

Vital and Health Statistics

Series 2, Number 178

February 2018



Estimating Usual Dietary Intake From National Health and Nutrition Examination Survey Data Using the National Cancer Institute Method

Data Evaluation and Methods Research



U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Copyright information

All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.

Suggested citation

Herrick KA, Rossen LM, Parsons R, Dodd KW. Estimating usual dietary intake from National Health and Nutrition Examination Survey data using the National Cancer Institute method. National Center for Health Statistics. Vital Health Stat 2(178). 2018.

Library of Congress Cataloging-in-Publication Data

Names: National Center for Health Statistics (U.S.), issuing body.
Title: Estimating usual dietary intake from National Health and Nutrition Examination Survey data using the National Cancer Institute method.
Other titles: Vital and health statistics. Series 2, Data evaluation and methods research ; no. 178. | DHHS publication ; no. (PHS) 2018-1378. 0276-4733
Description: Hyattsville, Maryland : U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, February 2018. | Series: Vital and health statistics. Series 2, data evaluation and methods research ; number 178 | Series: DHHS pub ; number 2018-1378 | Includes bibliographical references.
Identifiers: LCCN 2017058786| ISBN 9780840606877 (pbk.) | ISBN 0840606877 (pbk.)
Subjects: | MESH: National Health and Nutrition Examination Survey (U.S.) | Health Surveys | Diet Surveys | Eating | Statistics as Topic | United States
Classification: LCC RA408.5 | NLM W2 A N148vb no.178 2018 | DDC 614.4/2--dc23 LC record available at <https://lccn.loc.gov/2017058786>

For sale by the U.S. Government Printing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20402-9328
Printed on acid-free paper.

Vital and Health Statistics

Series 2, Number 178

Estimating Usual Dietary Intake From National Health and Nutrition Examination Survey Data Using the National Cancer Institute Method

Data Evaluation and Methods Research

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
February 2018
DHHS Publication No. 2018-1378

National Center for Health Statistics

Charles J. Rothwell, M.S., M.B.A., *Director*

Jennifer H. Madans, Ph.D., *Associate Director for Science*

Division of Health and Nutrition Examination Surveys

Kathryn S. Porter, M.D., M.S., *Director*

Ryne Paulose-Ram, Ph.D., *Associate Director for Science*

Division of Vital Statistics

Delton Atkinson, M.P.H., M.P.H., P.M.P., *Director*

Hanyu Ni, Ph.D., M.P.H., *Associate Director for Science*

Contents

Acknowledgments	v
Abstract	1
Introduction	1
NHANES Dietary Data Collection	1
Methods to Estimate Usual Dietary Intake	2
Methods	2
Analytic Sample	2
Covariates	3
Dietary Components and Nutrients	3
Implementing NCI Method With NHANES Data	4
Results	9
Discussion	11
Limitations	13
Summary	14
References	15
Appendix I. Nutrient Reference Values Used in This Report	28
Appendix II. Steps for Creating Poststratified Balanced Replicated Weights	29
Appendix III. SAS Syntax Used to Create Poststratified Adjusted Balanced Repeated Replication Weights for Use in Usual Intake Estimation	30
Appendix IV. National Cancer Institute Macros Used to Estimate Usual Intake	39
Appendix V. Usual Intake Estimation—Univariate—Calcium Example	40
Appendix VI. Usual Intake Estimation—Ratio and Bivariate—Percentage of Energy From Fat Example	43
Appendix VII. Empirical Approach to Selecting a Correlated or Uncorrelated Model	54

Text Figures

1. Estimated distributions of calcium using three different methods for men aged 40–59: National Health and Nutrition Examination Survey, 2011–2014	9
2. Percentage of U.S. population with calcium intake below estimated average requirement using different estimation methods, by age and sex: National Health and Nutrition Examination Survey, 2011–2014	10
3. Estimated distributions of fruit intake using three different methods for girls aged 2–5 years: National Health and Nutrition Examination Survey, 2009–2012	11
4. Percentage of U.S. population with fruit intake below the lower recommended range, by age and sex: National Health and Nutrition Examination Survey, 2009–2012	12
5. The relationship between bias and precision when comparing given day with usual intake methods for estimating intake	14

