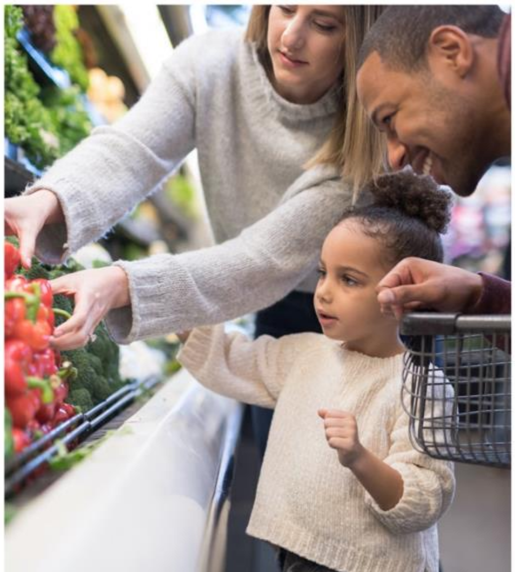
Indicators of Diet Quality, Nutrition, and Health for Americans by Program Participation Status, 2011–2016: SNAP Report

Appendix A. Data and Methods



Indicators of Diet Quality, Nutrition, and Health for Americans by Program Participation Status, 2011–2016: SNAP Report

Appendix A. Data and Methods

December 2021

Authors

Stacy Gleason
Dani Hansen
Breanna Wakar

Submitted to

USDA Food and Nutrition Service,
Office of Policy Support
1320 Braddock Place
Alexandria, VA 22310

Project Officer

Michael Burke

Submitted by

Insight Policy Research, Inc.
1901 North Moore Street
Suite 1100
Arlington, VA 22209

Project Director

Stacy Gleason

Contents

Appendix A. Data and Methods

All estimates in this report are based on NHANES data analyzed alone or in conjunction with various Food Patterns data sources. This appendix provides details on the data; estimation procedures for outcomes examined in this study, and statistical methods employed to produces results presented in the main report and appendices B through G.

A. Data Sources

1. NHANES Data

The NHANES is designed to assess the health and nutritional status of children and adults in the United States. CDC NCHS conducts the survey, which collects both interviews and physical examination data on a nationally representative sample of the U.S. population. The NHANES was administered periodically since 1971 and has been a continuous annual survey since 1999, with public datafiles released every 2 years. Each 2-year survey wave collects data on approximately 10,000 individuals. NCHS recommends combining two or more 2-year survey cycles of the continuous NHANES to increase sample size and produce estimates with greater statistical reliability. All the estimates in this report are based on three 2-year cycles of NHANES data (2011–2012, 2013–2014, and 2015–2016).[1]

NHANES includes a household interview conducted in respondents' homes and a physical examination conducted in MECs. Additional interviews are conducted at the time of the MEC exam, including a 24-hour dietary recall interview (i.e., the day 1 dietary recall). A second dietary recall interview is conducted by telephone 3 to 10 days after the MEC exam (i.e., the day 2 dietary recall). The study team used data from the following NHANES datafiles:

- Body Measures (BMX)
- Demographics Variables and Sample Weights (DEMO)
- Dietary Interview: Individual Foods–Day 1 (DR1IFF)
- Dietary Interview: Individual Foods–Day 2 (DR2IFF)
- Dietary Interview: Total Nutrient Intakes–Day 1 (DR1TOT)
- Dietary Interview: Total Nutrient Intakes–Day 2 (DR2TOT)
- Food Security (FSQ)
- Income (INQ)
- Reproductive Health (RHQ)

All the analyses in this report are based on data from NHANES respondents with complete day 1 dietary recall. To compute all dietary measures, the study team used day 1 dietary recall data. To estimate usual nutrient intakes, the team used day 2 dietary recall data in conjunction with day 1 dietary recall data to control for within-person day-to-day variance in nutrient intakes.

[1] Except serum vitamin D and B12 data in chapter 3, which use 2011–2014 data, 2015–2016 data were not yet published for this biomarker at the time of the analysis.

2. Food Patterns Equivalents Data

FPED (formerly the MyPyramid Equivalents Database) was used to obtain Food Patterns data for each food reported in the NHANES 2011–2016 data. Specifically, FPED provided data (USDA ARS, 2019) on the amounts of more than 30 Food Patterns components included in 100 grams of food (Bowman et al., 2008; Bowman et al., 2013). The Food Patterns components are defined as the number of cup equivalents of fruits, vegetables, and dairy; ounce equivalents of grains and protein foods; and teaspoon equivalents of added sugars. The study team linked each unique food reported in the NHANES 2011–2016 Individual Foods Files to the appropriate Food Patterns data source and computed the amounts of each Food Patterns component consumed based on the amount of food consumed by each individual.

B. Analysis Sample

The sample for most of the descriptive analyses included individuals aged 1 and older with complete dietary recalls, excluding breastfed children, infants (younger than 1 year old), and pregnant and breastfeeding women.[2] The sample for the matched analyses included individuals aged 1 and older.

