid serving was defined by the U.S. Department of Agriculture in the 1992 Dietary Guidelines Food Guide Pyramid as:

   - vegetables: 1 cup raw leafy, 1/2 cup of other vegetables, or 3/4 cup vegetable juice; and
   - fruit: 1 whole fruit, 1/2 cup of cut-up fruit, or 3/4 cup fruit juice.

   More recently, the 2005 Dietary Guidelines measure fruits and vegetables in cup equivalents. See MyPyramid for definitions of cup equivalents.

   Both metrics are provided for these 2005 data.

   For the milk group, a Pyramid serving is defined as:

   - 1 cup of milk or yogurt;
   - 1½ ounces of natural cheese; and
   - 2 ounces of processed cheese.

   For added sugar, a Pyramid serving is 1 teaspoon.

**Table 2- 2 Median Portion Size (Pk) in Grams per Mention by Age for Fiber and Calcium Analyses: Men**

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| Cooked Cereals ($P_1$) | 354.000000 | 219.630000 | 247.000000 | 247.000000 | 234.000000 | 239.000000 | 234.000000 |
| High-fiber Cereal ($P_2$) | 33.000000 | 33.000000 | 33.000000 | 33.000000 | 22.000000 | 22.000000 | 22.000000 |
| Moderate-fiber cereal ($P_3$) | 84.000000 | 66.000000 | 58.000000 | 64.125000 | 50.000000 | 47.000000 | 39.000000 |
| Low-fiber cereal ($P_4$) | 64.000000 | 54.000000 | 54.000000 | 40.625000 | 40.000000 | 30.375000 | 29.000000 |
| Milk ($P_5$) | 325.333333 | 268.400000 | 274.500000 | 244.000000 | 233.833333 | 206.000000 | 183.000000 |
| Regular Soda ($P_6$) | 453.866667 | 372.000000 | 372.000000 | 372.000000 | 370.200000 | 368.400000 | 368.000000 |
| Fruit Drinks ($P_7$) | 480.000000 | 376.500000 | 378.265000 | 372.000000 | 306.000000 | 248.000000 | 248.000000 |
| Fruit Juice ($P_8$) | 372.000000 | 311.250000 | 249.000000 | 249.000000 | 248.000000 | 186.750000 | 186.750000 |
| Fruit ($P_9$) | 131.750000 | 128.000000 | 123.200000 | 127.500000 | 122.000000 | 118.000000 | 114.250000 |
| Salad ($P_{10}$) | 30.000000 | 47.833333 | 42.500000 | 41.250000 | 42.500000 | 41.250000 | 44.666667 |
| French fries ($P_{11}$) | 112.500000 | 114.000000 | 100.000000 | 100.000000 | 85.500000 | 85.500000 | 97.000000 |
| Other potatoes ($P_{12}$) | 210.000000 | 193.000000 | 193.000000 | 161.000000 | 150.000000 | 127.000000 | 113.250000 |
| Dried beans ($P_{13}$) | 222.500000 | 188.095000 | 178.000000 | 189.750000 | 226.800000 | 199.333333 | 214.000000 |
| Other vegetables ($P_{14}$) | 61.250000 | 74.166667 | 75.625000 | 81.666667 | 76.000000 | 73.332500 | 70.000000 |
| Tomato sauce ($P_{15}$) | 63.000000 | 125.000000 | 125.000000 | 156.250000 | 122.500000 | 125.000000 | 125.000000 |

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| Salsa ($P_{16}$) | 62.250000 | 62.250000 | 49.275000 | 43.875000 | 16.000000 | 31.130000 | 31.130000 |
| Whole Grain Bread ($P_{17}$) | 56.000000 | 54.000000 | 52.000000 | 52.000000 | 51.000000 | 48.250000 | 48.000000 |
| Doughnuts, sweet rolls, muffins ($P_{18}$) | 71.000000 | 77.500000 | 72.800000 | 65.000000 | 63.000000 | 57.000000 | 57.000000 |
| Cookies, pie, cake, brownies ($P_{19}$) | 64.000000 | 66.000000 | 66.000000 | 73.733333 | 67.500000 | 64.000000 | 61.000000 |
| Cheese ($P_{20}$) | 35.440000 | 28.350000 | 30.470000 | 29.390000 | 28.350000 | 28.350000 | 28.350000 |

**Table 2- 3 Median Portion Size (Pk) in Grams per Mention by Age for Fiber and Calcium Analyses: Women**

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| Cooked Cereals ($P_1$) | 234.000000 | 234.000000 | 234.000000 | 234.000000 | 226.800000 | 234.000000 | 227.475000 |
| High-fiber Cereal ($P_2$) | 42.750000 | 42.750000 | 42.750000 | 42.750000 | 27.970000 | 27.970000 | 27.970000 |
| Moderate-fiber cereal ($P_3$) | 60.000000 | 57.000000 | 53.000000 | 49.500000 | 42.000000 | 39.083333 | 40.000000 |
| Low-fiber cereal ($P_4$) | 46.500000 | 37.500000 | 36.250000 | 33.000000 | 27.000000 | 26.000000 | 25.000000 |
| Milk ($P_5$) | 244.000000 | 244.000000 | 244.000000 | 214.250000 | 183.750000 | 183.000000 | 183.000000 |