Text Tables

A. Long-format data structure	5
B. Wide-format data structure	5

Detailed Tables

1. SAS macros developed by National Cancer Institute to run nonlinear mixed models used in this report	18
2. Calcium (grams per day): Mean, percentile, and percentage of persons meeting dietary reference intakes, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014	19
3. Total fat (grams per day): Mean and percentile, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014	21

4.	Energy (kilocalories per day): Mean and percentile, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014	22
5.	Percentage of energy intake from total fat (per day): Mean, percentile, and percentage of persons meeting dietary reference intakes, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014	23
6.	Total fruit (cup equivalents per day): Mean, percentile, and the percentage of persons below and above the 2015–2020 Dietary Guidelines recommended consumption range, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2009–2012	26

Appendix Tables

I.	Dietary reference intake for calcium (estimated average requirement and tolerable upper limit), percentage of energy from fat (acceptable macronutrient distribution range), and recommended intake range of fruit, by the Institute of Medicine Life Stage Group	28
II.	National Cancer Institute macros used to estimate usual intake	39

Acknowledgments

The authors of this report gratefully acknowledge the assistance of Naman Ahluwalia, Amy Branum, Ryne Paulose-Ram, and Cynthia Ogden in the preparation of this report.

Abstract

Background

Dietary recommendations are intended to be met based on dietary intake over long periods, as associations between diet and health result from habitual intake, not a single eating occasion or day of intake. Measuring usual intake directly is impractical for large population-based surveys due to the respondent burden associated with reporting habitual intake over longer periods. Therefore, analytical techniques were developed to estimate usual intake using as few as 2 days of 24-hour dietary recall data. With National Health and Nutrition Examination Survey (NHANES) data, this report demonstrates how to estimate usual intake using the National Cancer Institute (NCI) Method.

Methods

This report demonstrates how to estimate the usual intake of nutrients consumed daily or episodically using NHANES data. Means, percentiles, and the percentages above or below specified Dietary Reference Intake (DRI) values for given day, within-person mean (WPM), and estimates of usual intake are presented.

Results

Consistent with previous analyses, mean intakes were similar across methods. However, the distributions estimated by nonusual intake methods were wider compared with the NCI Method, which can lead to misclassification of the percentage of the population above or below certain DRIs.

Conclusion

Use of NHANES data to examine the proportion of the population at risk of insufficiency or excess of certain nutrients, with methods like given day and WPM that do not address within-person variation, may lead to biased estimates.

Keywords: diet • statistical distributions • National Health and Nutrition Examination Survey

Estimating Usual Dietary Intake From National Health and Nutrition Examination Survey Data Using the National Cancer Institute Method

by Kirsten A. Herrick, Ph.D., M.Sc., and Lauren M. Rossen, Ph.D., M.S., National Center for Health Statistics; Ruth Parsons, B.A., Information Management Services, Inc.; and Kevin W. Dodd, Ph.D., M.S., National Cancer Institute

Introduction

The National Health and Nutrition Examination Survey (NHANES) is a program of studies that has collected dietary information from respondents since the 1970s (1). NHANES, which began collecting data continuously in 1999, is a stratified, multistage probability survey of the U.S. civilian noninstitutionalized population (2). NHANES data are publicly released by the National Center for Health Statistics (NCHS) every 2 years on approximately 10,000 individuals. NHANES participants are administered a series of questionnaires in a detailed in-home interview, followed by a standardized physical examination and dietary intake interview at a specially designed mobile examination center (MEC).

Dietary intake data from NHANES are routinely used to inform and evaluate policy. For example, NHANES data have been used to evaluate the adequacy of food packages provided by the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) (3). The Dietary Guidelines for Americans also rely on NHANES data to document consumption patterns in the United States, and the Food and Drug Administration uses information from NHANES to determine whether changes

to the vitamin and mineral fortification regulations are needed (1).

NHANES Dietary Data Collection

NHANES has collected dietary data with two instruments—a food frequency questionnaire (FFQ) and a 24-hour recall. Since 1999, the 24-hour recall has been the primary source of dietary intake data for NHANES. Since 2002, NHANES has administered two 24-hour recalls to MEC participants: the first recall is administered in the MEC, and then a second recall is administered by phone 3–10 days later.