Children who consumed breast milk were excluded from the descriptive analysis sample because they had incomplete dietary recall data. Infants were excluded for three reasons:

- More than one-third of infants in the NHANES 2011–2016 had incomplete dietary recalls because they consumed breast milk. The records for these infants were missing values for the amounts of calories and nutrients consumed from breast milk because amounts were not quantified by respondents, and it was beyond the scope of this project to impute breast milk volumes.
- Many of the outcome measures used in this study did not apply to infants, such as the 2015–2020 DGA, HEI-2015, and BMI.
- Comparisons of the usual intakes of infants with the DRIs were limited because DRIs were defined for only a few nutrients for infants.

Pregnant and breastfeeding women 20–44 years old were excluded from both the descriptive and matched analysis samples because the dietary reference standards were different for these groups. However, pregnant or breastfeeding women 12–19 years old and aged 44 and older could not be identified in the NHANES 2011–2016 public use data. As a result, women of these ages were assumed to be neither pregnant nor breastfeeding.

C. Subgroups for Tabulation

The study team calculated descriptive estimates for the total U.S. population and for subgroups defined by program participation and income, and by age group and gender.

[2] Several of the outcome measures used in this study did not apply to children younger than 2 years old, such as the HEI, BMI, and percentage of calories from saturated fat and added sugars. Additional outcomes presented in chapter 3 did not apply to children or applied only to older children. The analyses for these measures were limited to the age groups to which they applied.

1. Program Participation and Income

SNAP participation was measured at the household level based on self-reported SNAP benefit receipt. The team defined SNAP participants as respondents who reported living in households currently receiving SNAP benefits, using the NHANES variable. The team identified nonparticipants as respondents who reported they had never received SNAP or that they had not received SNAP benefits in the last 12 months using the NHANES variables FSQ165 (ever received SNAP benefits) and FSQ171 or FSQ012—depending on year (received SNAP benefits in the last 12 months). To classify program nonparticipants as income-eligible or higher income, the team used household size and monthly income relative to the Federal Poverty Guidelines, using the NHANES family poverty income ratio variable INDFMMPC. Income-eligible nonparticipants were defined as individuals with annual income less than or equal to 130 percent of the relevant Federal Poverty Guidelines, whereas higher income nonparticipants were defined as individuals with annual income greater than 130 percent of the relevant Federal Poverty Guidelines, with no income cap.

2. Age Groups

The study team tabulated descriptive analysis results for three age groups:

- Children (1–18 years old)[3]
- Adults (19–59 years old)
- Older adults (aged 60 and older)

Ages were calculated based on age at the time of the MEC exam, when the first dietary recall was collected, rather than age at the time of the household interview.

D. Methods for Estimating Outcome Measures

The study team used several outcome measures to examine the diet quality of SNAP participants and nonparticipants. This section describes the methods used to construct each measure.

1. Usual Nutrient Intakes

To assess the prevalence of adequate nutrient intakes among SNAP participants and nonparticipants, the study team estimated usual nutrient intakes of vitamins, minerals, macronutrients, and other dietary components. The team then compared usual nutrient intake distributions with the DRIs and selected recommendations of the 2015–2020 DGA.

Dietary Reference Intakes

The DRIs, established by NASEM's, formerly IOM,[4] Food and Nutrition Board, provide guidelines on intake amounts appropriate for a given individual based on age, gender, and life stage (IOM, 1997; 1998; 2000; 2001; 2005; 2006; 2011; NASEM, 2019). The DRIs are the most up-to-date scientific standards for determining

[3] Some nutrition outcomes did not apply to children or applied only to older children; children were excluded from estimates when necessary.
[4] Now known as the Health and Medicine Division of the National Academies of Sciences, Engineering, and Medicine

whether diets provide enough nutrients to meet requirements without being excessive. Four different DRI standards were used to assess the usual nutrient intakes of SNAP participants and nonparticipants:

- EAR
- AI level
- CDRR
- AMDR

DRI values for each nutrient included in the analysis are shown in tables A.1 through A.3 for each age and gender group. When enough information is available about the distribution of nutrient requirements in the population, the DRIs define an EAR. The EAR is the average daily nutrient intake level estimated to meet the requirement of half the healthy individuals in a life stage and gender group. The EAR is used to assess the prevalence of inadequate intakes using the NASEM, formerly IOM, recommended "EAR cut-point method" (IOM, 2006, pp. 43–44). The method was used to analyze all nutrients for which EARs had been established. The EAR cut-point method assumes nutrient requirements are symmetrically distributed.[5]

When information on the distribution of requirements is insufficient to establish an EAR, the DRIs define an AI level. The AI is the level of intake assumed to be adequate based on observed or experimentally determined estimates of intake by apparently healthy people. AIs cannot be used to determine the proportion of a population with inadequate intakes. Instead, assessment focuses on comparison of mean usual intakes to an AI level. Populations with a mean usual intake equal to or greater than the population-specific AI can be assumed to have high levels of nutrient adequacy. However, when mean usual intakes fall below the AI, no firm conclusions can be drawn about the prevalence of adequate usual intakes.

