Dietary Screener in the National Health Interview Survey Cancer Control Supplement 2010

(NHIS 2010)

November 2019

DIETARY SCREENER: NHIS 2010

This information was originally published online on the Epidemiology and Genomic Research Program, National Cancer Institute's website as a reference for the Dietary Screener in the National Health Interview Survey Cancer Control Supplement 2010 (NHIS 2010). 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 the Internet at https://epi.grants.cancer.gov/diet/screeners/files.

Suggested citation for information contained in this report: The Dietary Screener in the National Health Interview Survey Cancer Control Supplement 2010. 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

DIETARY SCREENER: NHIS 2010

1. Overview

The National Health Interview Survey (NHIS) Cancer Control Supplement (CCS) is administered every five years and focuses on knowledge, attitudes, and practices in cancer-related health behaviors, screening, and risk assessment. The NHIS CCS covers a variety of topics, including diet and nutrition.

Diet can be assessed using a variety of tools. Food records and 24-hour dietary recalls assess the total diet over one or more days, and they are considered the most accurate instruments. In certain situations, however, such as when time is constrained or assessment of the total diet is not required, researchers have found that short dietary assessment instruments, often called screeners, are valuable. For example, screeners can be used for characterizing a population's median intakes or examining interrelationships between diet and other variables. Recognizing the need for these tools, the NCI's Risk Factor Assessment Branch (RFAB) has developed several dietary screeners. One of these is the 26-item Dietary Screener Questionnaire (DSQ), which asks about the frequency of consumption in the past month of selected foods and drinks. The DSQ captures intakes of fruits and vegetables, dairy/calcium, added sugars, whole grains/fiber, red meat, and processed meat.

The DSQ was included in both the National Health and Nutrition Examination Survey (NHANES) 2009-10 and the National Health Interview Survey Cancer Control Supplement (NHIS CCS) 2010.

You can view or print the DSQ from the National Cancer Institute's (NCI) Register of Validated Short Dietary Assessment Instruments.

Because screeners are shorter and less detailed than a total dietary assessment, some quantitative accuracy is sacrificed. However, calibrating a screener against the more precise 24-hour recall can help ensure that a screener is providing the best and most accurate estimates possible.

The NHANES 2009-10 collected two non-consecutive 24-hour recalls, in addition to the DSQ. The DSQ was calibrated to the more precise recalls, by developing scoring algorithms for each component of the DSQ. The scoring algorithms convert screener responses to estimates of dietary intake for fruits and vegetables (cup equivalents), dairy (cup equivalents), added sugars (tsp), whole grains (ounce equivalents), fiber (g), and calcium (mg). Responses to the red meat and processed mat questions may be used as qualitative indicators of intake frequency, but no scoring algorithms were developed for those particular dietary factors.

The DSQ administered in the NHIS CCS 2010 was nearly identical to the DSQ administered in the NHANES 2009-10. The only difference is that in the NHIS CCS 2010, two questions were asked to capture intakes of sports drinks, energy drinks, and fruitades, whereas only one question was asked in the NHANES 2009-10 DSQ. The instruments used in the 2010

NHIS CCS and NHANES 2009-10 are available through NCI's Register of Validated Short Dietary Assessment Instruments.

Questions Asked

Table 1- 1 Differences in questions asked in the NHIS CCS 2010 and the NHANES 2009-10 DSQ

NHANES 2009-10 DSQ	NHIS CCS 2010 DSQ
How often did you drink **sweetened** fruit drinks, sports or energy drinks, such as Kool-aid, lemonade, Hi-C, cranberry drink, Gatorade, Red Bull or Vitamin Water? Include fruit juices you made at home and added sugar to. Do **not** include diet drinks or artificially sweetened drinks.	How often did you drink SPORTS and ENERGY drinks such as Gatorade, Red Bull, and Vitamin water?
	How often did you drink sweetened fruit drinks, such as Kool-aid, cranberry and lemonade? Include fruit drinks you made at home and added sugar to.

The scoring algorithms for the NHANES 2009-10 DSQ have been applied to the NHIS CCS 2010 DSQ. At the time these scoring algorithms were developed, food pattern equivalents information from the NHANES 2009-10 24-hour recall data had not yet been released. Such information is necessary to perform direct calibration of the screener with the 24-hour recalls. Therefore, RFAB developed alternative scoring algorithms, based on earlier NHANES data. These algorithms are presented in the Data Processing and Scoring Procedures section. That section also describes in detail the procedures used to develop the algorithms. For those interested only in accessing the NHIS variables, please visit the Computed Variables page. Algorithms based on direct calibration, using NHANES 2009-2010 data, are also available on the NCI's Short Dietary Assessment Instruments site.

NCI's Short Dietary Assessment Instruments website also contains additional information about screeners and how they are used, including information about other screeners developed by NCI's RFAB.

2. Data Processing and Scoring Procedures

Data Processing

Our NCI research team followed several steps to formulate the Dietary Screener Questionnaire (DSQ) scoring algorithms. These steps are described for researchers who may be interested in the methodologic process our team used. However, it is not necessary for researchers to follow these steps; a SAS program that integrates these steps is publicly available, as are Computed Variables for the dietary variables in the NHIS CCS.

Our steps consisted of:

1. Converting Frequency Data to Daily Frequency
2. Identifying Extreme Exposure Values
3. Classifying Cereal Data.

Converting Frequency Data to Daily Frequency

Reported frequency responses were converted to daily estimates using the procedures described DSQ in the NHANES 2009-10: Data Processing and Scoring Procedures, information in the Appendix. The frequency responses for sports drinks/energy drinks and fruitades were summed, thus creating the fruitades/sports drinks variable parallel to that in the NHANES DSQ.

Identifying Extreme Exposure Values

The definition of extreme values for the NHANES 2009-10 DSQ was used in the NHIS CCS 2010 DSQ. These values and the number of excluded values in the NHIS CCS 2010 DSQ are presented below.

Table 2- 1 Definition of Extreme Values for NHIS 2010

Food Group	NHIS CCS 2010 (N=27,157 Adults 18+)	
	Maximum Acceptable Daily Frequency Value	Number of Excluded Values
Fruit	8	3
Fruit juice	8	2
Salad	5	55
Fried potatoes	5	12
Other potatoes	3	37

DIETARY SCREENER: NHIS 2010

Food Group	NHIS CCS 2010 (N=27,157 Adults 18+)	
	Maximum Acceptable Daily Frequency Value	Number of Excluded Values
Dried beans	4	29
Other vegetables	5	71
Tomato sauce	2	32
Salsa	3	38
Pizza	2	14
Soda	8	23
Sports drinks/Fruit drinks	7	27
Cookies, cake, pie	7	3
Doughnuts	5	3
Frozen desserts	5	9
Sugar/honey in coffee/tea	10	0
Candy	8	6
Any milk (not soy)	10	0
Cheese	6	43
Any cereal	7	9
Whole grain bread	6	55
Cooked whole grains	4	20
Popcorn	3	8
Red meat	6	23
Processed meat	4	16

Classifying Cereal Data

The DSQ, which was included in NHIS 2010, includes questions about cereal intake and allows respondents up to two responses on which cereals they consume. We classified each cereal reported first by hot or cold, and then along four dimensions: density of added sugars, whole grains, fiber, and calcium, as described in the DSQ in NHANES 2009-10: Data Processing and Scoring Procedures, included in the Appendix.

Scoring Procedures

Developing scoring algorithms

We developed scoring algorithms to convert screener responses to estimates of individual dietary intake for fruits and vegetables (cup equivalents), dairy (cup equivalents), added sugars (tsp), added sugars from sugar-sweetened beverages, whole grains (ounce equivalents), fiber (g), and calcium (mg) using the What We Eat in America 24-hour dietary recall data from the 2003-2006 NHANES. Equations were estimated in the NHANES 2003-2006, using SAS PROC REG.

Equations estimated in the NHANES 2003-2006, using SAS PROC REG

For cup equivalents of fruits and vegetables:

$$E([\text{Fruits and Veg}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG10}P_{10}]^{1/2})$$

Cup equivalents of fruits and vegetables were 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 in fruit and vegetable cup equivalents of group k; and k indexes the ten fruit and vegetable food groups. We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values.

For cup equivalents of fruits:

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

Cup equivalents of fruits 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 in fruit cup equivalents of group k; and k indexes the two fruit food groups: fruit, and 100% fruit juices. We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values.

For cup equivalents of vegetables:

$$E([\text{Veg}]^{1/2}) = b_0 + b_1([N_{FG3}P_3 + N_{FG4}P_4 + \ldots + N_{FG10}P_{10}]^{1/2})$$

Cup equivalents of 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 in vegetable cup equivalents of group k; and k indexes the eight vegetable food groups. We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values.

For cup equivalents of dairy:

$$E([\text{Dairy}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG4}P_4]^{1/2})$$

Cup equivalents of dairy were 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 in dairy cup equivalents of group k; and k indexes the four dairy food groups. We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 on the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values.

For teaspoons of added sugars:

$$E([\text{Added Sugars}]^{.33}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG13}P_{13}]^{.33})$$

Teaspoons of added sugars were cube-root-transformed to approximate normality; N_{FGk} is the usual number of times per day that an individual consumed food group k; P_k is the median portion size in added sugars teaspoons of group k; and k indexes the thirteen added sugars food groups.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values.