# THE FIVE FACTOR SCREENER: NHIS 2005

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| Regular Soda ($P_6$) | 372.000000 | 372.000000 | 370.200000 | 368.400000 | 330.666667 | 366.000000 | 368.200000 |
| Fruit Drinks ($P_7$) | 360.000000 | 341.000000 | 250.000000 | 250.000000 | 248.000000 | 240.000000 | 221.200000 |
| Fruit Juice ($P_8$) | 280.125000 | 249.000000 | 248.800000 | 233.250000 | 189.755000 | 186.600000 | 186.700000 |
| Fruit ($P_9$) | 118.000000 | 118.000000 | 118.000000 | 118.000000 | 118.000000 | 112.427143 | 109.000000 |
| Salad ($P_{10}$) | 33.750000 | 32.083333 | 47.000000 | 55.000000 | 43.750000 | 34.333333 | 41.250000 |
| French fries ($P_{11}$) | 79.500000 | 70.000000 | 70.000000 | 70.000000 | 66.000000 | 70.000000 | 64.000000 |
| Other potatoes ($P_{12}$) | 122.000000 | 127.000000 | 119.000000 | 113.000000 | 105.000000 | 105.000000 | 105.000000 |
| Dried beans ($P_{13}$) | 132.750000 | 126.500000 | 126.500000 | 141.750000 | 130.550000 | 172.000000 | 178.000000 |
| Other vegetables ($P_{14}$) | 56.000000 | 62.043333 | 64.415000 | 64.920000 | 65.000000 | 67.375000 | 71.333333 |
| Tomato sauce ($P_{15}$) | 125.000000 | 113.400000 | 62.500000 | 125.000000 | 62.500000 | 62.500000 | 62.500000 |
| Salsa ($P_{16}$) | 32.000000 | 31.130000 | 36.565000 | 27.847500 | 31.130000 | 16.000000 | 16.000000 |
| Whole Grain Bread ($P_{17}$) | 50.000000 | 48.000000 | 47.500000 | 45.000000 | 45.000000 | 42.400000 | 34.000000 |
| Doughnuts, sweet rolls, muffins ($P_{18}$) | 67.333333 | 58.000000 | 57.000000 | 58.500000 | 57.000000 | 59.000000 | 47.000000 |
| Cookies, pie, cake, brownies ($P_{19}$) | 56.700000 | 50.000000 | 48.800000 | 55.200000 | 57.000000 | 48.675000 | 52.666667 |
| Cheese ($P_{20}$) | 28.250000 | 24.000000 | 24.000000 | 26.250000 | 28.350000 | 26.250000 | 28.350000 |

**Table 2- 4 Median Portion Size (Pk) in Pyramid Servings* per Mention by Gender and Age for Fruits and Vegetables Analyses**

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| **Men** | | | | | | | |
| 100% fruit juice ($P_1$) | 2.000000 | 1.667500 | 1.335000 | 1.335000 | 1.334000 | 1.001000 | 1.001000 |
| Fruit ($P_2$) | 1.301000 | 1.301000 | 1.229571 | 1.227333 | 1.168000 | 1.168000 | 1.052333 |
| Salad ($P_3$) | 0.545000 | 0.708000 | 0.754500 | 0.750000 | 0.833500 | 0.750000 | 0.822500 |
| Fried potatoes ($P_4$) | 2.000000 | 2.000000 | 1.773000 | 1.710000 | 1.400000 | 1.250000 | 1.250000 |
| Other potatoes ($P_5$) | 2.000000 | 2.000000 | 1.999000 | 1.999000 | 1.914000 | 1.544000 | 1.508000 |
| Dried beans ($P_6$) | 1.374000 | 1.047000 | 1.065000 | 1.227000 | 1.000000 | 1.000000 | 1.114000 |
| Other Vegetables ($P_7$) | 0.750000 | 0.906000 | 0.974500 | 1.000000 | 1.000000 | 0.880000 | 0.833333 |
| Tomato sauce ($P_8$) | 0.500000 | 0.541000 | 0.541000 | 0.812000 | 0.541000 | 0.541000 | 0.541000 |
| Salsa ($P_9$) | 0.533000 | 0.533000 | 0.421500 | 0.386500 | 0.137000 | 0.266000 | 0.266000 |
| **Women** | | | | | | | |
| 100% fruit juice ($P_1$) | 1.500500 | 1.334000 | 1.334000 | 1.251250 | 1.019500 | 1.000500 | 1.000500 |
| Fruit ($P_2$) | 1.168000 | 1.168000 | 1.168000 | 1.168000 | 1.150500 | 1.083833 | 1.000000 |
| Salad ($P_3$) | 0.613500 | 0.572500 | 0.833333 | 1.000000 | 0.795500 | 0.625000 | 0.750000 |
| Fried potatoes ($P_4$) | 1.481000 | 1.365500 | 1.272000 | 1.400000 | 1.000000 | 1.026000 | 1.000000 |
| Other white potatoes ($P_5$) | 1.544000 | 1.544000 | 1.528000 | 1.544000 | 1.499000 | 1.516000 | 1.272000 |

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| Dried beans ($P_6$) | 0.964000 | 0.684000 | 0.800000 | 0.687000 | 0.822000 | 0.807000 | 1.000000 |
| Other Vegetables ($P_7$) | 0.702200 | 0.779333 | 0.792500 | 0.788500 | 0.774000 | 0.833000 | 0.856750 |
| Tomato sauce ($P_8$) | 0.541000 | 0.541000 | 0.273000 | 0.541000 | 0.500000 | 0.500000 | 0.500000 |
| Salsa ($P_9$) | 0.274000 | 0.266000 | 0.322500 | 0.238250 | 0.266000 | 0.137000 | 0.137000 |

* Using 1992 Food Guide Pyramid definitions of servings.