The CDRR is the intake above which intake reduction is expected to reduce chronic disease risk within an apparently healthy population. The CDRR replaced the tolerable upper intake level for sodium in 2019 (National Academies of Sciences, Engineering, and Medicine, 2019).

The DRIs also define AMDRs for intakes of macronutrients (total fat, carbohydrate, and protein) and key fatty acids (linoleic acid and linolenic acid). The AMDRs reflect a range of usual nutrient intakes associated with reduced risk of chronic disease while providing adequate intakes of other essential nutrients (IOM, 2000). AMDRs are expressed as percentages of total calorie intake because their requirements are not independent of each other nor of the total calorie requirement of the individual (IOM, 2006). A key feature of AMDRs is that they specify ranges of intake. Intakes that fall outside of these ranges (i.e., exceed the upper bound or fall below the lower bound) may increase risk of chronic disease.

[5] This assumption, however, does not hold for iron requirements among menstruating females. It is not appropriate to use the EAR cut-point method to estimate the prevalence of adequate iron intakes for menstruating females. For this reason, the full probability approach was used for females 9–50 years old (IOM, 2006).

The 2015–2020 DGA also includes quantitative recommendations for saturated fat and added sugar that encourage reduced intakes. The DGA recommendations are shown in table A.4.

Estimating usual nutrient intakes

The DRIs, which are used to assess the prevalence of inadequate and excessive nutrient intakes, are intended to be applied to measures of usual intakes or long-term averages of daily intakes. Therefore, information about the distribution of usual nutrient intakes is needed for assessing diets of population groups. Because individuals' nutrient intakes vary day to day, single 24-hour dietary recalls may lead to biased estimates of the distribution of usual intakes (NIH NCI, n.d.b). A second 24-hour dietary recall for at least a subset of the population allows for estimation of usual intakes (Dodd et al., 2006; Tooze et al., 2006).

Table A.1. Estimated Average Requirement by Age and Gender Groups

Age (Years) and Gender Group	Vitamin A (mcg RAE)	Vitamin C (mg)	Vitamin D (mcg)	Vitamin B6 (mg)	Vitamin B12 (mcg)	Vitamin E (mcg)	Folate (mcg DFE)	Niacin (mg)	Selenium (mcg)	Copper (mg)	Riboflavin (mg)	Thiamin (mg)	Calcium (mg)	Iron (mg)	Magnesium (mg)	Zinc (mg)	Carbohydrate (g)	Protein (g/kg body weight)
Children																		
1–3	210	13	10	0.4	0.7	5	120	5	17	0.26	0.4	0.4	500	3.0	65	2.5	100	0.87
4–8	275	22	10	0.5	1.0	6	160	6	23	0.34	0.5	0.5	800	4.1	110	4.0	100	0.76
Males																		
9–13	445	39	10	0.8	1.5	9	250	9	35	0.54	0.8	0.7	1,100	5.9	200	7.0	100	0.76
14–18	630	63	10	1.1	2.0	12	330	12	45	0.69	1.1	1.0	1,100	7.7	340	8.5	100	0.73
19–30	625	75	10	1.1	2.0	12	320	12	45	0.70	1.1	1.0	800	6.0	330	9.4	100	0.66
31–50	625	75	10	1.1	2.0	12	320	12	45	0.70	1.1	1.0	800	6.0	350	9.4	100	0.66
51–70	625	75	10	1.4	2.0	12	320	12	45	0.70	1.1	1.0	800	6.0	350	9.4	100	0.66
71+	625	75	10	1.4	2.0	12	320	12	45	0.70	1.1	1.0	1,000	6.0	350	9.4	100	0.66
Females																		
9–13	420	39	10	0.8	1.5	9	250	9	35	0.54	0.8	0.7	1,100	5.7	200	7.0	100	0.76
14–18	485	56	10	1.0	2.0	12	330	12	45	0.69	0.9	0.9	1,100	7.9	300	7.3	100	0.71
19–30	500	60	10	1.1	2.0	12	320	12	45	0.70	0.9	0.9	800	8.1	255	6.8	100	0.66
31–50	500	60	10	1.1	2.0	12	320	12	45	0.70	0.9	0.9	800	8.1	265	6.8	100	0.66
51–70	500	60	10	1.3	2.0	12	320	12	45	0.70	0.9	0.9	800	5.0	265	6.8	100	0.66
71+	500	60	10	1.3	2.0	12	320	12	45	0.70	0.9	0.9	1,000	5.0	265	6.8	100	0.66

DFE = dietary folate equivalents; g = grams; kg = kilograms; mcg = micrograms; mg = milligrams; RAE = retinol activity equivalents