For teaspoons of added sugars from sugar-sweetened beverages:

$$E([\text{Added Sugars}_{ssb}]^{.5}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + N_{FG3}P_3]^{.5})$$

The dependent variable added sugars from sugar-sweetened beverages included the following beverages: sodas, fruitades/sports drinks, and sugar in coffee/tea. Teaspoons of added sugars were square-root-transformed to approximate normality; N_{FGk} is the usual number of times per day that an individual consumed food group k; P_k is the median portion size in added sugars teaspoons of group k; and k indexes the three sugar-sweetened beverage food groups.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values.

For ounce-equivalents of whole grains:

$$E([\text{Whole Grains}]^{1/2}) = b_0 + b_1(N_{FG1}P_1) + b_2(N_{FG2}P_2) + \ldots + b_9(N_{FG9}P_9)$$

Ounce-equivalents of whole grains were 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 in whole grain ounce-equivalents of group k; and k indexes the nine whole grain food groups: hot cereals (2 groups), cold cereals (4 groups), brown rice, whole grain bread, and popcorn.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_k, k = 0, ..., 9 on the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values.

For grams of fiber:

$$E([\text{Fiber}]^{1/4}) = b_0 + b_1(N_{FG1}P_1) + b_2(N_{FG2}P_2) + \ldots + b_{29}(N_{FG29}P_{29})$$

Grams of fiber were quarter-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 in grams of group k; and k indexes the 28 food groups. Two food groups, tomato sauce, and dried beans, were frequently consumed in mixed dishes. Because we wanted to represent only the particular food, we estimated the number of grams of tomatoes and legumes, respectively, in these mixed dishes. For tomato sauce, this was done by multiplying the MPED 1 cup equivalent by 171 (1 MPED = 171 grams). For legumes, this was done by estimating the MPED one cup equivalent for legumes and multiplying by 212 (1 MPED = 212 grams).

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_k, k = 0,..., 29 on the NHANES 2003-2006 samples, stratifying by sex, and excluding extreme exposure values. We first included all 29 food groups in the regression model. After examining the results, we dropped food groups that failed to attain statistical significance at $\alpha = 0.25$ to form more parsimonious final models. In the fiber model, soda and sugar in coffee/tea were dropped for men, and salad and sugar in coffee/tea were dropped for women.

For milligrams of calcium:

$$E([\text{Calcium}]^{1/4}) = b_0 + b_1(N_{FG1}P_1) + b_2(N_{FG2}P_2) + \ldots + b_{29}(N_{FG29}P_{29})$$

Milligrams of calcium were quarter-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 in grams of group k; and k indexes the 29 food groups. Two food groups, tomato sauce, and dried beans, were frequently consumed in mixed dishes. Because we wanted to represent only the particular food, we estimated the number of grams of tomatoes and legumes, respectively, in these mixed dishes. For tomato sauce, this was done by multiplying the MPED 1 cup equivalent by 171 (1 MPED = 171 grams). For legumes, this was done by estimating the MPED one cup equivalent for legumes and multiplying by 212 (1 MPED = 212 grams).

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_k, k = 0, ..., 29 on the NHANES 2003-06 samples, stratifying by sex, and excluding extreme exposure values. We first included all 29 food groups in the regression model. After examining the results, we dropped food groups that failed to attain statistical significance at $\alpha = 0.25$ to form more parsimonious final models. In the calcium model, popcorn, brown rice, soda, sugar in coffee/tea, and other potatoes were dropped for men; and popcorn, brown rice, fruitades/sports drinks, doughnuts, and tomato sauce were dropped for women.

Estimating intakes

Using the equations above, we performed the following steps to estimate the individual intake of fruits and vegetables (cup equivalents), dairy (cup equivalents), added sugars (tsp), whole grains (ounce equivalents), fiber (g), and calcium (mg).

- **Estimation of Pk:**
  The median sex- and age-specific portion sizes for each food were estimated from NHANES 2003-2006. For fruit and vegetable variables, the unit was MyPyramid* cup equivalents (Table 2-2). For dairy, the unit was MyPyramid cup equivalents (Table 2-3). For added sugars and sugar-sweetened beverages the unit was MyPyramid teaspoons of added sugars (Table 2-4 and Table 2-5). For whole grains, the unit was grams (Table 2-6). For fiber and calcium, the unit was grams (Table 2-7).

  For fruits and vegetables, a MyPyramid cup equivalent is defined by the U.S. Department of Agriculture (USDA) and the U.S. Department of Health and Human Services (HHS) in the Dietary Guidelines for Americans, 2010 as:

  - vegetables: 1 cup raw or cooked; 1 cup vegetable juice; or 2 cups leafy salad greens
  - fruit: 1 cup raw or cooked; ½ cup dried fruits; or 1 cup fruit juice.

  For dairy, a MyPyramid cup equivalent is defined by the USDA and HHS in the Dietary Guidelines for Americans, 2010 as:

  - 1 cup milk, fortified soy beverage, or yogurt
  - 1 ½ ounces of natural cheese
  - 2 ounces of processed cheese

DIETARY SCREENER: NHIS 2010

The exposure sugar-sweetened beverages is defined in the Dietary Guidelines for Americans, 2010 as: "Liquids that are sweetened with various forms of sugars that add calories. These beverages include, but are not limited to, soda, fruitades and fruit drinks, and sports and energy drinks." For our analyses, we defined this exposure as including the above types of drinks plus coffees and teas when sweetened with sugar.

For whole grains, a MyPyramid ounce equivalent is defined by the USDA and HHS in the Dietary Guidelines for Americans, 2010 as:

- 1 one-ounce slice of whole grain bread
- 1 ounce of uncooked whole grain pasta or rice
- ½ cup of cooked whole grain rice, pasta, or cereal
- 1 whole grain tortilla (6-inch diameter)
- 1 whole grain pancake (5-inch diameter)
- 1 ounce (or about 1 cup cereal flakes) ready-to-eat whole grain cereal.

* MyPyramid is a food guidance system featured in the Dietary Guidelines for Americans, 2005. In order to examine food intakes in relation to MyPyramid, the MyPyramid Equivalent Database (MPED) was developed. It translates reported food intake into guidance-based food groups. The 2010 Dietary Guidelines used an update to USDA's food pattern guidance, and more recent versions of the MPED are now called Food Pattern Equivalent Databases (FPED).

Table 2- 2 Median Portion Size (Pk) in Cup Equivalents per Mention by Sex & Age for Fruits & Vegetables Analyses

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Men							
Fruit (P_1)	0.867300	0.911790	0.867300	0.867300	0.813000	0.715500	0.713000
100% Fruit juice (P_2)	1.480620	1.458000	1.437000	1.147570	0.999440	0.812045	0.780813
Salad (P_3)	0.291000	0.375000	0.390500	0.400500	0.499950	0.407500	0.388810
Fried potatoes (P_4)	0.622000	0.663435	0.548000	0.548055	0.545000	0.455000	0.444000
Other potatoes (P_5)	0.836000	0.767025	0.874650	0.930000	0.809676	0.809676	0.536000
Dried beans (P_6)	0.563000	0.524700	0.597726	0.563000	0.480533	0.377000	0.411000
Other vegetables (P_7)	0.333401	0.410997	0.388250	0.410901	0.431000	0.442680	0.375000
Tomato sauce (P_8)	0.612000	0.688000	0.637962	0.925000	0.861500	0.861800	0.486000
Salsa (P_9)	0.177021	0.212160	0.198000	0.265000	0.206392	0.132000	0.231000
Pizza (P_{10})	0.647000	0.729000	0.694861	0.693760	0.759000	0.542720	0.293000

DIETARY SCREENER: NHIS 2010

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Women							
Fruit (P_1)	0.764240	0.750000	0.739187	0.775915	0.742000	0.633000	0.632000
100% Fruit juice (P_2)	1.242252	1.003392	1.000000	1.000000	0.813000	0.749580	0.750000
Salad (P_3)	0.318225	0.395500	0.500000	0.500000	0.437502	0.375000	0.333500
Fried potatoes (P_4)	0.480000	0.446660	0.444000	0.385000	0.455275	0.444000	0.320000
Other potatoes (P_5)	0.749700	0.500000	0.697000	0.555500	0.570960	0.666000	0.515000
Dried beans (P_6)	0.390000	0.342000	0.321000	0.333000	0.341398	0.323400	0.305000
Other vegetables (P_7)	0.308000	0.361100	0.333200	0.389714	0.396680	0.389083	0.347200
Tomato sauce (P_8)	0.624000	0.480000	0.388000	0.537240	0.563000	0.710000	0.336660
Salsa (P_9)	0.129000	0.212160	0.174877	0.206392	0.138705	0.231000	0.097852
Pizza (P_{10})	0.433541	0.444000	0.432000	0.431609	0.488653	0.569016	0.390861

Table 2- 3 Median Portion Size (Pk) in Cup Equivalents per Mention by Sex & Age for Dairy Analyses

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Men							
Cheese (P_1)	0.708341	0.584210	0.648000	0.595350	0.586278	0.518500	0.505700
Pizza (P_2)	1.317000	1.500000	1.190850	1.325659	1.166539	1.113000	0.726591
Milk (P_3)	1.313025	1.313025	1.311965	1.000400	1.000000	0.833333	0.781563
Frozen desserts (P_4)	0.261840	0.238967	0.264000	0.264000	0.271000	0.220448	0.226000
Women							
Cheese (P_1)	0.500250	0.471000	0.470000	0.469920	0.461080	0.469920	0.432075
Pizza (P_2)	0.878526	0.853000	0.851482	0.968774	0.988000	0.992000	0.713000
Milk (P_3)	1.042083	1.000000	0.968000	0.999000	0.750300	0.750000	0.708500
Frozen desserts (P_4)	0.198000	0.201000	0.205000	0.226980	0.192889	0.221000	0.194480