NHANES is designed so that all of the days of the week and seasons of the year are represented in the sample. This feature ensures that a single 24-hour recall can be used to estimate the average consumption of a population or subgroup (4–8) under two key assumptions. The first assumption is that the 24-hour recall provides an unbiased measure of true intake on a population basis (i.e., free from systematic error [9,10]). The second assumption is that seasons of the year and days of the week are adequately represented across the population. However, a single 24-hour recall, commonly referred to as “given day,” is an imprecise measure of an individual’s usual intake due in large part to day-to-day variation in individual diets.

An important consideration regarding measuring dietary intake is minimization of measurement error. All self-report measures of dietary intake are subject to systematic and random measurement error (1,11–14). Since 24-hour recalls are limited to foods consumed in the previous day, the magnitude of systematic errors related to recall bias is reduced compared with FFQs (9). However, dietary reporting from 24-hour recalls is impacted by the combination of day-to-day variation and variation due to other random errors, generally considered together as “within-person variation” (9).

One goal of measuring dietary intake is to capture usual or habitual intake or the dietary intake of a person over long periods of time (4,6). This is important because dietary recommendations are intended to be met over time, and diet–health relationships are often the product of long-term exposures, not a single eating occasion (15). For this reason, comparing nutrients from a single day with Dietary Reference Intakes (DRIs) is not recommended to evaluate dietary adequacy. However, given that most large-scale population-based surveys, like NHANES, use only 1 or 2 days of 24-hour recalls as the primary source to measure dietary intake, more sophisticated analytic methods, employed post-data collection, are necessary to describe usual intake using NHANES dietary data (16–19).

Methods to Estimate Usual Dietary Intake

Different analytic methods for estimating usual intake have been developed over the past 3 decades to address the issue of within-person variation found in dietary intake data as assessed by 24-hour recalls. They include within-person mean, the National Research Council/Institute of Medicine (NRC/IOM) methods, the Iowa State University (ISU) method, the Statistical Program to Assess Dietary Exposure (SPADE), and the National Cancer Institute (NCI) Method (5,9,20–25), all of which are described briefly below.

The simplest method for estimating usual intake is to average multiple 24-hour recalls per individual and call this within-person mean (WPM) an approximation of usual intake. In the context of NHANES, this means averaging data from both days of dietary intake to estimate usual intake. This can lead to biased estimates of inadequate or excess intake because it does not: (a) distinguish within-person from between-person variation; (b) account for consumption amounts that are at the extremes, either zero or extremely positively skewed; (c) allow for adjustment by covariates; or (d) allow for the correlation between the amount consumed and the probability to consume (9). Theoretically, with a larger number of recalls to average across, it would be possible to mitigate points a, b, and d. However, since NHANES only collects two recalls, use of WPM to address these issues is limited.

The NRC/IOM methods represent an innovation in estimating usual intake, because within-person variation was recognized and partitioned for the first time (9). The ISU method, implemented by using PC-Side, software for intake distribution estimation developed by ISU (26), accounts for within-individual variation and can be used in complex surveys (27). However, if many 24-hour recalls have no consumption of a food item, it cannot account for the large number of zero consumers, and adjustment by covariates is limited (5).

The NCI Method (5) has evolved from earlier methods, including the ISU method and NRC/IOM methods, and has been extended for use in complex surveys, such as NHANES. This method can incorporate covariates and therefore provides estimates of usual intake for subpopulations (28), handles situations where the amount consumed and probability to consume are correlated (5), deals with foods or nutrients where a significant proportion of the population reports no consumption of the item of interest on a given day (29,30), and models usual intake in relation to health outcomes (23). SPADE is comparable to the NCI Method, except it can be implemented using the statistical software

R as a package called SPADE.RIVM (25,31).

The goal of this report is to demonstrate how to estimate usual intake appropriately by adapting one particular method that can be used with complex surveys, the NCI Method, using NHANES dietary data by providing examples of estimating usual intake of various nutrients and foods. Therefore, the focus of this report is not to describe dietary intake for the U.S. population but to supplement NCI’s User’s Guide (32), NHANES web tutorials (33), and examples from NCI’s website (34–36). This is done by providing specific examples of using usual intake methods with NHANES data, including some of the technical issues that can potentially arise when using the NCI Method.