**Table 2- 5 Median Portion Size (Pk) in Cup Servings** per Mention by Gender and Age for Fruits and Vegetables Analyses**

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| Men | | | | | | | |
| 100% fruit juice ($P_1$) | 1.499160 | 1.250580 | 1.000980 | 1.000980 | 1.000176 | 0.750735 | 0.750735 |
| Fruit ($P_2$) | 0.999580 | 0.933450 | 0.867300 | 0.867300 | 0.867300 | 0.774916 | 0.657060 |
| Salad ($P_3$) | 0.272700 | 0.353970 | 0.377235 | 0.374963 | 0.416640 | 0.375000 | 0.411323 |
| Fried potatoes ($P_4$) | 0.721125 | 0.727700 | 0.641000 | 0.641000 | 0.548055 | 0.480750 | 0.499980 |
| Other potatoes ($P_5$) | 1.000400 | 1.140030 | 0.999600 | 0.999600 | 0.999490 | 0.833175 | 0.754400 |
| Dried beans ($P_6$) | 0.717550 | 0.551540 | 0.566720 | 0.612360 | 0.500250 | 0.502285 | 0.575360 |
| Other Vegetables ($P_7$) | 0.387675 | 0.473920 | 0.499840 | 0.500240 | 0.499905 | 0.460585 | 0.416899 |

# THE FIVE FACTOR SCREENER: NHIS 2005

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| **Tomato sauce ($P_8$)** | 0.249900 | 0.271250 | 0.271250 | 0.406875 | 0.271250 | 0.271250 | 0.271250 |
| **Salsa ($P_9$)** | 0.266430 | 0.266430 | 0.210897 | 0.195683 | 0.068480 | 0.133236 | 0.133236 |
| **Women** | | | | | | | |
| **100% fruit juice ($P_1$)** | 1.124370 | 1.000960 | 1.000176 | 0.938130 | 0.764776 | 0.750728 | 0.750434 |
| **Fruit ($P_2$)** | 0.749235 | 0.867300 | 0.844838 | 0.789970 | 0.742350 | 0.712640 | 0.620475 |
| **Salad ($P_3$)** | 0.306788 | 0.286335 | 0.416625 | 0.499950 | 0.397688 | 0.312469 | 0.374963 |
| **Fried potatoes ($P_4$)** | 0.509595 | 0.455110 | 0.448700 | 0.448700 | 0.394856 | 0.444260 | 0.444260 |
| **Other white potatoes ($P_5$)** | 0.782020 | 0.876945 | 0.771260 | 0.771260 | 0.749700 | 0.771260 | 0.644235 |
| **Dried beans ($P_6$)** | 0.492150 | 0.341550 | 0.430530 | 0.345763 | 0.430685 | 0.430530 | 0.500400 |
| **Other Vegetables ($P_7$)** | 0.364468 | 0.395882 | 0.404303 | 0.408330 | 0.416913 | 0.436560 | 0.452214 |
| **Tomato sauce ($P_8$)** | 0.271250 | 0.271250 | 0.136710 | 0.271250 | 0.249900 | 0.249900 | 0.249900 |
| **Salsa ($P_9$)** | 0.136960 | 0.133236 | 0.163080 | 0.119187 | 0.133236 | 0.068480 | 0.068480 |

** Using 2005 MyPyramid definitions of cups of fruits and vegetables.

**Table 2- 6 Median Portion Size (Pk) in Pyramid Servings* per Mention by Gender and Age for Dairy Analyses**

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| **Men** | | | | | | | |
| **Milk ($P_1$)** | 1.250500 | 1.083000 | 1.100400 | 1.000000 | 0.916667 | 0.833333 | 0.750000 |
| **Cheese ($P_2$)** | 0.741000 | 0.641333 | 0.667000 | 0.600000 | 0.575000 | 0.499000 | 0.370000 |
| **Women** | | | | | | | |
| **Milk ($P_1$)** | 1.000000 | 1.000000 | 0.999000 | 0.874000 | 0.750000 | 0.718750 | 0.750000 |
| **Cheese ($P_2$)** | 0.517000 | 0.470000 | 0.494000 | 0.494000 | 0.470000 | 0.379000 | 0.494000 |

* Using 1992 Food Guide Pyramid definitions of servings.

**Table 2- 7 Median Portion Size (Pk) in Pyramid Servings* per Mention by Gender and Age for Added Sugar Analyses**

| Food Group | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-99 |
| **Men** | | | | | | | |
| **Soda ($P_1$)** | 11.835000 | 9.990000 | 9.947000 | 9.683000 | 9.683000 | 9.631000 | 9.605000 |
| **Fruit drinks ($P_2$)** | 9.627000 | 8.561000 | 8.985000 | 8.194000 | 6.815000 | 5.463000 | 5.307000 |
| **Doughnuts, sweet rolls, muffins ($P_3$)** | 4.3080000 | 4.196000 | 3.707000 | 3.095000 | 2.897000 | 2.837000 | 2.781500 |
| **Cookies, pie, cake, brownies ($P_4$)** | 5.189000 | 5.027000 | 4.845000 | 4.716000 | 4.730000 | 4.428000 | 3.968000 |
| **Women** | | | | | | | |
| **Soda ($P_1$)** | 9.815000 | 9.683000 | 9.683000 | 9.644000 | 8.443500 | 8.370000 | 9.683000 |
| **Fruit drinks ($P_2$)** | 7.997000 | 7.876000 | 6.418000 | 6.002000 | 6.418000 | 5.116000 | 5.116000 |
| **Doughnuts, sweet rolls, muffins ($P_3$)** | 2.966000 | 2.966000 | 2.797000 | 2.966000 | 2.498000 | 2.627000 | 2.627000 |
| **Cookies, pie, cake, brownies ($P_4$)** | 4.133000 | 3.650000 | 3.842000 | 3.719000 | 4.027500 | 3.571000 | 3.166500 |

* Using 1992 Food Guide Pyramid definitions of servings.