Table 2- 4 Median Portion Size (Pk) in Teaspoons per Mention by Sex & Age for Added Sugars Analyses

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Men							
Hot cereals (P_1)	1.758667	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Cold cereals (P_2)	3.840000	3.939187	2.724000	2.515333	1.594763	1.449000	1.127508
Soda (P_3)	11.601377	10.544000	10.327300	9.302000	8.994700	8.994700	7.889000
Sugar/honey in coffee/tea (P_4)	4.241000	3.807000	3.395571	3.000186	2.000000	2.120544	2.222844
Fruitades/sports drinks (P_5)	8.559500	8.232800	10.264077	8.113458	7.558200	5.439624	5.439624
Candy (P_6)	3.475000	3.259468	3.031745	3.054975	2.974000	2.359090	2.445380
Doughnuts (P_7)	4.282600	4.262250	3.708000	3.517150	4.289115	2.789390	3.127400
Cookies, cake, pie, brownies (P_8)	3.761000	3.756900	4.387860	3.735000	3.192000	3.230824	2.667387
Frozen desserts (P_9)	5.052000	5.262000	5.052000	4.680000	4.508965	3.691000	3.849000
Women							
Hot cereals (P_1)	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Cold cereals (P_2)	2.827000	2.585755	2.393280	1.843222	0.793000	0.972800	0.789570
Soda (P_3)	9.346200	8.947000	7.953000	8.617500	7.865340	7.295500	5.917000

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Sugar/honey in coffee/tea (P_4)	3.053667	2.884500	2.827392	2.153000	1.983456	1.982667	1.382000
Fruitades/sports drinks (P_5)	7.258570	7.493000	6.451665	6.230040	6.798000	4.536750	4.113113
Candy (P_6)	2.703000	2.121000	2.275000	2.248500	2.098600	1.858320	1.686480
Doughnuts (P_7)	4.191040	3.112720	3.127000	2.781216	3.242000	3.150500	2.383229
Cookies, cake, pie, brownies (P_8)	2.661000	3.340980	3.651000	3.208920	2.821134	2.729500	2.572640
Frozen desserts (P_9)	3.509870	3.875520	3.300180	3.094000	2.949000	3.350500	3.300180

Table 2- 5 Median Portion Size (Pk) in Teaspoons per Mention by Sex & Age for Sugar-Sweetened Beverage Analyses

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Men							
Soda (P_1)	11.601377	10.544000	10.327300	9.302000	8.994700	8.994700	7.889000
Sugar in coffee/tea (P_2)	4.241000	3.807000	3.395571	3.000186	2.000000	2.120544	2.222844
Fruitades/sports drinks (P_3)	8.559500	8.232800	10.264077	8.113458	7.558200	5.439624	5.439624
Women							
Soda (P_1)	9.346200	8.947000	7.953000	8.617500	7.865340	7.295500	5.917000
Sugar in coffee/tea (P_2)	3.053667	2.884500	2.827392	2.153000	1.983456	1.982667	1.382000

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Fruitades/sports drinks (P_3)	7.258570	7.493000	6.451665	6.230040	6.798000	4.536750	4.113113

Table 2- 6 Median Portion Size (Pk) in Grams per Mention by Sex & Age for Whole Grains Analyses

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Men							
Hot cereals (P_1)	308.000000	256.000000	242.000000	234.000000	234.000000	234.000000	215.345000
Cold cereals (P_2)	59.095000	63.440000	56.720000	56.310000	50.750000	43.030000	39.000000
***Brown rice (P_3)**	170.630000	170.630000	159.000000	159.000000	115.785000	115.785000	115.785000
Whole grain bread (P_4)	52.000000	50.666667	52.000000	52.000000	50.666667	43.000000	44.000000
Popcorn (P_5)	25.380000	54.310000	56.250000	39.880000	37.500000	51.840000	44.565000
Women							
Hot cereals (P_1)	248.630000	211.750000	234.000000	175.500000	181.500000	204.750000	167.340000
Cold cereals (P_2)	41.000000	45.000000	45.000000	39.380000	39.380000	30.625000	30.000000
***Brown rice (P_3)**	139.130000	139.130000	164.000000	164.000000	98.000000	98.000000	98.000000
Whole grain bread (P_4)	48.000000	45.500000	42.250000	46.000000	40.000000	37.000000	29.000000
Popcorn (P_5)	25.380000	35.000000	30.400000	26.500000	22.800000	28.000000	28.000000

*Because of the small number of consumers of brown rice, certain age/sex groups were collapsed: by sex, ages 18-37; by sex, ages 38-57; by sex, ages 58 +.

Table 2- 7 Median Portion Size (Pk) in Grams per Mention by Sex & Age for Fiber & Calcium Analyses

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Men							
Hot cereals (P_1)	308.000000	256.000000	242.000000	234.000000	234.000000	234.000000	215.345000
Cold cereals (P_2)	59.095000	63.440000	56.720000	56.310000	50.750000	43.030000	39.000000
Milk (P_3)	325.333333	320.250000	320.250000	250.000000	244.000000	205.116000	198.250000
Soda (P_4)	495.637500	461.250000	450.266667	370.900000	370.000000	368.400000	335.500000
100% Fruit juices (P_5)	372.000000	372.000000	357.940000	287.083333	248.000000	201.500000	193.750000
Sugar added to coffee/tea (P_6)	253.300000	83.975000	39.540000	25.200000	10.400000	12.500000	12.600000
Fruitades/sports drinks (P_7)	496.000000	453.130000	540.310000	458.560000	352.190000	252.000000	248.000000
Fruit (P_8)	126.666667	127.500000	128.000000	120.835000	118.000000	109.500000	109.500000
Salad (P_9)	30.000000	41.250000	41.250000	40.315000	53.845000	43.500000	41.250000
Fried potatoes (P_{10})	99.065000	114.000000	85.500000	88.000000	85.000000	85.500000	72.750000

DIETARY SCREENER: NHIS 2010

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Other potatoes (P11)	147.000000	157.500000	173.000000	160.500000	157.500000	142.500000	120.375000
Dried beans (P12)	119.356000	111.236400	126.717912	119.356000	101.872890	79.924000	87.132000
***Brown rice (P13)**	170.630000	170.630000	159.000000	159.000000	115.785000	115.785000	115.785000
Other vegetables (P14)	73.757143	80.500000	77.908571	74.250000	75.422500	75.100000	60.830000
Salsa (P15)	26.700000	32.000000	31.130000	43.380000	31.130000	29.250000	32.000000
Pizza (P16)	261.380000	318.090000	255.000000	265.070000	255.820000	213.120000	132.320000
Tomato sauce (P17)	104.652000	117.648000	109.091425	158.175000	147.316500	147.367800	83.106000
Cheese (P18)	37.044000	29.767500	32.400000	29.414338	29.988000	25.901400	25.850000
Whole grain bread (P19)	52.000000	50.666667	52.000000	52.000000	50.666667	43.000000	44.000000
Candy (P20)	28.000000	26.000000	25.333333	24.675000	22.500000	20.000000	20.925000
Doughnuts (P21)	93.000000	83.250000	68.510000	64.500000	80.500000	60.000000	65.000000
Cookies, cake, pie, brownies (P22)	53.000000	53.333333	66.000000	50.000000	45.000000	48.000000	40.000000
Frozen desserts (P23)	132.000000	124.875000	116.380000	116.380000	118.000000	95.500000	93.670000
Popcorn (P24)	25.380000	54.310000	56.250000	39.880000	37.500000	51.840000	44.565000

DIETARY SCREENER: NHIS 2010

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Women							
Hot cereals (P_1)	248.630000	211.750000	234.000000	175.500000	181.500000	204.750000	167.340000
Cold cereals (P_2)	41.000000	45.000000	45.000000	39.380000	39.380000	30.625000	30.000000
Milk (P_3)	266.875000	245.000000	244.000000	244.000000	187.577500	183.750000	183.000000
Soda (P_4)	392.100000	369.000000	368.200000	368.400000	366.000000	307.000000	248.000000
100% Fruit juices (P_5)	313.875000	263.500000	248.800000	248.800000	202.310000	186.000000	186.600000
Sugar added to coffee/tea (P_6)	220.880000	132.865000	89.900000	21.000000	30.000000	42.380000	8.400000
Fruitades/sports drinks (P_7)	357.940000	355.500000	332.060000	321.560000	252.800000	205.560000	183.600000
Fruit (P_8)	118.000000	115.750000	107.000000	117.000000	107.000000	100.500000	104.282500
Salad (P_9)	33.606667	41.250000	55.000000	46.876667	48.125000	41.250000	32.655000
Fried potatoes (P_{10})	72.500000	68.333333	70.000000	61.000000	72.000000	66.000000	57.000000
Other potatoes (P_{11})	144.750000	122.000000	127.000000	125.000000	118.130000	116.220000	112.500000
Dried beans (P_{12})	82.680000	72.504000	68.052000	70.596000	72.376418	68.560800	64.660000
***Brown rice (P_{13})**	139.130000	139.130000	164.000000	164.000000	98.000000	98.000000	98.000000
Other vegetables (P_{14})	63.950000	66.500000	58.907500	62.480000	67.500000	63.416667	61.080000