Estimates of intake, calculated by the NCI Method, are compared with intake estimates from simpler methods, (WPM and given day) for illustrative purposes. Additionally, this report provides an overview of the analytic and programming considerations that an analyst may encounter when applying the NCI Method to NHANES dietary data, such as how to create and use balanced repeated replication (BRR) weights to account for the complex survey design. Syntax and instructions are provided to estimate usual intake, standard error, and the proportion above or below a threshold using available NCI SAS macros (34).

Methods

Analytic Sample

All participants who took part in the examination component of NHANES were eligible for the dietary interview. This analysis was limited to all respondents aged 2 years and over who completed at least one dietary recall that was reliable and met the minimum criteria of completeness, namely that four of the five steps in the Automated Multiple-Pass Method (AMPM) were completed and that foods and beverages consumed were identified (37,38). This analysis used data from the 2011–2014 NHANES to estimate the usual intake of

selected nutrients (i.e., calcium, total fat, total energy, and percentage of energy from fat). Data from the 2009–2012 survey were used to estimate the usual intake of one type of food (i.e., fruits), because the Food Patterns Equivalents Database (FPED) for 2013–2014, which is used to quantify total fruit consumption (39), was not available for use at the time of this analysis. The unweighted total examination response rate in 2009–2014 for participants aged 2 years and over ranged from 68.5% to 76.0% (40). The NHANES protocol was approved by the NCHS Research Ethics Review Board. Informed consent was obtained from persons aged 18 and over. Written parental consent and child assent was obtained for participants aged 2–17.

Of the 18,259 respondents aged 2 years and over examined in the 2009–2012 NHANES, 1,131 persons did not provide a dietary recall, 142 supplied a recall but were deemed unreliable, and 11 reported consuming some breast milk and were therefore excluded. The amount of breast milk consumed is not quantified in the 24-hour recall, therefore, these records contain missing values for the amount consumed and for the amount of energy and nutrients from breast milk. Consequently, no total nutrient intakes (contained in the Total Nutrient Intakes files) are available for respondents who reported consuming any breast milk. An additional 168 women who were either breastfeeding or pregnant were excluded, because the eating patterns for this group can differ from the general public, and they have separate DRIs for many nutrients.

Of the 17,912 respondents aged 2 years and over examined in the 2011–2014 NHANES, 1,700 persons did not provide a dietary recall, 200 supplied a recall but were deemed unreliable, and 12 respondents reported consuming breast milk and were therefore excluded. An additional 170 women who were either breastfeeding or pregnant were excluded. The final analytic samples, after exclusions, comprised 16,807 respondents in the 2009–2012 survey and 15,830 respondents in the 2011–2014 survey.

Covariates

The NCI Method allows for adjustment of dietary intake by covariates. Moreover, covariates are used to create the BRR weights necessary to use the NCI Method with NHANES data, and covariates also guide the presentation of results. The tables in this report present age groups at the time of interview (i.e., 2–5, 6–11, 12–19, 20–39, 40–59, and 60 and over), consistent with NHANES sample-selection methods (41). Self-reported race and Hispanic origin was categorized as non-Hispanic white, non-Hispanic black, non-Hispanic Asian, Hispanic, and other (includes multiracial groups) when creating BRR weights for the 2011–2014 NHANES. Self-reported race and Hispanic origin was categorized as non-Hispanic white, non-Hispanic black, Hispanic, and other (includes multiracial groups) when creating BRR weights for the 2009–2012 NHANES. Usual intake estimation with BRR weights accounts for day of the week, nonresponse, noncoverage, and unequal probabilities of selection and the complex sample design of NHANES. More details on BRR weights is provided in the Methods section and in [Appendices II](#) and [III](#).

Dietary Components and Nutrients

During the NHANES dietary interview, trained interviewers, using the computer-assisted dietary interview system AMPM, collect the type and quantity of all foods and beverages consumed in a single 24-hour period, specifically the 24-hour period before the dietary interview (from midnight to midnight), at the MEC. AMPM includes standard probes to reduce the probability of missed or forgotten food items in a multiple-pass method designed to maximize complete and accurate reporting (38). A second telephone-administered 24-hour recall is collected (3–10 days after the first 24-hour recall at the MEC). Proxies, 92% of whom were the participant's mother, report dietary intake for children aged 5 years and under and assist with the dietary

interview for children aged 6–11. Dietary intake is self-reported for participants aged 12 and over (37).