3. For fiber (grams), estimation of $b_0$ and $b_1$, the model is:

$$E([\text{Dietary Factor}]^{1/3}) = b_0 + (b_1[N_{FG1}P_1] + b_2[N_{FG2}P_2] + \ldots + b_{20}[N_{FG20}P_{20}])$$

For calcium (mg) estimation of $b_0$ and $b_1$, the model is:

$$E([\text{Dietary Factor}]^{1/4}) = b_0 + (b_1[N_{FG1}P_1] + b_2[N_{FG2}P_2] + \ldots + b_{20}[N_{FG20}P_{20}])$$

For fiber and calcium, estimation of $b_k$, k = 0, …, 20: the values for each parameter, for each gender, are in the following table:

**Table 2- 8 Estimated Regression Coefficients for Foods as Predictors of Fiber (gm) and Calcium (mg), by Gender**

| Parameter | Men | | Women | |
|---|---|---|---|---|
| | Fiber (cube root) | Calcium (quarter root) | Fiber (cube root) | Calcium (quarter root) |
| **Intercept ($b_0$)** | 2.015301 | 4.482732 | 1.838259 | 4.155762 |
| **Cooked Cereals ($b_1$)** | 0.000558 | 0.000318 | 0.000671 | 0.000484 |
| **High-fiber Cereal ($b_2$)** | 0.011463 | 0.006716 | 0.019873 | 0.006744 |
| **Moderate-fiber cereal ($b_3$)** | 0.003515 | -0.000355 | 0.004688 | 0.000074 |
| **Low-fiber cereal ($b_4$)** | -0.000425 | -0.002023 | 0.001493 | -0.001305 |
| **Milk ($b_5$)** | 0.000180 | 0.002204 | 0.000169 | 0.002580 |
| **Regular Soda ($b_6$)** | 0.000043 | 0.000089 | --- | 0.000095 |
| **Fruit Drinks ($b_7$)** | 0.000141 | 0.000105 | 0.000115 | 0.000326 |
| **Fruit juice ($b_8$)** | 0.000166 | 0.000123 | 0.000229 | 0.000195 |
| **Fruit ($b_9$)** | 0.000985 | 0.000170 | 0.001009 | 0.000264 |
| **Salad ($b_{10}$)** | -0.000447 | -0.000938 | --- | -0.000723 |
| **French fries ($b_{11}$)** | 0.001517 | 0.001159 | 0.001381 | 0.000414 |
| **Other potatoes ($b_{12}$)** | 0.000720 | 0.000349 | 0.000693 | 0.000489 |
| **Dried beans ($b_{13}$)** | 0.002156 | 0.000511 | 0.003217 | 0.001035 |
| **Other vegetables ($b_{14}$)** | 0.000899 | 0.000400 | 0.000925 | 0.000396 |

| Parameter | Men | | Women | |
|---|---|---|---|---|
| | Fiber (cube root) | Calcium (quarter root) | Fiber (cube root) | Calcium (quarter root) |
| **Tomato sauce ($b_{15}$)** | 0.001315 | 0.000948 | 0.001204 | 0.000287 |
| **Salsa ($b_{16}$)** | 0.003632 | 0.002179 | 0.003239 | 0.002679 |
| **Whole grain bread ($b_{17}$)** | 0.002927 | 0.000406 | 0.003401 | 0.000680 |
| **Doughnuts, sweet rolls, muffins ($b_{18}$)** | 0.000979 | 0.001198 | 0.001683 | 0.001873 |
| **Cookies, pie, cake, brownies ($b_{19}$)** | 0.001006 | 0.001771 | 0.001377 | 0.002451 |
| **Cheese ($b_{20}$)** | 0.001206 | 0.014186 | 0.000513 | 0.015442 |

4. For Pyramid servings of fruits and vegetables, estimation of $b_0$ and $b_1$, the model is:

$$E([\text{Dietary Factor}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG9}P_9]^{1/2})$$

For Pyramid servings of fruits and vegetables, including and excluding French fries, for each gender, the estimates of the parameters are:

**Table 2- 9 Estimated Regression Coefficients for Sum of Foods Predicting Servings of Total Fruits and Vegetables and Fruits and Vegetables Excluding French Fries, by Gender**

| Parameter | Men | Women |
|---|---|---|
| **Summary Variable with French fries** | | |
| Intercept ($b_0$) | 0.704319 | 0.658819 |
| $b_1$ | 0.835532 | 0.796243 |
| **Summary Variable excluding French fries** | | |
| Intercept ($b_0$) | 0.729653 | 0.639540 |
| $b_1$ | 0.822694 | 0.804796 |

5. For cups of fruits and vegetables (2005 MyPyramid definition), including and excluding French fries, for each gender, the estimates of the parameters are:

**Table 2- 10 Estimated Regression Coefficients for Sum of Foods Predicting Cups of Total Fruits and Vegetables and Fruits and Vegetables Excluding French Fries, by Gender**

| Parameter | Men | Women |
|---|---|---|
| **Summary Variable with French fries** | | |
| Intercept ($b_0$) | 0.529258 | 0.502480 |
| $b_1$ | 0.839721 | 0.792683 |
| **Summary Variable excluding French fries** | | |
| Intercept ($b_0$) | 0.559458 | 0.495205 |
| $b_1$ | 0.819797 | 0.794978 |