DIETARY SCREENER: NHIS 2010

Food Group	Age Group (years)						
	18-27	28-37	38-47	48-57	58-67	68-77	>78
Salsa (P15)	22.500000	31.130000	29.250000	31.130000	20.750000	32.000000	17.800000
Pizza (P16)	180.960000	168.390000	161.005000	157.790000	176.000000	250.830000	145.395000
Tomato sauce (P17)	106.704000	82.080000	66.348000	91.868040	96.273000	121.410000	57.568860
Cheese (P18)	25.384278	24.250000	23.496000	23.500000	23.187780	23.500000	21.700000
Whole grain bread (P19)	48.000000	45.500000	42.250000	46.000000	40.000000	37.000000	29.000000
Candy (P20)	22.750000	19.800000	20.500000	18.000000	18.000000	16.000000	14.000000
Doughnuts (P21)	81.500000	64.000000	65.000000	64.000000	66.000000	62.000000	56.000000
Cookies, cake, pie, brownies (P22)	37.000000	44.550000	51.680000	48.750000	41.000000	40.000000	40.000000
Frozen desserts (P23)	92.375000	99.750000	92.000000	88.750000	89.000000	88.665000	78.000000
Popcorn (P24)	25.380000	35.000000	30.400000	26.500000	22.800000	28.000000	28.000000

*Because of the small number of consumers of brown rice, some age/sex groups were combined: by sex, ages 18-37 years; by sex, ages 38-57; by sex, ages 58+.

- **Estimation of b0 and b1:**

Table 2- 8 Estimated Regression Coefficients for Sum of Foods Predicting Square Root Cup Equivalents of Total Fruits & Vegetables Including/Excluding French Fries by Sex

For MyPryamid cup equivalents of **fruits and vegetables**, estimation of b_0 and b_1, the model is:

$$E\left([\text{Fruits and Veg}]^{1/2}\right) = b_0 + b_1\left([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG10}P_{10}]^{1/2}\right)$$

For MyPyramid cup equivalents of **fruits and vegetables**, including and excluding French fries, for each sex, the estimates of the parameters are:

Parameter	Men	Women
Summary Variable including French fries		
Intercept (b_0)	0.569237	0.475853
b_1	0.811703	0.835344
Summary Variable excluding French fries		
Intercept (b_0)	0.578786	0.482116
b_1	0.805061	0.831300

Table 2- 9 Estimated Regression Coefficients for Sum of Foods Predicting Square Root Cup Equivalents of Fruits by Sex

For MyPryamid cup equivalents of **fruits**, estimation of b_0 and b_1, the model is:

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

For MyPyramid cup equivalents of **fruits**, for each sex, the estimates of the parameters are:

Parameter	Men	Women
Intercept (b_0)	0.197558	0.219778
b_1	0.904235	0.868067

Table 2- 10 Estimated Regression Coefficients for Sum of Foods Predicting Square Root Cup Equivalents of Vegetables including Legumes & Fried Potatoes by Sex

For MyPryamid cup equivalents of **vegetables**, estimation of b_0 and b_1, the model is:

$$E([Vegetables]^{1/2}) = b_0 + b_1([N_{FG3}P_3 + N_{FG4}P_4 + ... + N_{FG10}P_{10}]^{1/2})$$

For MyPyramid cup equivalents of **vegetables**, for each sex, the estimates of the parameters are:

Parameter	Men	Women
Intercept (b_0)	0.576577	0.489698
b_1	0.735432	0.754079

Table 2- 11 Estimated Regression Coefficients for Sum of Foods Predicting Square Root Cup Equivalents of Dairy by Sex

For MyPyramid cup equivalents of **dairy**, estimation of b_0 and b_1, the model is:

$$E([Dairy]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + ... + N_{FG4}P_4]^{1/2})$$

For MyPyramid cup equivalents of **dairy**, for each sex, the estimates of the parameters are:

Parameter	Men	Women
Intercept b_0	0.354193	0.388273
b_1	0.889021	0.832711

Table 2- 12 Estimated Regression Coefficients for Sum of Foods Predicting Cube Root Teaspoons of Added Sugars by Sex

For MyPyramid teaspoons of **added sugars including cereals**, estimation of b_0 and b_1, the model is:

$$E([Dietary\ Factor]^{1/3}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + ... + N_{FG13}P_{13}]^{1/3})$$

For MyPyramid teaspoons of **added sugars**, for each sex, the estimates of the parameters are:

Parameter	Men	Women
Intercept b_0	0.909718	0.868961

Parameter	Men	Women
b_1	0.763567	0.749127

Table 2- 13 Estimated Regression Coefficients for Sum of Foods Predicting Square Root Teaspoons of Added Sugars from Sugar-Sweetened Beverages by Sex

For MyPyramid teaspoons of **added sugars from sugar-sweetened beverages**, estimation of b_0 and b_1, the model is:

$$E([\text{Dietary Factor}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + N_{FG3}P_3]^{1/2})$$

Food groups used are: soda, fruitades/sports drinks, and sugar in coffee/tea. For MyPyramid teaspoons of **added sugars from sugar-sweetened beverages**, for each sex, the estimates of the parameters are:

Parameter	**Men**	**Women**
Intercept b_0	0.013561	-0.023338
b_1	1.059944	1.054829

Table 2- 14 Estimated Regression Coefficients for Foods as Predictors of Square Root of Ounce-Equivalents of Whole Grains by Sex

For MyPyramid ounce-equivalents of **whole grains**, estimation of b_0 and $b_1 ... b_9$, the model is:

$$E([\text{Whole Grains}]^{.50}) = b_0 + b_1N_{FG1}P_1 + b_2N_{FG2}P_2 + ... + b_9 N_{FG9}P_9$$

For MyPyramid ounce-equivalents of **whole grains**, for each sex, the estimates of the parameters are:

Parameter	**Men**	**Women**
Intercept (b_0)	0.266412	0.262635
Hot cereals: no whole grain (b_1)	-0.000254	-0.000084
Hot cereals: any whole grain (b_2)	0.004534	0.005142
Cold cereals: no whole grain (b_3)	-0.001101	-0.001681
Cold cereals: 2nd quartile whole grain (b_4)	0.011122	0.010002

Parameter	Men	Women
Cold cereals: 3rd quartile whole grain (b_5)	0.012334	0.013282
Cold cereals: 4th quartile whole grain (b_6)	0.016685	0.018480
Brown rice (b_7)	0.008077	0.005883
Whole grain bread (b_8)	0.010057	0.010776
Popcorn (b_9)	0.017884	0.022795

Table 2- 15 Estimated Regression Coefficients for Foods as Predictors of Quarter Root Fiber (g) by Sex

For grams of **fiber**, estimation of b_0 and $b_1 ... b_{29}$, the model is:

$$E([Fiber]^{.25}) = b_0 + b_1 N_{FG1} P_1 + b_2 N_{FG2} P_2 + ... + b_9 N_{FG29} P_{29}$$

For quarter root grams of **fiber**, for each sex, the estimates of the parameters are:

Parameter	Men	Women
Intercept (b_0)	1.677395	1.601281
Hot cereals, 1st tertile fiber (b_1)	0.000034	---
Hot cereals, 2nd tertile fiber (b_2)	0.000459	0.000621
Hot cereals, 3rd tertile fiber (b_3)	0.000551	0.000660
Cold cereals, 1st quartile fiber (b_4)	0.000091	-0.000746
Cold cereals, 2nd quartile fiber (b_5)	0.000951	0.001748
Cold cereals, 3rd quartile fiber (b_6)	0.002346	0.002819
Cold cereals, 4th quartile fiber (b_7)	0.003411	0.003646
Milk (b_8)	0.000070	0.000111
Soda (b_9)	---	-0.000032
100% Fruit juices (b_{10})	0.000077	0.000103
Sugar added to coffee/tea (b_{11})	---	---
Fruitades/sports drinks (b_{12})	0.000053	-0.000023
Fruit (b_{13})	0.000693	0.000715

Parameter	Men	Women
Salad (b_{14})	-0.000156	---
Fried potatoes (b_{15})	0.000672	0.000904
Other potatoes (b_{16})	0.000400	0.000239
Dried beans (b_{17})	0.002147	0.002593
Brown rice (b_{18})	0.001032	0.001046
Other vegetables (b_{19})	0.000569	0.000597
Salsa (b_{20})	0.002049	0.002128
Pizza (b_{21})	0.000426	0.000336
Tomato sauce (b_{22})	0.001026	0.001066
Cheese (b_{23})	0.000401	0.000745
Whole grain bread (b_{24})	0.000949	0.001257
Candy (b_{25})	0.001882	0.001622
Doughnuts (b_{26})	0.000574	0.000840
Cookies, cake, pie, brownies (b_{27})	0.000603	0.000816
Frozen desserts (b_{28})	0.000320	0.000327
Popcorn (b_{29})	0.003181	0.004668

Table 2- 16 Estimated Regression Coefficients for Foods as Predictors of Quarter Root Calcium (mg) by Sex

For milligrams of **calcium**, estimation of b_0 and $b_1 ... b_{29}$, the model is:

$$E([\text{Calcium}]^{.25}) = b_0 + b_1 N_{FG1} P_1 + b_2 N_{FG2} P_2 + \ldots + b_{29} N_{FG29} P_{29}$$