Table A.2. Adequate Intake Level and Chronic Disease Risk Reduction Level by Age and Gender Group

Age (Years) and Gender Group	Potassium (mg)	Sodium AI (mg)	Sodium CDRR (mg)	Fiber (g)	Linoleic Acid (g)	Linolenic Acid (g)	Choline (mg)
Children							
1–3	2,000	800	1,200	19	7	0.7	200
4–8	2,300	1,000	1,500	25	10	0.9	250
Males							
9–13	2,500	1,200	1,800	31	12	1.2	375
14–18	3,000	1,500	2,300	38	16	1.6	550
19–30	3,400	1,500	2,300	38	17	1.6	550
31–50	3,400	1,500	2,300	38	17	1.6	550
51–70	3,400	1,500	2,300	30	14	1.6	550
71+	3,400	1,500	2,300	30	14	1.6	550
Females							
9–13	2,300	1,500	1,800	26	10	1.0	375
14–18	2,300	1,500	2,300	26	11	1.1	400
19–30	2,600	1,500	2,300	25	12	1.1	425
31–50	2,600	1,500	2,300	25	12	1.1	425
51–70	2,600	1,500	2,300	21	11	1.1	425
71+	2,600	1,500	2,300	21	11	1.1	425

AI = adequate intake; CDRR = chronic disease risk reduction; g = grams; mg = milligrams

Table A.3. Acceptable Macronutrient Distribution Range by Age Group

Age Group (Years)	Total Fat	Linoleic Acid	Linolenic Acid	Carbohydrate	Protein
Percent of Total Calories					
Children 1–3	30–40	5–10	0.6–1.2	45–65	5–20
Children 4–18	25–35	5–10	0.6–1.2	45–65	10–30
Adults 19+	25–35	5–10	0.6–1.2	45–65	10–35

Table A.4. 2015–2020 DGA for Saturated Fat and Added Sugars by Age Group

Age Group (Years)	Saturated Fat (Percent of Total Calories)	Added Sugars (Percent of Total Calories)
2–20	< 10	< 10
51+	< 10	< 10

DGA = Dietary Guidelines for Americans

The study team used the method developed by NCI to estimate the usual intake distributions, mean intakes, and percentages of individuals above, below, or within the standards established in the DRIs or recommended in the 2015–2020 DGA. The NCI method involves the use of two SAS macros—Distrib and Mixtran—available in NCI's data and fits the model (Parsons et al., 2009). Distrib uses the parameters estimated by Mixtran to estimate the usual intake statistics through simulation. Distrib also provides the estimated percentage of the population whose intake falls below a given value (e.g., a DRI value or DGA recommendation). To estimate standard errors for the estimated percentiles and percentages, the team used the balanced repeated replication method.

2. Usual Intake of Calories

Usual intake of calories was computed using the NCI Mixtran and Distrib SAS macros. Mixtran transforms the data and fits the model used for calculating the estimates. Distrib uses the parameters estimated by Mixtran to calculate the mean and distribution of the variable of interest based on the model established for the population being examined.

3. Body Mass Index

Weight status is defined using BMI, a measure of the relationship between height and weight. BMI is a widely accepted index for classifying the weight status of individuals as underweight, healthy weight, overweight, or obese. The NHANES collects body measurement data during the MEC exam, including body weight and height (or recumbent length for children younger than 2 years old). These data are available in the NHANES Body Measures Files. These files also include a variable for BMI, defined as follows:

$$\text{BMI} = \text{weight in kilograms} \div (\text{height in meters})^2$$

Methods for classifying the weight status of individuals based on BMI differ for adults and children. The study team classified adult weight status using the BMI variable from the NHANES data and the BMI cutoffs specified by CDC (2015). For children, CDC recommends using BMI to screen for overweight and obesity beginning at age 2 (CDC, 2018). Because children grow at different rates at different times, children's weight status is determined by using BMI-for-age percentiles that account for a child's age and gender. The team used the SAS program provided on CDC's website to estimate BMI-for-age percentiles for children. The team classified children's weight status based on comparison of BMI-for-age percentiles with the CDC standards. Children younger than 2 years old and individuals with missing BMI or height and weight data were excluded from the analysis. Table 3.2 in chapter 3 shows reference values for BMI.

4. Other Indicators of Nutrition and Health

Waist circumference, WHtR, and disease risk were assessed for all adults aged 19 and older. Unlike BMI alone, these measures account for abdominal obesity. Waist circumference was categorized into normal risk or high risk based on the cutoffs for men and women specified by CDC (2015). Waist-to-height ratios were assessed compared with the generally accepted cutoff of < 0.5 for normal risk (Ashwell & Gibson, 2016). Disease risk, which takes into account both an individual's BMI category and waist circumference category, was categorized as normal risk, increased risk, high risk, very high risk, or extremely high risk (NIH NHLBI, n.d.b).