6. For Pyramid servings of dairy, estimation of $b_0$ and $b_1$, the model is:

$$E([\text{Dietary Factor}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2]^{1/2})$$

For Pyramid servings of dairy, for each gender, the estimates of the parameters are:

**Table 2- 11 Estimated Regression Coefficients for Sum of Foods Predicting Servings of Dairy, by Gender**

| Parameter | Men | Women |
|---|---|---|
| Intercept ($b_0$) | 0.417414 | 0.385301 |
| $b_1$ | 0.831739 | 0.782852 |

7. For Pyramid teaspoons of added sugar, estimation of $b_0$ and $b_1$, the model is:

$$E([\text{Dietary Factor}]^{1/3}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG4}P_4]^{1/3})$$

For Pyramid teaspoons of added sugar, for each gender, the estimates of the parameters are:

**Table 2- 12 Estimated Regression Coefficients for Sum of Foods Predicting Teaspoons of Added Sugar, by Gender**

| Parameter | Men | Women |
|---|---|---|
| Intercept ($b_0$) | 1.672746 | 1.591494 |
| $b_1$ | 0.534485 | 0.491231 |

# 3. Uses of Screener Estimates

## Introduction

Dietary intake estimates derived from the Five-Factor Screener are rough estimates of usual intake of fruits and vegetables, fiber, calcium, servings of dairy, and added sugar. These estimates are not as accurate as those from more detailed methods (e.g., 24-hour recalls). However, Validation Results suggests that the estimates may be useful to characterize a population's median intakes, to discriminate among individuals or populations with regard to higher vs. lower intakes, to track dietary changes in individuals or populations over time, and to allow examination of interrelationships between diet and other variables. In addition, diet estimates from the Cancer Control Supplement (CCS) could be used as benchmark national data for smaller surveys, for example, in a particular state.

## Variance-Adjustment Factor

### What is the variance adjustment estimate and why do we need it?

Data from the Five-Factor Screener are individuals' reports about their intake and, like all self-reports, contain some error. The algorithms we use to estimate servings of fruits and vegetables, grams of fiber, mg of calcium, servings of dairy, and teaspoons of added sugar calibrate the data to 24-hour recalls. The screener estimate of intake represents what we expect the person would have reported on his 24-hour recall, given what he reported on the individual items in the screener. As a result, the mean of the screener estimate of intake should equal the mean of the 24-hour recall estimate of intake in the population. (It would also equal the mean of true intake in the population if the 24-hour recalls were unbiased. However, there are many studies suggesting that recalls underestimate individuals' true intakes).

When describing a population's distribution of dietary intakes, the parameters needed are an estimate of central tendency (i.e. mean or median) and an estimate of spread (variance). The variance of the screener, however, is expected to be smaller than the variance of true intake, since the screener prediction formula estimates the conditional expectation of true

intake given the screener responses, and in general the variance of a conditional expectation of a variable is smaller than the variance of itself. As a result, the screener estimates of intake cannot be used to estimate quantiles (other than median) or prevalence estimates of true intake without an adjustment. Procedures have been developed to estimate the variance of true intake using data from 24-hour recalls, by taking into consideration within person variability [1, 2]. We extended these procedures to allow estimation of the variance of true intake using data from the screener. The resulting variance adjustment factor adjusts the screener variance to approximate the variance of true intake in the population.

## How did we estimate the variance adjustment factors?

We have estimated the adjustment factors in the two external validation datasets available to us: the Observing Protein and Energy Nutrition Study (OPEN) [3] and the Eating at America's Table Study (EATS) [4]. The results indicate that the adjustment factors differ by gender and dietary variable. Under the assumption that the variance adjustment factors appropriate to the 2005 National Health Interview Study (NHIS) are similar to those in these two external datasets, the variance-adjusted screener estimate of intake should have variance closer to the estimated variance of true intake that would have been obtained from repeat 24-hour recalls.

**Table 3- 1 Variance Adjustment Factors for the NHIS Multifactor Screener**

| Nutrient | Gender | Variance Adjustment Factor |
|---|---|---|
| **Total Fruit & Vegetable Intake** | Male | 1.2 |
| | Female | 1.0 |
| **Fruit & Vegetable Intake (excluding fried potatoes)** | Male | 1.2 |
| | Female | 1.0 |
| **Fiber Intake** | Male | 1.4 |
| | Female | 1.2 |
| **Calcium Intake** | Male | 1.0 |
| | Female | 0.9 |
| **Dairy Intake** | Male | 1.1 |
| | Female | 1.1 |
| **Added Sugar Intake** | Male | 1.5 |
| | Female | 1.3 |

## How do you use the variance adjustment estimates?

To estimate quantile values or prevalence estimates for an exposure, you should first adjust the screener so that it has approximately the same variance as true intake.

Adjust the screener estimate of intake by:

- multiplying intake by an adjustment factor (an estimate of the ratio of the standard deviation of true intake to the standard deviation of screener intake); and
- adding a constant so that the overall mean is unchanged.

The formula for the variance-adjusted screener is:

$$\text{variance-adjusted screener} = (\text{variance adjustment factor}) * (\text{unadjusted screener} - \text{mean}_{\text{unadj scr.}}) + \text{mean}_{\text{unadj scr.}}$$

This procedure is performed on the normally distributed version of the variable (i.e., Pyramid servings of fruits and vegetables is square-rooted; fiber is cube rooted; calcium is quarter-rooted; dairy is square-rooted; and added sugar is cube rooted). The results can then be back-transformed (e.g. cubed, squared, etc.) to obtain estimates in the original units.