For quarter root of grams of **calcium**, for each sex, the estimates of the parameters are:

Parameter	Men	Women
Intercept (b_0)	4.607439	4.392113
Hot cereals, 1st tertile fiber (b_1)	0.000364	-0.000237
Hot cereals, 2nd tertile fiber (b_2)	-0.000383	0.000202

DIETARY SCREENER: NHIS 2010

Parameter	Men	Women
Hot cereals, 3rd tertile fiber (b_3)	0.001234	0.001473
Cold cereals, 1st quartile fiber (b_4)	-0.004061	-0.003941
Cold cereals, 2nd quartile fiber (b_5)	-0.001969	-0.000746
Cold cereals, 3rd quartile fiber (b_6)	-0.002168	0.000393
Cold cereals, 4th quartile fiber (b_7)	0.003378	0.005292
Milk (b_8)	0.002110	0.002154
Soda (b_9)	---	-0.000041
100% Fruit juices (b_{10})	0.000476	0.000543
Sugar added to coffee/tea (b_{11})	---	0.000287
Fruitades/sports drinks (b_{12})	0.000143	---
Fruit (b_{13})	0.000375	0.000389
Salad (b_{14})	-0.001216	-0.001232
Fried potatoes (b_{15})	0.000843	0.001641
Other potatoes (b_{16})	---	-0.000375
Dried beans (b_{17})	0.001151	0.001799
Brown rice (b_{18})	---	---
Other vegetables (b_{19})	0.000890	0.000869
Salsa (b_{20})	0.004009	0.004581
Pizza (b_{21})	0.002236	0.002744
Tomato sauce (b_{22})	0.001162	---
Cheese (b_{23})	0.012615	0.014689
Whole Grain Bread (b_{24})	0.000390	0.001564
Candy (b_{25})	0.002980	0.003592
Doughnuts (b_{26})	0.000407	---
Cookies, cake, pie, brownies (b_{27})	0.000576	0.001495
Frozen desserts (b_{28})	0.002096	0.002727

Parameter	Men	Women
Popcorn (b_{29})	---	---

3. 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 and vegetables, dairy, added sugars from sugar-sweetened beverages, and whole grains, use the square-root transformation. For added sugars, use the cube-root transformation. For fiber and calcium, use the quarter-root transformation. After analyses, the result variables can be back-transformed for easier interpretation.

The computed diet variables for the NHIS CCS 2010 DSQ are available for download in two formats: SAS transport and comma-separated values. The files include the following variables:

- **HHX** - Household identifier
- **FMX** - Family serial number
- **FPX** - Person number (in family)
- **FCE** - Cup equivalents of fruits per day
- **VCE** - Cup equivalents of vegetables per day (including legumes and French fries)
- **FVCE** - Cup equivalents of fruits and vegetables (including legumes and French fries) per day
- **FVCENoFF** - Cup equivalents of fruits and vegetables (including legumes and excluding French fries) per day
- **Fiber** - Fiber (g) per day
- **Dairy** - Cup equivalents of dairy per day
- **Calcium** - Calcium (mg) per day
- **Sugar** - Added sugars (tsp) per day
- **SSB** - Added sugars (tsp) from sugar-sweetened beverages
- **WHGrain**- Ounce equivalents of whole grains per day

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

DIETARY SCREENER: NHIS 2010

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 (CSV) File (nhis2010dietvars.02-05-2013) - This zip file contains the comma-separated values file, which includes 13 variables, 27,157 records, and an additional record for the variable names.
- SAS Transport File (nhis2010dietvars.02-05-2013.v9x) - The SAS transport file includes 27,157 records and 13 variables. To access the SAS dataset, unzip the file, then use proc cimport.

  For example:

  ```
  proc cimport file='nhis2010dietvars.?.v9x' data=dietvars'
  ```

- SAS Program (create.nhis2010dietvars.02-05-2013) - This is the SAS program that created the dataset. This is for reference, only; there should be no need to run this program.
- Supplement Excel Files (nhis2010dietvars.excel.02-05-2013) - This zip file contains 4 Excel files that are required to run the SAS program.

APPENDIX

Data Processing and Scoring Procedures

Two different data processing and scoring procedures have been developed for the NHANES 2009-10 Dietary Screener Questionnaire (DSQ). The earlier procedures, developed prior to release of the NHANES 2009-10 dietary data, were based on NHANES 2003-2006 24-hour dietary recall (24HR) data. Because the DSQ was not administered in the 2003-2006 NHANES, a pseudoscreener was constructed using frequency information from the 24HR and then related to dietary intake. The content on those web pages has been archived, including the section reproduced in this appendix, and can now be viewed in PDF format in its original context on the NCI's Short Dietary Assessment Instruments website.

The current procedures are based on NHANES 2009-10 24HR and DSQ data, and directly model the relationship between DSQ responses and 24HR intake. This set of procedures is preferable for new analyses of DSQ data.

A. Converting Frequency Data to Daily Frequency

Frequency information on the DSQ in NHANES 2009-2010 was collected using a rate and time unit (e.g., 3 times per week). A different response format consisting of a set number of frequency categories is used in the self-administered paper questionnaire. In both cases, the frequency responses are converted to a common unit of time, i.e. times per day.

For frequency category responses, we used the following conversions. Note that the frequency categories differ somewhat between foods and beverages.

Table A_ 1 Converting frequency data to daily frequencies

Frequency Category	Daily Frequency	
	Foods	Beverages
Never	0	0
1 time last month	0.033	0.033
2-3 times last month	0.083	0.083
1 time per week	0.143	0.143
2 times per week	0.286	0.286
3-4 times per week	0.5	0.5

Frequency Category	Daily Frequency	
	Foods	Beverages
5-6 times per week	0.786	0.786
1 time per day	1	1
2 or more times per day	2	-
2-3 times per day	-	2.5
4-5 times per day	-	4.5
6 or more times per day	-	6

If using the rate and time unit format, we used the following procedures:

Table A_ 2 Converting frequency data to daily frequency using rate and time unit format

Time Period Reported	Daily Frequency
Day	As reported
Week	Reported frequency divided by 7
Month	Reported frequency divided by 30

B. Identifying Extreme Exposure Values

There are various perspectives on whether to exclude potentially unlikely exposure values. If the researcher chooses to do so, several approaches exist for identifying extreme values. We examined the plausibility of the reported frequencies for each food item in the NHANES 2009-10 DSQ and chose to identify extreme values using a method that identifies them based on their actual distribution in the sample, but also minimizes the number of values identified.

Because all items have a real probability of being consumed zero or a small number of times, we focused on the higher ends of the distribution, and examined the times per day distributions. For each food item, we identified discontinuous points of the distributions. We defined the highest frequency before the discontinuity as the maximum acceptable value. Values above this maximum were top-coded (not excluded).

Table B_ 1 Maximum acceptable values in NHANES 2009-10

Food Group	NHANES 2009-10 (N=8541)	
	Maximum Acceptable Daily Frequency Value	Number of Identified Values
Fruit	8	4
Fruit juice	8	8
Salad	5	0
Fried potatoes	5	1
Other potatoes	3	2
Dried beans	4	0
Other vegetables	5	0
Tomato sauce	2	1
Salsa	3	2
Pizza	2	0
Soda	8	23
Fruitades/sports drinks	7	6
Cookies, cake, pie	7	0
Doughnuts	5	0
Ice cream	5	0
Sugar/honey in coffee/tea	10	8
Candy	8	4
Any milk (not soy)	10	12
Cheese	6	2
Any cereal	7	1
Whole grain bread	6	2
Cooked whole grains	4	1
Popcorn	3	1
Red meat	6	2
Processed meat	4	1

Because the NHANES is a national representative survey and this screener was administered to children aged 2 years through adults aged 69 years, we expect this guidance to be appropriate for most U.S. populations.

C. Classifying Cereal Data

DSQ - CURRENT METHODS

The DSQ includes questions about cereal intake and allows respondents to choose up to two responses from a listing of cereals as to which they consumed most frequently. A total of 391 cereal names, including generic or not further specified ones, are on the listing, and 283 are distinct entities.

We classified each of the 283 distinct cereals listed along four dimensions: density of added sugars, whole grains, fiber, and calcium. In the scoring algorithms, all categories have different regression coefficients that relate them to the various dietary factors.

Processing of the cereal data consisted of:

1. **Identification of distinct cereals.** Determine cereal categories based on nutrient density. For each nutrient, i.e. added sugars, whole grains, fiber, and calcium, we ordered listed cereals by density (nutrient/100 grams). We then divided each distribution into tertiles. Note: this density categorization was based on the cereal composition and not the absolute frequency of reported consumption.
2. **Application of this classification to all listed cereals.** Thus, each cereal listed is coded along the following attributes: category for added sugars; category for whole grains; category for fiber; and category for calcium.
3. **Weighting each cereal frequency according to order of report.** For those respondents who reported two different cereal types, we assumed that the first cereal reported was the most frequently consumed and the second was less frequently consumed. Accordingly, we weighted the first cereal at 0.75 and the second at 0.25. For those who reported only one cereal type, no weighting was necessary.

Following are the classification criteria for cereals by nutrient density. Note that any given cereal may fall into different tertiles for different nutrients.