Blood pressure measures are taken from examination data and measured for individuals 8 and older. Up to four blood pressure readings are taken during the examination and combined into a single average reading using the following method (CDC NCHS, 2011):

- If only one blood pressure reading is obtained, that reading is the average.
- If more than one blood pressure reading is obtained, the first reading is excluded from the average. This means if only two blood pressure readings are obtained, the second reading is the average.
- If three blood pressure readings are obtained, the average is calculated using the second and third reading.
- Any diastolic readings of zero are excluded from the diastolic average.

Several additional indicators of health and nutrition collected in the NHANES laboratory dataset were utilized. Table A.5 presents these variables and their respective age groups.

Table A.5. Blood Biomarkers and Biochemical Indicators

Measure	Minimum Age Included in Analysis
Blood Biomarkers	
HDL cholesterol	6 years
LDL cholesterol	12 years
TG	12 years
Glucose	12 years
Hemoglobin	1 year
Biochemical Indicators	
Vitamin B12	19 years
Vitamin D	1 year
Iodine	6 years

HDL = high-density lipoprotein; LDL = low-density lipoprotein; TG = triglyceride

5. Calories From Saturated Fats, Added Sugars, and Alcohol

To assess the consumption of calories from saturated fats, added sugars, and alcohol (previously referred to as empty calories), the study team estimated the percentage contribution of each to total calorie intakes. Estimates were based on a single day of intake.[6] Children younger than 2 years old were excluded from the analysis because the 2015–2020 DGA did not apply to them. To construct this measure, the team obtained data on alcohol and total calories from the NHANES Individual Foods Files and Total Nutrients Files and obtained data on saturated fats, added sugars, and alcohol from the FPED data sources (USDA ARS, 2019) described previously. The results display the calories consumed from saturated fat and added sugar as a proportion of total calories among all school children by participation/eligibility status and age group.

6. Food Choices Defined Using the Supermarket Aisle Approach

To examine the food choices of SNAP participants and nonparticipants, the study team categorized all foods reported in day 1 dietary recalls according to the food groups and subgroups defined in the supermarket aisle approach used by Cole and Fox (2008a). This approach categorizes foods into major food groups and subgroups based on supermarket groupings. The team made slight modifications to the food groups defined by Cole and Fox (2008a) to reflect the types of foods reported in the NHANES 2011–2016. Grains were classified as whole grains if at least 50 percent of the total grains were whole grains (using the FPED data sources; see USDA ARS, 2019). Vegetables that were not categorized separately by type were assigned to the "other raw" or "other cooked" vegetable groups. Changes to food groups from the previous round were as follows:

- "Other cooked" and "other raw" vegetables were no longer differentiated based on high versus low nutrient levels.
- Soy milk was removed from the dairy group and instead included in the newly created plant-based beverages group.
- Soups were no longer differentiated as meat soups, grain soups, or vegetable soups but instead categorized as soups/stews.

The study team estimated the percentages of individuals consuming 1 or more foods (in any amount) from the 11 major supermarket aisle food groups on the day covered in the dietary recall. For each supermarket aisle subgroup, the percentage of individuals consuming one or more foods from the subgroup among those who consumed any foods in the corresponding major group was estimated. For example, the percentage of individuals consuming each grain subgroup is conditional on consuming any grains. This approach allows for comparison of food choices among SNAP participants and nonparticipants while controlling for different overall levels of consumption at the major food group level. All the supermarket aisle food groups and subgroups reflect foods consumed as discrete items except when eaten in combination (i.e., the mixed dishes category).

7. Average Amounts of Food Consumed From Supermarket Aisle Food Groups

The study team examined the mean amounts of food consumed by SNAP participants and nonparticipants on the day covered in the dietary recall from each of the major food groups and

[6] The 2015–2020 DGA does not make recommendations for "empty calories." Instead, the DGA describes calories from saturated fats, added sugars, and alcohol, along with refined grains and any other excess calories, as "calories for other uses."

subgroups defined in the supermarket aisle approach. Amounts were estimated in both grams and Food Patterns units among the total population and among consumers only. To construct these measures, the NHANES Individual Foods Files, FPED data, and major food groups and subgroups defined in the supermarket aisle approach were used. To estimate average amounts consumed in grams, gram amounts for foods reported consumed within each food group and subgroup were summed to create daily totals for each individual. To estimate amounts consumed in Food Patterns units, FPED data were used to obtain cup and ounce equivalents data for foods in the milk and milk products, fruits, vegetables, meat and meat alternatives, and grains groups and their associated subgroups. Food Patterns units for each food group and subgroup were summed to create daily totals in cup or ounce equivalents for each individual. For foods that were reported as multiple components but counted as one item in the food choices analysis, the team summed the gram and Food Patterns units for all components reported, so foods were handled the same way in both analyses.