The variance adjustment procedure is used to estimate prevalence of obtaining recommended intakes for the 2000 NHIS in Thompson et al., [5].

## When do you use variance adjustment estimates?

The appropriate use of the screener information depends on the analytical objective. Following is a characterization of suggested procedures for various analytical objectives.

**Table 3- 2 Suggested procedures for various analytical objectives**

| Analytical Objective | Procedure |
|---|---|
| Estimate mean or median intake in the population or within subpopulations. | First, transform the variable to normalized version. Then, use the unadjusted screener estimate of intake. |
| Estimate quantiles (other than median) of the distribution of intake in the population; estimate prevalence of attaining certain levels of dietary intake. | Use the variance-adjusted screener estimate. |

| Analytical Objective | Procedure |
|---|---|
| Classify individuals into exposure categories (e.g., meeting recommended intake vs. not meeting recommended intake) for later use in a regression model. | Use the variance-adjusted screener estimates to determine appropriate classification into categories. |
| Use the screener estimate as a continuous covariate in a multivariate regression model. | First, transform the variable to normalized version. Then, use the unadjusted screener estimate. |

## Attenuation of Regression Parameters Using Screener Estimates

When the screener estimate of dietary intake is used as a continuous covariate in a multivariate regression, the estimated regression coefficient will typically be attenuated (biased toward zero) due to measurement error in the screener. The "attenuation factor" [6] can be estimated in a calibration study and used to deattenuate the estimated regression coefficient (by dividing the estimated regression coefficient by the attenuation factor).

We estimated attenuation factors in the OPEN and EATS studies (see below). If you use these factors to deattenuate estimated regression coefficients, note that the data come from relatively small studies that consist of fairly homogeneous samples (primarily white, well-educated individuals).

**Table 3- 3 Estimated attenuation factors for screener predicted intake in the OPEN and EATS studies**

| Gender | Square-Root Fruit & Veg | Square-Root Fruit & Veg (excluding French Fries) | Cube-Root Fiber | Quarter-Root Calcium | Square-Root Dairy | Cube Root Added Sugar |
|---|---|---|---|---|---|---|
| **Men** | 0.75 | 0.79 | 0.75 | 0.57 | 0.80 | 0.93 |
| **Women** | 0.64 | 0.66 | 0.67 | 0.45 | 0.73 | 0.88 |

If you categorize the screener values into quantiles and use the resulting categorical variable in a linear or logistic regression, the bias (due to misclassification) is more complicated because the categorization can lead to differential misclassification in the screener [7]. Although methods may be available to correct for this [8, 9] , it is not simple, nor are we comfortable suggesting how to do it at this time.

Even though the estimated regression coefficients are biased (due to measurement error in the screener or misclassification in the categorized screener), tests of whether the regression coefficient is different from zero are still valid. For example, if one used the SUDAAN REGRESS procedure with fruit and vegetable intake (estimated by the screener) as a covariate in the model, one could use the Wald statistic provided by SUDAAN to test whether the regression coefficients were statistically significantly different from zero. This assumes that there is only one covariate in the model measured with error; when there are multiple covariates measured with error, the Wald test that a single regression coefficient is zero may not be valid, although the test that the regression coefficients for all covariates measured with error are zero is still valid.

# 4. Validation Results

Risk Factor Assessment Branch staff have assessed indirectly the validity of parts of the Five-Factor Screener in two studies: NCI's Observing Protein and Energy (OPEN) Study [3] and the Eating at America's Table Study (EATS) [4]. In both studies, multiple 24-hour recalls in conjunction with a measurement error model were used to assess validity.

**Table 4- 1 Estimated mean fruit and vegetables (F&V) servings*, Fiber, Calcium, Dairy Servings, and Added Sugar from 24HR and screener and de-attenuated Pearson correlation coefficient between true intake and screener, by gender: OPEN**

| Dietary Factor | N | Mean (95% CI) | | | De-attenuated Pearson Correlation Coefficient (SEE) |
|---|---|---|---|---|---|
| | | 24HR | Screener | Difference | |
| **Total Pyramid servings of F&V* (square root)** | | | | | |
| **Men** | 260 | 2.48 | 2.28 | -0.21[2] | 0.58 (0.066) |
| **Women** | 221 | 2.29 | 2.14 | -0.16[2] | 0.73 (0.078) |
| **Pyramid servings of F&V* -- Fries (square root)** | | | | | |
| **Men** | 260 | 2.43 | 2.22 | -0.21[2] | 0.61 (0.067) |
| **Women** | 221 | 2.25 | 2.11 | -0.15[2] | 0.74 (0.070) |
| **Grams of Fiber (cube root)** | | | | | |
| **Men** | 260 | 2.78 | 2.56 | -0.22[2] | 0.52 (0.59) |
| **Women** | 221 | 2.55 | 2.36 | -0.18[2] | 0.54 (0.70) |

| Dietary Factor | N | Mean (95% CI) | | | De-attenuated Pearson Correlation Coefficient (SEE) |
|---|---|---|---|---|---|
| | | 24HR | Screener | Difference | |
| **Mgs of Calcium (fourth root)** | | | | | |
| **Men** | 260 | 5.44 | 5.20 | -0.25[2] | 0.59 (0.066) |
| **Women** | 221 | 5.18 | 4.87 | -0.31[2] | 0.44 (0.080) |
| **Tsps of Added Sugar (cube root)** | | | | | |
| **Men** | 260 | 2.61 | 2.58 | -0.04 | 0.68 (0.039) |
| **Women** | 222 | 2.39 | 2.29 | -0.10[2] | 0.66 (0.045) |
| **Servings of Dairy (square root)** | | | | | |
| **Men** | 260 | 1.13 | 1.13 | -0.00 | 0.64 (0.041) |
| **Women** | 221 | 1.03 | 0.99 | -0.04 | 0.64 (0.44) |