Table C_ 1 Classification Criteria for Hot and Cold Cereals with Regard to Added Sugars Density

Cereal	Density (tsp added sugars/100 grams)	No. of cereals
Lowest tertile added sugars	≤0.71	94
Second tertile added sugars	0.72 - 5.49	95
Highest tertile added sugars	>5.49	94

Table C_ 2 Classification Criteria for Cereals with Regard to Whole Grain Density

Cereal	Density (ounce-equivalents of whole grains/100 grams)	No. of cereals
Lowest tertile whole grain	≤0.21	94
Second tertile whole grain	0.22 - 1.40	94
Highest tertile whole grain	>1.40	95

Table C_ 3 Classification Criteria for Cereals with Regard to Fiber Density

Cereal	Density (grams of fiber/100 grams)	No. of cereals
Lowest tertile fiber	≤2.1	92
Second tertile fiber	2.2 - 7.3	96
Highest tertile fiber	>7.3	95

Table C_ 4 Classification Criteria for Hot and Cold Cereals with Regard to Calcium Density

Cereal	Density (milligrams of calcium/100 grams)	No. of cereals
Lowest tertile calcium	≤21	95
Second tertile calcium	22-100	94
Highest tertile calcium	>100	94

Cereal data can be reported in different formats. The format used in the NHANES DSQ was 8-digit FNDDS food codes, whereas cereal data reported on the self-administered web version is in a character (i.e. letters) format. Table 5 (XLSX) in the SAS folder on NCI's Short Dietary Assessment Instruments website shows the food codes and attributes for the cereals reported in both formats.

DSQ - EARLIER METHODS

The DSQ includes questions about cereal intake and allows respondents up to two responses on which cereals they consume. We classified each cereal reported first by hot or cold, and then along four dimensions: density of added sugars, whole grains, fiber, and calcium.

In the scoring algorithms, hot and cold cereals have different portion sizes. In addition, all categories have different regression coefficients that relate them to the various exposures. Initial coding of the cereal data must:

1. Classify each by hot or cold.
2. Classify each by nutrient density (nutrient/100 grams). We classified all hot and cold cereals reported on the DSQ in the NHANES 2009-10 into categories based on their nutrient density of four exposures: added sugars, whole grains, fiber, and calcium. The density categorization was based on the cereal composition and not the absolute frequency of reported consumption. Thus, each cereal reported is coded along the following attributes: hot or cold; category for added sugars; category for whole grains; category for fiber; and category for calcium.
3. Weight each according to order of report. For those respondents who reported two different cereal types, we assumed that the first cereal reported was the most frequently consumed and the second was less frequently consumed. Accordingly, we weighted the first cereal at 0.75 and the second at 0.25. For those who reported only

one cereal type, no weighting was necessary. These weights were applied to all relevant cereal categories in the scoring algorithms.

Following are the classification criteria for cereal by hot or cold and by nutrient density. Note that each cereal may fall in different n-tiles for different nutrients.

Table C_ 5 Classification Criteria for Hot and Cold Cereals with Regard to Added Sugars Density

Type of Cereal	Density (tsp added sugars/100 grams)	No. of cereals
Hot cereals: no added sugars	0	33
Hot cereals: any added sugars	> 0	11
Cold cereals: 1st quartile added sugars	< 3.0	71
Cold cereals: 2nd quartile added sugars	3.01 to 5.25	68
Cold cereals: 3rd quartile added sugars	5.26 - 9.0	71
Cold cereals: 4th quartile added sugars	> 9.0	69

Table C_ 6 Classification Criteria for Hot and Cold Cereals with Regard to Whole Grain Density

Type of Cereal	Density (ounce-equivalents of whole grains/100 grams)	No. of cereals
Hot cereals: no whole grain	0	12
Hot cereals: any whole grain	> 0	32
Cold cereals: no whole grain	0	88
Cold cereals: 2nd quartile whole grain	0.06 - 1.077	51

Type of Cereal	Density (ounce-equivalents of whole grains/100 grams)	No. of cereals
Cold cereals: 3rd quartile whole grain	1.07705 - 1.82	69
Cold cereals: 4th quartile whole grain	> 1.82	71

Table C_ 7 Classification Criteria for Hot and Cold Cereals with Regard to Fiber Density

Type of Cereal	Density (grams of fiber/100 grams)	No. of cereals
Hot cereals: 1st tertile fiber	< 1.0	15
Hot cereals: 2nd tertile fiber	1.01 - 1.7	13
Hot cereals: 3rd tertile fiber	> 1.7	16
Cold cereals: 1st quartile fiber	< 3.1	69
Cold cereals: 2nd quartile fiber	3.103 - 5.99	69
Cold cereals: 3rd quartile fiber	6.0 - 9.99	72
Cold cereals: 4th quartile fiber	> 10.0	69

Table C_ 8 Classification Criteria for Hot and Cold Cereals with Regard to Calcium Density

Type of Cereal	Density (milligrams of calcium/100 grams)	No. of cereals
Hot cereals: 1st tertile calcium	< 9.99	18
Hot cereals: 2nd tertile calcium	10-64	11
Hot cereals: 3rd tertile calcium	> 65	15

Type of Cereal	Density (milligrams of calcium/100 grams)	No. of cereals
Cold cereals: 1st quartile calcium	< 24.9	70
Cold cereals: 2nd quartile calcium	25.0 - 46.99	70
Cold cereals: 3rd quartile calcium	47.0 - 332.99	68
Cold cereals: 4th quartile calcium	> 333	71

Cereal data can be reported in different formats. The format used in the NHANES DSQ was 8-digit FNDDS food codes. Table 5 (XLSX/CSV) in the SAS folder shows the food codes and attributes for the cereals reported in the 2009-10 NHANES.

Cereal data reported on the self-administered web is in a character (i.e. letters) format. Table 6 (XLSX/CSV) in the SAS folder shows the food codes and attributes for the cereal choices in the self-administered web version of the questionnaire.

D.Developing Scoring Algorithms

DSQ - CURRENT METHODS

We developed scoring procedures to convert screener responses to estimates of individual dietary intake for fruits and vegetables (cup equivalents), dairy (cup equivalents), added sugars (tsp), whole grains (ounce equivalents), fiber (g), and calcium (mg) using the DSQ and What We Eat in America 24-hour dietary recall data from the 2009-2010 NHANES. The following equations were estimated in the NHANES 2009-2010.

For cup equivalents of fruits and vegetables including legumes and French fries

For cup equivalents of fruits and vegetables including legumes and excluding French fries

For cup equivalents of fruits

For cup equivalents of vegetables including legumes and French fries

For cup equivalents of vegetables including legumes and excluding French fries

For cup equivalents of dairy

For teaspoon equivalents of added sugars

For teaspoon equivalents of added sugars from sugar-sweetened beverages

For ounce-equivalents of whole grains

For grams of fiber

For milligrams of calcium

For cup equivalents of total fruits and vegetables including legumes and French fries

For each sex:
E (Fruits and Vegetables) = $b_0 + b_1Kid + b_2Teen + b_3N_{FG3}P_3 + b_4N_{FG4}P_4 + ... + b_{12}N_{FG12}P_{12}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in fruit and vegetable cup equivalents of group k from the 24HR; and k indexes the ten fruit and vegetable food groups.

For cup equivalents of fruits and vegetables including legumes and excluding French fries

For each sex:
E (Fruits and Vegetables, no fries) = $b_0 + b_1Kid + b_2Teen + b_3N_{FG3}P_3 + b_4N_{FG4}P_4 + b_5N_{FG5}P_5 + b_7N_{FG7}P_7 + ... + b_{12}N_{FG12}P_{12}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;

- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in fruit and vegetable cup equivalents of group k from the 24HR; and k indexes the nine fruit and vegetable food groups.

For cup equivalents of fruits

For each sex:

$E\ (\text{Fruits}) = b_0 + b_1\text{Kid} + b_2\text{Teen} + b_3N_{FG3}P_3 + b_4N_{FG4}P_4$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in fruit cup equivalents of group k from the 24HR; and k indexes the two fruit food groups.

For cup equivalents of vegetables including French fries and legumes

For each sex:

$E\ (\text{Vegetables}) = b_0 + b_1\text{Kid} + b_2\text{Teen} + b_5N_{FG5}P_5 + b_6N_{FG6}P_6 + \ldots + b_{12}N_{FG12}P_{12}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;

- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in vegetable cup equivalents of group k from the 24HR; and k indexes the eight vegetable food groups.

For cup equivalents of vegetables including legumes and excluding French fries

$E\ (\text{Vegetables, no FF}) = b_0 + b_1\text{Kid} + b_2\text{Teen} + b_5N_{FG5}P_5 + b_7N_{FG7}P_7 + b_8N_{FG8}P_8 + \ldots + b_{12}N_{FG12}P_{12}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in vegetable cup equivalents of group k from the 24HR; and k indexes the seven vegetable food groups.

For cup equivalents of dairy

For each sex:
$E\ (\text{Dairy}) = b_0 + b_1\text{Kid} + b_2\text{Teen} + b_8N_{FG8}P_8 + b_{13}N_{FG13}P_{13} + b_{14}N_{FG14}P_{14} + b_{15}N_{FG15}P_{15}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in dairy cup equivalents of group k from the 24HR; and k indexes the four dairy food groups.

For teaspoon equivalents of added sugars

For each sex:

$E\text{ (Added sugars)} = b_0 + b_1Kid + b_2Teen + b_{15}N_{FG15}P_{15} + b_{16}N_{FG16}P_{16} + ... + b_{24}N_{FG24}P_{24}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in teaspoon equivalents of group k from the 24HR; and k indexes the ten added sugar food groups.