The study team then estimated the mean amounts of grams and Food Patterns units for the total population, which included all individuals regardless of whether the food group or subgroup was consumed. To estimate the average amounts consumed among consumers only, the team included only those individuals who reported consuming the specific food group or subgroup. The estimates reflect average daily amounts of foods consumed on the day covered in the dietary recall. The results for the average amounts of foods consumed from supermarket aisle food groups should not be used to represent total food group intake or compared with USDA Food Patterns recommendations. Total food group intakes for each USDA Food Patterns group were not estimated for this study but were estimated by USDA using NHANES 2011–2011, 2013–2014, and 2015–2016 data. These intakes are available online at the following links:

- https://www.ars.usda.gov/ARSUserFiles/80400530/pdf/fped/Table_1_FPED_GEN_1112.pdf
- https://www.ars.usda.gov/ARSUserFiles/80400530/pdf/fped/Table_1_FPED_GEN_1314.pdf
- https://www.ars.usda.gov/ARSUserFiles/80400530/pdf/fped/Table_1_FPED_GEN_1516.pdf

8. Healthy Eating Index

To estimate mean HEI-2015 and HEI-2010 component and total scores, the study team used the following resources developed by NCI (NIH NCI DCCPS, 2020a) and available on its website:

- Three SAS macros that allocate beans and peas to the protein/meat and beans and vegetables components and apply the HEI scoring algorithm (hei2010.beanspeas.allocation.macro.sas, hei2010.score.macro.v1.2.sas, and hei2015.score.macro.sas)

NCI's SAS programs and macros are designed to estimate mean HEI component and total scores and corresponding standard errors and confidence intervals using 1 day of dietary intake data from the NHANES (NIH NCI, 2020). The SAS code uses SAS survey procedures to account for the complex survey design and a Monte Carlo simulation step to compute standard errors. The SAS programs read in the variables needed from the NHANES Individual Foods Files and Total Nutrient Intakes Files and the variables needed from FPED. The study team adapted NCI's SAS code to calculate HEI scores for the NHANES 2011–2016 and to import the Food Patterns data sources (described previously).

The SAS programs use the population ratio method (rather than using means of individual scores or means of individual ratios) and 1 day of dietary intake data to estimate mean component and total HEI scores. In this method, the ratio between the group's total intake of a food group or nutrient of interest and its total calorie intake is computed. This convention is usually suggested largely because of two factors: (1) it reduces possible bias resulting from correlations between an individual's 1-day food-or-nutrient-to-energy ratio and calorie intake, and (2) there is usually less score truncation in the HEI scoring system for the group-level HEI measure than in the mean of the individual-level HEI scores (Freedman et al., 2008).

E. Statistical Methods

The study team produced all estimates for this report using SAS (version 9.4). Sample weights were used to account for sample design and nonresponse. Information about the NHANES survey design (strata and primary sampling units) was used for estimating variances and testing statistical significance. Therefore, the SAS procedures used included SURVEYREG and SURVEYMEANS.

The NHANES analytic guidelines recommend calculating standard errors using procedures that account for the complex sampling design effect to produce asymptotically unbiased estimates of the variance. Following the NHANES guidelines, the team estimated standard errors using replicate weights that accounted for the complex survey design. Standard errors are included in appendix tables only.

1. Sampling Weights

The study team applied weights reflecting the sampling design of the NHANES to project sample statistics to population statistics. The study team constructed 6-year weights according to the NHANES analytic guidelines because all estimates were based on three cycles of NHANES data.[7] The NHANES provides several weights for use in analyzing each wave of data, including full sample 2-year interview weights, full sample 2-year examination weights, day 1 dietary sample weights, and day 2 dietary sample weights. Because the team limited its analytic sample to NHANES respondents with complete and reliable day 1 dietary recall data, it primarily used the day 1 dietary sample weights. Day 1 weights adjust for the nonresponse in the day 1 dietary recall and the differential allocation by day of the week for the dietary intake data collection. For the usual intakes analysis, which used both day 1 and day 2 dietary recall data, the team also used the day 2 dietary sample weights. This weight incorporated adjustments for the additional nonresponse in the day 2 dietary recall and for the proportion of weekend-weekday combinations of day 1 and day 2 recalls.

2. Age-Adjusted Statistics

The study team used age adjustment to produce descriptive estimates for all ages, children, adults, and older adults, and separately for all individuals, men/boys, and women/girls. For all outcomes except usual nutrient intakes, when adjusting estimates for all individuals, the team used a single weight for everyone in a particular age group rather than separate weights for males and females. For usual nutrient intakes, age adjustment was applied separately for males and females, and age adjustment weights were then used to create the combined group estimates.

Age adjustment eliminates differences between comparison groups resulting solely from differences in the age distributions of the groups. The age-adjusted estimates were calculated as the weighted average of estimates computed for each DRI age group (or portion of DRI age group) using weights equal to the

[7] Except serum vitamin D data in chapter 3, which uses 2011–2014 data, 2015–2016 data were not yet published for this biomarker at the time of the analysis.

proportion of the 2010 U.S. population within each age group. Table A.6 shows the population distribution used for age adjustment. Two approaches were used for age adjustment.