Four nutrients and one food were chosen for the examples in this report, because they are commonly represented in nutritional epidemiological analyses and demonstrate a variety of analyses that can be accomplished with the NCI Method to estimate usual intake with NHANES data. These examples also supplement the examples provided on NCI's website (34–36). These nutrients are calcium (grams), energy (kilocalorie), fat (grams), percentage of energy from fat, and fruit (cup equivalents). The first four represent ubiquitously consumed nutrients (i.e., nutrients that are consumed by nearly everyone, every day), and fruit represents a food that is episodically consumed (i.e., occasionally consumed by a smaller segment of the overall population). Approximately 25% of the population reported no consumption of fruit on both 24-hour recall days. The percentage of energy from fat represents a special type of ubiquitously consumed nutrient, because it is a ratio of means, total fat consumed divided by total energy consumed. Different methods are required for estimating the usual intake of ubiquitously and episodically consumed nutrients, foods, and ratios of nutrients, which are described below. Estimates of energy intake are presented because they are a component used to calculate energy from total fat (13,14).

Day 1 (37,42) and Day 2 (43,44) Total Nutrients files from the 2011–2014 NHANES were used for the analysis of calcium, energy, fat, and the percentage of energy from fat described in this report. These are summary files that describe the total energy and nutrients from the foods and beverages captured in the Individual Food files. For the estimation of the cup equivalents of fruit consumed, the 2009–2012 FPED files were used (available from the United States Department of Agriculture (USDA) website [45]). The FPED files use the Individual Food files and disaggregate each reported food and beverage into its component parts, which are then organized into 37 groups based on the USDA's Food Patterns

(FP), developed from the 2010 Dietary Guidelines for Americans. One FP of interest is the consumption of fruit, which is analyzed in this report.

Dietary Reference Intakes

The estimated population distribution of usual intake can be used to assess dietary intake relative to population-based recommendations more appropriately than WPM or given-day methods. To demonstrate this, estimated usual nutrient intakes were compared with sex and age-specific DRIs established by IOM (17,46–48). The DRIs presented here include the Estimated Average Requirements (EARs) and Tolerable Upper Intake Level (UL). EAR is “the average daily intake level estimated to meet the requirement of half of the healthy individuals in a particular life stage and gender group” and is considered the best measure of population adequacy of nutrient intake (46). Intakes lower than EAR indicate the estimated prevalence of inadequate intakes within a group. The proportion of the population with intakes greater than UL identifies those with excessive intakes who are potentially at risk of adverse effects (19).

In contrast to DRIs, which address single nutrients, Acceptable Macronutrient Distribution Ranges (AMDRs) have been established for the energy derived from macronutrients, such as the percentage of energy from fat. Ranges are used, rather than a discrete number, because energy needs vary by individual. AMDR is the range of intake for a particular energy source that is associated with reduced risk of chronic disease while providing intakes of essential nutrients (17). Consumption in excess of the upper range of AMDR is potentially associated with increased risk for some diet-related chronic diseases, while consumption below the lower range of AMDR may be related to insufficient intakes of essential nutrients (49).

There are no DRIs for fruit, however, the 2015–2020 Dietary Guidelines for Americans promote recommended ranges of intake by age and sex that can be used like DRIs (15). For fruit intake, this report presents the estimated percentages

below the lower range and above the upper range of consumption. [Appendix I](#) provides details of the IOM DRIs by age and sex that are used in this analysis.

Importantly, DRIs are meant to be applied to true usual intakes for a given population. Given day and WPM do not meet that standard, however, they are presented for illustrative purposes (16–19). Additionally, in the analysis presented in this report, DRIs were applied to estimates of usual intake distributions, generated by the NCI Method under the assumption that 24-hour recalls are an unbiased measure of true population intake.