For teaspoon equivalents of added sugars from sugar-sweetened beverages

For each sex:

$E\text{ (Added sugars)} = b_0 + b_1Kid + b_2Teen + b_{16}N_{FG16}P_{16} + b_{17}N_{FG17}P_{17} + b_{18}N_{FG18}P_{18}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in teaspoon equivalents of group k from the 24HR; and k indexes the three added sugar from beverages food groups.

For ounce-equivalents of whole grains

For each sex:
E (Whole grains) = $b_0 + b_1Kid + b_2Teen + b_{25}N_{FG25}P_{25} + b_{26}N_{FG26}P_{26} + ... + b_{30}N_{FG30}P_{30}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the average daily frequency an individual reported consumption of food group k on the DSQ; P_k is the median portion size in grams of group k from the 24HR; and k indexes the six whole grain food groups.

For grams of fiber

For each sex:
E (Fiber) = $b_0 + b_1Kid + b_2Teen + b_3N_{FG3}P_3 + b_4N_{FG4}P_4 + ... + b_{21}N_{FG21}P_{21} + b_{28}N_{FG28}P_{28} + b_{29}N_{FG29}P_{29} + ... + b_{33}N_{FG33}P_{33}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the usual average daily frequency number of times per day an individual consumed reported consumption of food group k on the DSQ; P_k is the median portion size in grams of group k from the 24HR; and k indexes the twenty-five food groups.

For milligrams of calcium

For each sex:

$E\ (\text{Calcium}) = b_0 + b_1\text{Kid} + b_2\text{Teen} + b_3N_{FG3}P_3 + b_4N_{FG4}P_4 + \ldots + b_{21}N_{FG21}P_{21} + b_{28}N_{FG28}P_{28} + b_{29}N_{FG29}P_{29} + b_{30}N_{FG30}P_{30} + b_{34}N_{FG34}P_{34} + b_{35}N_{FG35}P_{35} + b_{36}N_{FG36}P_{36}$

$b_{29}N_{FG29}P_{29} + \ldots + b_{33}N_{FG33}P_{33}$

- Where E is the expected value;
- N_{FGk} is the daily frequency of intake of Food Group k;
- P_k is the sex-age specific portion size of Food Group k;
- Kid is 1 if aged 2-11 years and otherwise is 0;
- Teen is 1 if aged 12-17 years and otherwise is 0;
- b is the estimated regression coefficient for each term.

The dependent variable is estimated from modelled 24HR data and is the exposure of interest; N_{FGk} is the usual average daily frequency number of times per day an individual consumed reported consumption of food group k on the DSQ; P_k is the median portion size in grams of group k from the 24HR; and k indexes the twenty-five food groups.

DSQ - EARLIER METHODS

We developed scoring procedures to convert screener responses to estimates of individual dietary intake for fruits and vegetables (cup equivalents), dairy (cup equivalents), added sugars (tsp), whole grains (ounce equivalents), fiber (g), and calcium (mg) using the What We Eat in America 24-hour dietary recall data from the 2003-2006 NHANES. The following equations were estimated in the NHANES 2003-2006, using SAS PROC REG.

For cup equivalents of fruits and vegetables

For cup equivalents of dairy

For teaspoons of added sugars, not including cereals

For teaspoons of added sugars including cereals

For teaspoons of added sugars from sugar-sweetened beverages

For ounce-equivalents of whole grains

For grams of fiber

For milligrams of calcium

For cup equivalents of fruits and vegetables

$$E([\text{Fruits and Veg}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG10}P_{10}]^{1/2})$$

Cup equivalents of fruits and vegetables were 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 in fruit and vegetable cup equivalents of group k; and k indexes the ten fruit and vegetable food groups. We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex and age group (ages 2-17 years; 18+), and excluding extreme exposure values.

For cup equivalents of dairy

$$E([\text{Dairy}]^{1/2}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + \ldots + N_{FG4}P_4]^{1/2})$$

Cup equivalents of dairy were 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 in dairy cup equivalents of group k; and k indexes the four dairy food groups. We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 on the NHANES 2003-2006 samples, stratifying by sex and age group (ages 2-17 years; 18+), and excluding extreme exposure values.

For teaspoons of added sugars, not including cereals

$$E([\text{Added Sugars}]^{.33}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + ... + N_{FG7}P_7]^{.33})$$

Teaspoons of added sugars were cube-root-transformed to approximate normality; N_{FGk} is the usual number of times per day that an individual consumed food group k; P_k is the median portion size in added sugars teaspoons of group k; and k indexes the six or seven added sugars food groups.

The item sugar in coffee/tea was not asked of children younger than age 12 years in NHANES. Thus, separate algorithms, not including the sugar in coffee/tea item, were estimated for children less than age 12 years.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex and age group (ages 2-11 years; 12-17; 18+), and excluding extreme exposure values.

Users of the DSQ may prefer to include the sugar in coffee/tea item for children. If so, another algorithm is available that provides regression coefficients for all children and adolescents (ages 2-17 years).

For teaspoons of added sugars including cereals

$$E([\text{Added Sugars}]^{.33}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + ... + N_{FG13}P_{13}]^{.33})$$

Teaspoons of added sugars were cube-root-transformed to approximate normality; N_{FGk} is the usual number of times per day that an individual consumed food group k; P_k is the median portion size in added sugars teaspoons of group k; and k indexes the twelve or thirteen added sugars food groups: hot cereals (2 groups), cold cereals (4 groups), and the six or seven added sugars food groups.

The item sugar in coffee/tea was not asked of children younger than age 12 years in NHANES. Thus, separate algorithms, not including the sugar in coffee/tea item, were estimated for children younger than age 12 years.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex and age group (ages 2-11 years; 12-17; 18+), and excluding extreme exposure values.

Users of the DSQ may prefer to include the sugar in coffee/tea item for children. If so, another algorithm is available that provides regression coefficients for all children and adolescents (ages 2-17 years).

For teaspoons of added sugars from sugar-sweetened beverages

$$E([\text{Added Sugars}_{ssb}]^{.5}) = b_0 + b_1([N_{FG1}P_1 + N_{FG2}P_2 + N_{FG3}P_3]^{.5})$$

The dependent variable added sugars from sugar-sweetened beverages included the following beverages: sodas, fruitades/sports drinks, and sugar in coffee/tea. Teaspoons of added sugars were square-root-transformed to approximate normality; N_{FGk} is the usual number of times per day that an individual consumed food group k; P_k is the median portion size in added sugars teaspoons of group k; and k indexes the two or three sugar sweetened beverage food groups.

The item sugar in coffee/tea was not asked of children younger than age 12 years in NHANES. Thus, separate algorithms, not including the sugar in coffee/tea item, were estimated for children less than age 12 years.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_1 in the NHANES 2003-2006 samples, stratifying by sex and age group (ages 2-11 years; 12-17; 18+), and excluding extreme exposure values.

Users of the DSQ may prefer to include the sugar in coffee/tea item for children. If so, another algorithm is available that provides regression coefficients for all children and adolescents (ages 2-17 years).

For ounce-equivalents of whole grains

$$E([\text{Whole Grains}]^{1/2}) = b_0 + b_1(N_{FG1}P_1) + b_2(N_{FG2}P_2) + \ldots + b_9(N_{FG9}P_9)$$

Ounce-equivalents of whole grains were 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 in grams of group k; and k indexes the nine whole grain food groups: hot cereals (2 groups), cold cereals (4 groups), brown rice, whole grain bread, and popcorn.

We calculated weighted least-squares estimates of the regression coefficients b_0 and bk, k = 0, ..., 9 on the NHANES 2003-2006 samples, stratifying by sex and age group (ages 2-17 years; 18+), and excluding extreme exposure values.

For grams of fiber

$$E([\text{Fiber}]^{1/4}) = b_0 + b_1(N_{FG1}P_1) + b_2(N_{FG2}P_2) + \ldots + b_{29}(N_{FG29}P_{29})$$

Grams of fiber were quarter-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 in grams of group k; and k indexes the 29 food groups, which includes three hot cereal and four cold cereal variables. Two food groups, tomato sauce and dried beans, were frequently consumed in mixed dishes. Because we wanted to represent only the particular food, we estimated the number of grams of tomatoes and legumes, respectively, in these mixed dishes. For tomato sauce, this was done by multiplying the MPED 1 cup equivalent by 171 (1 MPED = 171 grams). For legumes, this was done by estimating the MPED one cup equivalent for legumes and multiplying by 212 (1 MPED = 212 grams).

The item sugar in coffee/tea was not asked of children less than 12 years of age in the NHANES. Thus, separate algorithms, not including the sugar in coffee/tea item, were estimated for children less than 12 years.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_k, k = 0, ... , 29 (or 28, when sugar in coffee/tea was not asked), stratifying by sex and age group (ages 2-11 years, 12-17 years; 18+), and excluding extreme exposure values. We first included all 29 (28) food groups in the regression model. After examining the results, we dropped food groups that failed to attain statistical significance at $\alpha = 0.25$ to form more parsimonious final models. In the fiber model, milk, fruitades/sports drinks, and salad were dropped for boys ages 2-11 years; sugar in coffee/tea was dropped for boys ages 12-17 years; soda and sugar in coffee/tea were dropped for men; brown rice, milk, ice cream, fruitades/sports drinks, candy, and salad were dropped for girls ages 2-11 years; sugar in coffee/tea was dropped for girls ages 12-17 years; and salad and sugar in coffee/tea were dropped for women.