Table A.6. 2010 Census Population for DRI Age Groups

Age Group (years)	Population	Percent
DRI Age Groups		
1–3	12,194,039	4.0
4–8	20,263,474	6.6
9–13	20,659,565	6.8
14–18	21,621,091	7.1
19–30	51,558,750	16.9
31–50	84,115,923	27.6
51–70	68,598,943	22.5
71+	25,789,600	8.5
Other Age Groups		
2–3	8,215,969	2.7
51–59	37,302,635	12.2
60–70	31,296,308	10.3

DRI = Dietary Reference Intake

The study team used the first approach for the HEI-2015, HEI-2010, and usual nutrient intakes outcomes. For this approach, the mean score for each DRI age group was calculated. The mean score for each comparison group was computed as the weighted average of the age group estimates for that group, using census proportions. The team used the same set of weights for each comparison group and used the following equation to calculate standard errors for HEI-2015, HEI-2010, and usual nutrient intakes:

$$\sqrt{\sum_{i=1}^{J}\left[\left(SE_{X_i}\right)^2 \times (K_i)^2\right]}$$

In this equation, SE_{xi} is the standard error for DRI age group "i," and K_i is the census proportion adjustment for that age group.

The second approach was used for the calories from saturated fats, added sugars and alcohol; BMI; and food choices outcomes. For this approach, the outcome was first calculated for each individual. SAS procedures were used to calculate age-adjusted estimates and standard errors. Census proportion adjustments for each DRI age group were incorporated into PROC SURVEYMEANS and PROC SURVEYREG. Output from running PROC SURVEYREG provided separate estimates and standard errors for all individuals, SNAP participants, income-eligible nonparticipants, and higher income nonparticipants. Age adjustment was not performed for multivariable analysis.

For the descriptive analysis, age adjustment was not applied to the average amounts of foods consumed. Insufficient sample sizes prevented the computation of reliable estimates for numerous components of this analysis. For many of the food subgroups included, specific age groups had zero participants consuming food in that subgroup. When no one in an age or comparison group consumed a food, there was insufficient variation to use age-adjustment procedures.

3. Statistical Significance

For the descriptive analyses, the study team conducted *t*-tests to determine whether differences in outcomes between program participants and each group of nonparticipants (income-eligible nonparticipants and higher income nonparticipants) reached statistical significance. Because of the large number of *t*-tests conducted (comparing SNAP participants and each group of nonparticipants, overall and by age group and gender), the team urges caution in interpreting results; a proportion of these tests would be expected to be significant just by chance. All tables throughout the report indicate statistically significant differences at the .05, .01, and .001 levels and statistically significant differences at the .05 level or better.

For the matched analyses, different sets of *t*-tests were conducted. For analyses comparing matched SNAP participants and nonparticipants, only one *t*-test was conducted for each nutrition outcome to determine whether the difference between matched participants and nonparticipants was significant. For the analyses described in appendix G, the study team conducted *t*-tests to determine whether differences in outcomes between SNAP participants and matched nonparticipants were statistically significant.

4. Indicators of Statistical Reliability

The study team tested all estimates for statistical reliability according to recommendations in the NHANES analytic guidelines on variance estimation. These guidelines recommend estimates have a relative standard error of 30 percent or less rather than a minimum sample size. Because the design effect is highly variable for different variables within each 2-year cycle of the continuous NHANES, the analytic guidelines do not set a single minimum sample size for analysis (CDC NCHS, 2020d). Estimates in each table were flagged with a "**u**" to indicate the estimate was statistically unreliable if the coefficient of variation (ratio of the standard error to the mean expressed as a percent) was greater than 30 percent. Unreliable estimates are not discussed in this report.

5. Propensity Score Estimation and Matching

Simple differences in nutrition outcomes observed between groups of participants and nonparticipants may reflect differences in demographic, economic, or household characteristics of the groups rather than an effect of program participation. When people with certain characteristics (which are also related to the outcomes of interest) are more likely to participate in a program, this is known as selection bias. The only method that would provide a true assessment of the impact of program participation on nutrition outcomes would be randomly assigning people to the two groups, an option impossible to implement. Without this option, one can use a nonexperimental method such as multivariable analysis.

The study team used a propensity score approach (Mabli et al., 2010) to account for differences in the characteristics of the comparison groups rather than controlling for those differences within multivariate regression models. The objective of propensity score matching is to achieve balance on the observed covariates and generate comparison groups that would have been expected in a randomized experiment. This method was selected because the computational methods used to estimate the nutrition outcomes were too complex to incorporate into a regression modeling framework.

The three steps to using this approach are as follows:

1. Estimate the propensity scores using available covariates.
2. Match the comparison groups based on those scores.
3. Use the newly formed comparison groups in the nutrition outcome analyses.

These steps are described in general and then described regarding how they were implemented specifically for the matched analyses comparing nutrition outcomes of matched SNAP participants and income-eligible nonparticipants.