Implementing NCI Method With NHANES Data

The NCI Method uses mixed effects models, which contain both random and fixed effects, to estimate usual intake by separating and removing the within-person variation from the between-person variation. For episodically consumed foods or nutrients, this method involves two-part models. The first part uses logistic mixed-effects models to estimate the probability (ranging from 0 to 1) to consume a food or nutrient. The second part uses linear mixed-effects models to estimate the amount of the food or nutrient consumed. The two parts are linked through the inclusion of person-specific random effects and by the optional inclusion of shared covariates (5). For ubiquitously consumed foods or nutrients (e.g., those consumed by nearly everyone in the sample), the NCI Method involves only the “amount” model or the linear mixed-effects model because the probability to consume is 1 or nearly close to 1. Additional steps are required beyond the two-part model to estimate distributions of usual intake using Markov chain Monte Carlo (MCMC) simulations, which are described below.

There are three major components necessary to run the NCI usual intake models using NHANES data:

1. Multiple 24-hour dietary recalls on a proportion of the population
2. Poststratified, integerized BRR weights

3. SAS macros developed by NCI to run nonlinear mixed models or the NCI Method (34) ([Table 1](#))

Number of 24-hour recalls

To utilize the NCI Method, there must be multiple 24-hour dietary recalls on a proportion of the population. A minimum of two dietary recalls is necessary to separate the within-person from between-person variation. However, the sample size for analyses using the NCI Method is determined by the total number of respondents who have at least 1 day of dietary recall. When using NHANES data, this means the total number of respondents in the analytic sample will be equal to the number with a Day 1 recall. There is no rule for the proportion of the population with more than one recall needed to use the NCI Method. However, the variability in the nutrient is linked to the proportion of the population with duplicate recalls needed to produce stable and reliable estimates: the more variability, the greater the proportion of the population with duplicate recalls needed. Typically, approximately 85% of respondents in NHANES have two dietary recalls (37,42,50). For nutrients and foods consumed by a majority of the population every day, 85% of respondents with duplicate records is sufficient to implement the NCI Method.

- The structure of the two Total Nutrient Intakes files (Days 1 and 2) of dietary recall data must be in long format. This means that each day of a person’s dietary recall is represented by a separate row of data, and a variable on the file indicates the day. For NHANES data, each participant’s identification number (i.e., SEQN) would thus be listed twice if that participant completed 2 days of dietary recall; the first row represents recall data from Day 1 and the second row represents recall data from Day 2 ([Table A](#)). The contrasting format would be a wide file, where the Day 2 dietary data would be added as additional variables to the individual-level record. SEQN would only appear

once in the file, and dietary data would be represented in separate columns ([Table B](#)).

Balanced repeated replication

To analyze NHANES data appropriately, the differential weighting of individuals and the correlation of individuals within primary sampling units (PSU) and strata must be taken into account. Given-day intake or WPMs can be estimated using standard SUDAAN or SAS SURVEY commands, which include the design variables (PSU, STRATA) and the dietary sample weights provided by NHANES. However, the NCI Method uses the SAS procedure, **NLMIXED**, to estimate a nonlinear mixed-effects model. **NLMIXED** allows for sample weights but does not explicitly account for the complex survey design (PSU and STRATA). In order to account for both the differential weighting and the complex survey design of NHANES, BRR weights must be used with the **NLMIXED** procedure. The **NLMIXED** procedure accounts for differential weighting by using the REPLICATE statement.

BRR is a statistical method for estimating the sampling variability of a statistic, taking into account complex sample design through replicate weights. Instead of a single weight per individual (as is used in Taylor series linearization and the standard NHANES weights), BRR uses multiple weights or replicates for each individual in the survey. BRR weights are derived from the original sample weight and the design variables PSU and STRATA. [Appendix II](#) describes the procedure to create BRR weights, while [Appendix III](#) provides syntax for the SAS procedure, **PROC SURVEYFREQ**, which was used to create the BRR weights described in this report. In order to calculate BRR weights, users must specify the number of BRR replicate weights needed and the perturbation factor.

In BRR, one-half of the sample is used at a time: one of two PSUs from each stratum. In standard BRR, the sample weights of the observations in the selected PSUs are doubled, while the sample weights of the observations in

Table A. Long-format data structure

SEQN	Day	Nutrient_1	Nutrient_2
1	1	500 grams	10 grams
1	2	520 grams	15 grams
2	1	870 grams	2 grams
2	2	820 grams	7 grams

Table B. Wide-format data structure

SEQN	Nutrient_1_Day1	Nutrient_1_Day2	Nutrient_2_Day1	Nutrient_2_Day2
1	500 grams	520 grams	10 grams	15 grams
2	870 grams	820 grams	2 grams	7 grams

the PSU not chosen are set to zero. The number of replicates (R) that are needed for BRR is the smallest integer that is divisible by 4 and is greater than or equal to S , where S is the number of strata.