Users of the DSQ may prefer to include the sugar in coffee/tea item for children. If so, another algorithm is available that provides regression coefficients for all children and adolescents (ages 2-17 years). For children and adolescents combined, salad and candy were excluded for both boys and girls.

For milligrams of calcium

$$E([\text{Calcium}]^{1/4}) = b_0 + b_1(N_{FG1}P_1) + b_2(N_{FG2}P_2) + \ldots + b_{29}(N_{FG29}P_{29})$$

Milligrams of calcium were quarter-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 in grams of group k; and k indexes the 29 food groups, which includes three hot cereal and four cold cereal variables. Two food groups, tomato sauce and dried beans, were frequently consumed in mixed dishes. Because we wanted to represent only the particular food, we estimated the number of grams of tomatoes and legumes, respectively, in these mixed dishes. For tomato sauce, this was done by multiplying the MPED 1 cup equivalent by 171 (1 MPED = 171 grams). For legumes, this was done by estimating the MPED one cup equivalent for legumes and multiplying by 212 (1 MPED = 212 grams).

The item sugar in coffee/tea was not asked of children less than 12 years of age in the NHANES. Thus, separate algorithms, not including the sugar in coffee/tea item, were estimated for children less than 12 years.

We calculated weighted least-squares estimates of the regression coefficients b_0 and b_k, k = 0, ... , 29 (or 28) on the NHANES 2003-2006 samples, stratifying by sex and age group (ages 2-11 years, 12-17 years; 18+), and excluding extreme exposure values. We first included all 29 (28) food groups in the regression model. After examining the results, we dropped food groups that failed to attain statistical significance at $\alpha = 0.25$ to form more parsimonious final models. In the calcium model, doughnuts, salad, and French fries were dropped for boys ages 2-11; sugar in coffee/tea, fruit, and other potatoes were dropped for boys ages

12-17; popcorn, brown rice, soda, sugar in coffee/tea, and other potatoes were dropped for men; fruitades/sports drinks, candy, fruit, other potatoes, and cookies/cake/pie were dropped for girls ages 2-11; popcorn, candy, other potatoes, French fries, fruitades/sports drinks, and tomato sauce were dropped for girls ages 12-17; and popcorn, brown rice, fruitades/sports drinks, doughnuts, and tomato sauce were dropped for women.

Users of the DSQ may prefer to include the sugar in coffee/tea item for children. If so, another algorithm is available that provides regression coefficients for all children and adolescents (ages 2-17 years). For children and adolescents combined, popcorn, sugar in coffee/tea, fruit, and other potatoes were excluded for boys; popcorn, doughnuts, candy, other potatoes, and fruitades/sports drinks were excluded for girls.

E. Computed Variables

DSQ - CURRENT METHODS

The computed dietary variables for the Dietary Screener Questionnaire in NHANES 2009-10 are available for download from the NCI's Short Dietary Assessment Instruments website in two formats -- SAS transport and comma-separated values (CSV). The files include the following variables:

- **SEQN** - Unique individual identifier

Variables to estimate mean intakes:

- **DSQfvl** - Predicted intake of fruits and vegetables including legumes and French fries (cup equivalents) per day
- **DSQfvlnf** - Predicted intake of fruits and vegetables including legumes and excluding French fries (cup equivalents) per day
- **DSQfrt** - Predicted intake of fruits (cup equivalents) per day
- **DSQvlall** - Predicted intake of vegetables including legumes and French fries (cup equivalents) per day
- **DSQvlnf** - Predicted intake of vegetables including legumes and excluding French fries (cup equivalents) per day
- **DSQdairy** - Predicted intake of dairy (cup equivalents) per day
- **DSQsug** - Predicted intake of total added sugars (tsp equivalents) per day
- **DSQssb** - Predicted intake of added sugars from sugar-sweetened beverages (tsp equivalents) per day

- **DSQwhgr** - Predicted intake of whole grains (ounce equivalents) per day
- **DSQfib** - Predicted intake of fiber (gm) per day
- **DSQcalc** - Predicted intake of calcium (mg) per day

Variables to estimate low threshold prevalence:

- **DSQfvl_low** - Predicted probability of eating less than 1.7 cup equivalents of fruits and vegetables including legumes and French fries
- **DSQfvlnf_low** - Predicted probability of eating less than 1.7 cup equivalents of fruits and vegetables including legumes and excluding French fries
- **DSQfrt_low** - Predicted probability of eating less than 0.5 cup equivalents of fruits
- **DSQvlall_low** - Predicted probability of eating less than 1.0 cup equivalents of vegetables including legumes and French fries
- **DSQvlnf_low** - Predicted probability of eating less than 1.0 cup equivalents of vegetables including legumes and excluding French fries
- **DSQdairy_low** - Predicted probability of eating less than 1.2 cup equivalents of dairy
- **DSQsug_low** - Predicted probability of eating less than 11 tsp equivalents of total added sugars
- **DSQssb_low** - Predicted probability of eating less than 3 tsp equivalents of added sugars from sugar-sweetened beverages
- **DSQwhgr_low** - Predicted probability of eating less than 0.3 ounce equivalents of whole grains
- **DSQfib_low** - Predicted probability of eating less than 12 grams of fiber
- **DSQcalc_low** - Predicted probability of eating less than 800 milligrams of calcium

Variables to estimate high threshold prevalence:

- **DSQfvl_high** - Predicted probability of eating equal or more than 3.2 cup equivalents of fruits and vegetables including legumes and French fries
- **DSQfvlnf_high** - Predicted probability of eating equal or more than 3.2 cup equivalents of fruits and vegetables including legumes and excluding French fries
- **DSQfrt_high** - Predicted probability of eating equal or more than 1.4 cup equivalents of fruits
- **DSQvall_hig**h - Predicted probability of eating equal or more than cup equivalents of 1.8 of vegetables including legumes and French fries
- **DSQvlnf_high** - Predicted probability of eating equal or more than 1.8 cup equivalents of vegetables including legumes and excluding French fries
- **DSQdairy_high** - Predicted probability of eating equal or more than 2.4 cup equivalents of dairy
- **DSQsug_high** - Predicted probability of eating equal or more than 23 tsp equivalents of total added sugars
- **DSQssb_high** - Predicted probability of eating equal or more than 11 tsp equivalents of added sugars from sugar-sweetened beverages
- **DSQwhgr_high** - Predicted probability of eating equal or more than 1.0 ounce equivalents of whole grains

- **DSQfib_high** - Predicted probability of eating equal or more than 19 grams of fiber (gm)
- **DSQcalc_high** - Predicted probability of eating equal or more than 1100 milligrams of calcium (mg)

The datasets are sorted in ascending order by the ID variable SEQN.

- Comma-separated Values File [CSV]
  This file contains the comma-separated values file, which includes 34 variables, 7,861 records, and an additional record for the variable names.
- SAS Transport File [XPT]
  The SAS transport file includes 7,861 records and 34 variables. To access the SAS dataset, unzip the file, then use proc cimport.

  For example:
  ```
  proc cimport file='nh0910.DSQvars.xpt' data=dietvars;
  ```
- Supplement Excel Files [ZIP]
  This zip file contains 3 Excel files that are required to run the SAS program.
- Note that 680 respondents did not answer the DSQ. When one or more values is missing in the estimation of the computed diet variable, that value for that variable will also be missing.

DSQ – EARLIER METHODS

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 and vegetables, dairy, added sugars from sugar sweetened beverages, and whole grains, use the square-root transformation; for added sugars, use the cube-root transformation; for fiber and calcium, use the quarter-root transformation. After analyses, the result variables can be back-transformed for easier interpretation.

The computed diet variables for the Dietary Screener Questionnaire in NHANES 2009-10 for the Earlier Methods are available for download from the NCI's Short Dietary Assessment Instruments website in two formats -- SAS transport and comma-separated values (CSV). The files include the following variables:

- **SEQN** - Unique individual identifier
- **predfib** - Predicted fiber (gm) per day
- **predcalc** - Predicted calcium (mg) per day
- **predsug** - Predicted added sugars (tsp) per day
- **predsugnc** - Predicted added sugars (tsp) per day, not using cereal
- **predwhgr** - Predicted ounce equivalents of whole grains per day
- **preddairy** - Predicted cup equivalents of dairy per day

- **predfvl** - Predicted cup equivalents of fruits and vegetables (including legumes) per day
- **predfvlnf** - Predicted cup equivalents of fruits and vegetables (including legumes) except French fries per day
- **predssb** - Predicted added sugars (tsp) from sugar-sweetened beverages

The datasets are sorted in ascending order by the ID variable SEQN.

- Comma-separated Values File [ZIP]
  This zip file contains the comma-separated values file, which includes 10 variables, 8,541 records, and an additional record for the variable names.
- SAS Transport File [SAS]
  The SAS transport file includes 8,541 records and 10 variables. To access the SAS dataset, unzip the file, then use proc cimport.

  For example:
  proc cimport file='nh0910.dietvars.?.v9x' data=dietvars'
- SAS Program [SAS]
  This is the SAS program that created the dataset. This is for reference, only; there should be no need to run this program.
- Supplement Excel Files [ZIP]
  This zip file contains 3 Excel files that are required to run the SAS program.
- Note that 680 respondents did not answer the DSQ. When one or more values is missing in the estimation of the computed diet variable, that value for that variable will also be missing.