A propensity can be defined as the probability of an individual being assigned to a particular "treatment" group, given a set of observed covariates:

$$p(x) = \Pr(T = 1 | X = x)$$

In this equation, *T* is the binary treatment group, and *X* is a set of observed covariates.

The purpose of the propensity score estimation and matching was to minimize the selection bias inherent in the descriptive comparison groups. To accomplish this, one must include as many variables as possible that might explain differences between the comparison groups. These variables should be associated with both the "treatment" (SNAP participation status) and the "outcome" (each dietary outcome; Stuart, 2010). Because the list of relevant variables available in NHANES likely does not account for all possible confounders, this study's findings do not indicate causality (the impact of participating in SNAP).

The study team used logistic regression modeling to compute a score for each respondent included in the study, representing the likelihood (expressed as a proportion) the respondent would be a member of a particular comparison group, based on the respondent's characteristics. The team began by including in the logistic regression model a set of characteristics found to be strong predictors of program participation (Condon et al., 2015a), with variables available in the NHANES 2011–2016. These factors are summarized in table A.7. Not all these variables were retained in the final model (as described in more detail later in this section).

The team used the estimated propensity scores to reduce the analytic sample to those individuals with similar characteristics in all comparison groups. The propensity score estimation and matching resulted in comparison groups that were similar based on the characteristics the team entered into each of the three logistic regression models, with dissimilar individuals discarded. These new comparison groups were then used for each of the dietary outcome analyses. The use of these new comparison groups adjusted, or controlled, for the variables marked in table A.7. All nutritional outcome differences were tested statistically using two-sample, two-tailed *t*-tests.

The following sections describe the process for propensity score estimation; matching based on those scores, including the resultant sample size; and analyses.

Propensity score estimation

The study team estimated propensity scores representing the likelihood of being a SNAP participant, given a set of characteristics, for NHANES 2011–2016 respondents aged 1 and older.

The variables in the far left column of table A.7 were entered into a logistic regression model, with the independent variable being a dichotomous indicator for SNAP participation status (SNAP participant versus income-eligible nonparticipant). For the logistic regression model to converge, some variables were removed. These variables were identified by their extreme odds ratios and very large standard errors, indicators of a lack of variability in their distribution. In some cases, the removed variables were highly associated with participation in SNAP (e.g., health insurance type). In others, they were so highly correlated with other variables in the models, only one of those variables could be retained. The variables in the final model are indicated in table A.7 in the column marked “Comparison 1.” The final model was used to estimate a propensity score for each participant and nonparticipant. The propensity score is the estimated likelihood an individual from either group may be a SNAP participant given their characteristics.

Table A.7. Variables Entered Into and Retained in the Propensity Score Estimation Model

Initially Entered Into All Propensity Score Estimation Models	Variables Retained in Final Model
NHANES 2011–2016	
Gender	●
Race/ethnicity	●
Age	●
U.S. citizenship	●
Education, highest grade completed	●
Total number of people in household	●
Marital status (aged 20 and older)	
Employment, type of work done last week	
Employment, hours worked past week	
Ratio of income to poverty	●
Annual household income	
Annual family income	
Monthly family income	●
Total savings/cash assets for the family	●
Income from Supplemental Security Income	●
Income from State/county cash assistance	●
Number of months working in main job	
Type of health insurance	
General health condition	
Consumer Behavior Questionnaire	
Money spent at supermarket/grocery store	●
Money spent on nonfood items	●
Money spent on food at other stores	●
Money spent on eating out	●
Money spent on carryout/delivered foods	●

Propensity score matching

The next step was to match individuals using 1:1 nearest-neighbor matching with replacement, implemented by the %PSMatching SAS macro (Coca-Perraillon, 2007). This nearest-neighbor algorithm matched each participant with the nonparticipant, resulting in the smallest between-propensity-score difference. All nonparticipants not matched to a SNAP participant were discarded.

The study team opted for 1:1 matching, meaning each SNAP participant was matched with one income-eligible nonparticipant; this implies the final sample size will not exceed twice the number of SNAP participants. The team also allowed for replacement—after a nonparticipant was selected, they were returned to the pool of possible matches—so each nonparticipant was able to be selected more than once if they were the best match for a participant. This optimized the quality of the matching, allowing each participant to be matched to the best possible nonparticipant, not just the best possible nonparticipant remaining in the sample. This process of matching reduced the sample to those individuals whose propensity scores (and, hence, characteristics) were very similar to each other.

Changes in the size of the analytic sample are presented in table A.8. The differences between the pre- and post-estimation sample sizes are because of missing information on covariates.

Table A.8. Sample Size Pre- and Post-Matching

Participant Group	Before Propensity Score Estimation	After Propensity Score Estimation	After Matching
Matched SNAP participants	4,988	2,983	2,983
Matched income-eligible nonparticipants	4,551	2,277	1,301

Analyses

Differences in nutrition and health outcomes between the two matched groups were tested statistically using two-sample, two-tailed t-tests.