* Using 1992 Food Guide Pyramid definitions of servings.
[1] $p < 0.05$
[2] $p < 0.01$

**Table 4- 2 Estimated mean F&V servings*, Fiber, Calcium, Dairy Servings, and Added Sugar from 24HR and screener and de-attenuated Pearson correlation coefficient between true intake and screener, by gender: EATS**

| Dietary Factor | N | Mean (95% CI) | | | De-attenuated Pearson Correlation Coefficient (SEE) |
|---|---|---|---|---|---|
| | | 24HR | Screener | Difference | |
| **Servings of Total F&V* (square root)** | | | | | |
| **Men** | 184 | 2.44 | 2.36 | -0.08[1] | 0.70 (0.058) |
| **Women** | 247 | 2.07 | 2.11 | 0.04 | 0.54 (0.060) |
| **Servings of F&V* -- Fries (square root)** | | | | | |
| **Men** | 184 | 2.34 | 2.29 | -0.05 | 0.72 (0.054) |
| **Women** | 247 | 2.01 | 2.06 | 0.05 | 0.55 (0.058) |

| Dietary Factor | N | Mean (95% CI) | | | De-attenuated Pearson Correlation Coefficient (SEE) |
|---|---|---|---|---|---|
| | | 24HR | Screener | Difference | |
| **Grams of Fiber (cube root)** | | | | | |
| **Men** | 184 | 2.73 | 2.59 | -0.12[1] | 0.60 (0.059) |
| **Women** | 247 | 2.41 | 2.35 | -0.06[2] | 0.55 (0.054) |
| **Mgs of Calcium (fourth root)** | | | | | |
| **Men** | 184 | 5.42 | 5.26 | -0.04[2] | 0.60 (0.061) |
| **Women** | 247 | 5.01 | 4.93 | -0.07 | 0.56 (0.053) |
| **Tsps of Added Sugar (cube root)** | | | | | |
| **Men** | 446 | 2.64 | 2.67 | 0.03 | 0.59 (0.037) |
| **Women** | 519 | 2.32 | 2.35 | 0.03 | 0.66 (0.032) |
| **Servings of Dairy (square root)** | | | | | |
| **Men** | 446 | 1.15 | 1.17 | 0.03 | 0.74 (0.32) |
| **Women** | 519 | 0.95 | 1.00 | 0.05[2] | 0.73< (0.029) |

* Using 1992 Food Guide Pyramid definitions of servings.

[a] Estimated from multiple 24-hour dietary recalls in a measurement error model.

[1] $p < 0.05$

[2] $p < 0.01$

**Table 4- 3 Median intakes of fruits and vegetables (Pyramid servings*) and added sugar (teaspoons) for NHANES 2001-02, NHIS 2005, and CHIS 2005 by gender and race/ethnicity**

| Gender and Race/Ethnicity | Pyramid Servings* of Fruits and Vegetables | | Teaspoons of Added Sugar | |
|---|---|---|---|---|
| | NHANES 2001-02 (24HR) | NHIS 2005 (9-item screener) | NHANES 2001-02 (24HR) | NHIS 2005 (4-item screener) |
| **Men** | | | | |
| **Non-Hispanic White** | 4.7 | 5.2 | 20.5 | 19.0 |

| Gender and Race/Ethnicity | Pyramid Servings* of Fruits and Vegetables | | Teaspoons of Added Sugar | |
|---|---|---|---|---|
| | NHANES 2001-02 (24HR) | NHIS 2005 (9-item screener) | NHANES 2001-02 (24HR) | NHIS 2005 (4-item screener) |
| **Non-Hispanic Black** | 4.2 | 5.3 | 23.6 | 21.4 |
| **Hispanic** | 5.2 | 5.8 | 21.6 | 21.4 |
| **Women** | | | | |
| **Non-Hispanic White** | 4.2 | 4.4 | 14.0 | 12.8 |
| **Non-Hispanic Black** | 3.5 | 4.3 | 19.3 | 15.5 |
| **Hispanic** | 4.0 | 4.6 | 16.3 | 14.5 |

* Using 1992 Food Guide Pyramid definitions of servings.

**Table 4- 4 Median intakes of fiber (gm) and calcium (mg) for NHANES 2003-04 and NHIS 2005, by gender and race/ethnicity**

| Gender and Race/Ethnicity | Fiber (gm) | | Calcium (mg) | |
|---|---|---|---|---|
| | NHANES 2003-04 (24HR) | NHIS 2005 (18-item screener) | NHANES 2003-04 (24HR) | NHIS 2005 (2-item screener) |
| **Men** | | | | |
| **Non-Hispanic White** | 15.7 | 17.9 | 1061 | 876 |
| **Non-Hispanic Black** | 12.7 | 17.4 | 787 | 753 |
| **Hispanic** | 17.2 | 22.3 | 917 | 864 |
| **Women** | | | | |
| **Non-Hispanic White** | 12.5 | 13.9 | 688 | 653 |

| Gender and Race/Ethnicity | Fiber (gm) | | Calcium (mg) | |
|---|---|---|---|---|
| | NHANES 2003-04 (24HR) | NHIS 2005 (18-item screener) | NHANES 2003-04 (24HR) | NHIS 2005 (2-item screener) |
| **Non-Hispanic Black** | 9.7 | 13.4 | 542 | 584 |
| **Hispanic** | 13.3 | 16.8 | 703 | 691 |

These validation results suggest that dietary exposure estimates computed for the 2005 NHIS Cancer Control Supplement (CCS) may be useful to compare subgroup means, especially for populations consuming mainstream diets. The estimates may be less useful for populations with more ethnic diets, including Asian and possibly Latino populations. Although significant error may be associated with these estimates of diet, we believe the exposure estimates still substantially reflect what individuals are actually consuming.

# 5. Computed Variables

NOTE: The dietary variables provided here are in their natural units. For most analyses, however, they must be transformed first, to approximate normal distributions. For fruits, vegetables, and dairy, use the square-root transformation; for fiber and added sugar, use the cube-root transformation; for calcium, use the quarter-root transformation. After analyses, the result variables can be back-transformed for easier interpretation.

The computed diet variables for the 2005 NHIS Five-Factor Screener are available for download in two formats -- SAS transport and comma-separated values (CSV). The files include the following variables:

- **FV** - Pyramid servings (1992 definition) of fruits and vegetables per day
- **FVAdj** - Adjusted Pyramid servings (1992 definition) of fruits and vegetables per day
- **FVNoFF** - Pyramid servings (1992 definition) of fruits and vegetables excluding French fries per day
- **FVNoFFAdj** - Adjusted Pyramid servings (1992 definition) of fruits and vegetables excluding French fries per day
- **FCE** - Cup equivalents of fruits (2005 definition) per day
- **FCEAdj** - Adjusted cup equivalents of fruits (2005 definition) per day
- **FVCE** - Cup equivalents of fruits and vegetables (2005 definition) per day

- **FVCEAdj** - Adjusted cup equivalents of fruits and vegetables (2005 definition) per day
- **FVCENoFF** - Cup equivalents of fruits and vegetables except French fries (2005 definition) per day
- **FVCENoFFAdj** - Adjusted cup equivalents of fruits and vegetables excluding French fries (2005 definition) per day
- **VCE** - Cup equivalents of vegetables (2005 definition) per day
- **VCEAdj** - Adjusted cup equivalents of vegetables (2005 definition) per day
- **VCENoFF** - Cup equivalents of vegetables excluding French fries (2005 definition) per day
- **VCENoFFAdj** - Adjusted cup equivalents of vegetables excluding French fries (2005 definition) per day
- **Fiber** - Fiber (gm) per day
- **FiberAdj** - Adjusted fiber (gm) per day
- **Dairy** - Pyramid servings (1992 definition) of dairy per day
- **DairyAdj** - Adjusted Pyramid servings (1992 definition) of dairy per day
- **Calcium** - Calcium (mg) per day
- **CalciumAdj** - Adjusted calcium (mg) per day
- **Sugar** - Added sugar (tsp) (1992 definition) per day
- **SugarAdj** - Adjusted added sugar (tsp) (1992 definition) per day
- **HHX** - Household identifier
- **FMX** - Family serial number
- **FPX** - Person number (in family)

These datasets and SAS program for this screener are in a folder on NCI's Short Dietary Assessment Instruments website.

The datasets are sorted in ascending order by the ID variables HHX, FMX, and FPX. All numeric variables have been rounded to the nearest 0.000001.

- Comma-separated Values File (nhis2005dietvars.08-15-2007.csv) - This zip file contains the comma-separated values file, which includes 25 variables, 31,428 records, and an additional record for the variable names.
- SAS Transport File (nhis2005dietvars.08-15-2007.v8x) - The SAS transport file includes 31,428 records and 15 variables. To access the SAS dataset, unzip the file, then use proc cimport.

For example:

```
proc cimport file='nhis2005dietvars.08-15-2007.v8x' data=dietvars'
```

- SAS Program (create.nhis2005dietvars.03-29-2007.v8x.sas)- This is the SAS program that created the dataset. This is just for reference; there should be no need to run this program.

# 6. References


1. National Research Council. *Nutrient Adequacy: Assessment Using Food Consumption Surveys.* Washington, DC: *National Academy Press*, 1986. Available for download from: https://www.nap.edu/catalog/618/nutrient-adequacy-assessment-using-food-consumption-surveys.
2. Institute of Medicine. *Dietary Reference Intakes: Applications in Dietary Assessment*. Washington, DC: *National Academy Press*, 2000. Available for download from: https://www.nap.edu/catalog/9956/dietary-reference-intakes-applications-in-dietary-assessment.
3. Subar AF, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol*, 2003. 158(1):1-13.
4. Subar AF, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *Am J Epidemiol*, 2001. 154(12): 1089-99.
5. Thompson FE, et al. Dietary intake estimates in the National Health Interview Survey, 2000: methodology, results, and interpretation. *J Am Diet Assoc*, 2005. 105(3): 352-63.
6. Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat Med*, 1989. 8(9): 1051-69; discussion 1071-3.
7. Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. *Am J Epidemio*l, 1991. 134(10): 1233-44.
8. Flegal KM, Brownie C, Haas JD. The effects of exposure misclassification on estimates of relative risk. *Am J Epidemiol*, 1986. 123(4): 736-51.
9. Morrissey MJ, Spiegelman D. Matrix methods for estimating odds ratios with misclassified exposure data: extensions and comparisons. *Biometrics*, 1999. 55(2): 338-44.