NHANES typically has 15 strata (S) per survey cycle. Consequently, for 2 years of NHANES data, this number is 16. There is additionally always a base run, so the total number of BRR weights created are the numbers below, plus 1.

- 4 years of NHANES = 32 BRR weights
 - 4 years = $15 + 15$ strata = 30, $R > 30$ and divisible by 4, therefore 32 is the smallest integer
- 6 years of NHANES = 48 BRR weights
- 8 years of NHANES = 60 BRR weights
- 10 years of NHANES = 76 BRR weights
- 12 years of NHANES = 92 BRR weights

A modification of BRR, called Fay's method, is typically used when implementing the NCI Method with NHANES data. With Fay's modification, the sample weights from BRR are not zero-weighted in one-half of the sample and double-weighted in the other half. Instead, they are weighted by a factor F (F is a proportion that ranges between 0 and 1). This factor weighs down one-half of the sample by F , and the other half is weighted up by $1 - F$. For example, when $F = 0.3$, the corresponding perturbation factor is 70%. Weights are decreased by

30% in the one-half sample and increased by 70% in the other one-half sample. This is the standard perturbation factor used in other reports and analyses of NHANES dietary intake data (51,52).

Special consideration related to BRR weights

Prior to creating BRR weights, analysts must ensure that there are only two strata per PSU, as there are occasions when NHANES has three PSUs in a single stratum. All participants within one randomly selected PSU should be reassigned into one of the remaining two PSUs in that stratum. The whole PSU is reassigned at one time, rather than randomly assigning individuals in the third PSU to PSU 1 or 2. Sample code for this reassignment and recoding of the PSUs is shown in [Appendix III](#).

BRR weights created using the original NHANES dietary sample weight and survey design information will not necessarily reflect the same population totals as the original sample weights. This is because one-half of the PSU is up-weighted and the other half is down-weighted. Therefore, the BRR weights are no longer necessarily representative of the U.S. population. This can be observed by looking at the weighted population totals of all combinations of the common domains used to create the representative sampling frame (i.e., age, sex, race and Hispanic origin, and comparing BRR weights with the original sample weight [referred to as control] frequencies [syntax provided in [Appendix III](#)] demonstrates how to obtain

control frequencies]). It is necessary to poststratify the BRR weights so that each of the replicate weights will produce estimates that are representative of the U.S. population, consistent with the original sample weight. In other words, the age-sex-race and Hispanic-origin-specific population totals produced by the original sample weight should be matched by each of the BRR weights. It is an additional requirement of the macros used in the NCI Method that the BRR weights be integerized (i.e., rounded to the nearest whole number).

For the purposes of this analysis, two sets of BRR weights were generated. For the estimation of nutrients for the 2011–2014 NHANES (calcium, energy, total fat, and the percentage of energy from fat), all possible combinations of poststratification cells by age, sex, and race and Hispanic origin were created. The following categories were used for race and Hispanic origin: non-Hispanic white, non-Hispanic black, non-Hispanic Asian, Hispanic, and other combined, and for age groups: 2–5, 6–11, 12–19, 20–30, 40–59, and 60 years and over. For the estimation of the usual intake of fruit, all possible combinations of poststratification cells by sex, race and Hispanic origin, and age were created. The following categories were used for race: non-Hispanic white, non-Hispanic black, Hispanic, and other combined, and for the age groups: 1–3, 4–8, 9–13, 14–18, 19–30, 31–50, 50–69, and 70 years and over. The race and Hispanic-origin categories reflect the categories used to create a representative sample in NHANES for their respective survey cycles. The 2011–2014 age categories reflect the traditional sampling and reporting categories for NHANES, while the age categories for 2009–2012 reflect the Life Stage Groups used by the National Academies of Science, Engineering, and Medicine (formerly the Institute of Medicine) for DRIs ([Appendix I](#)). Different age groups were used to examine the ease of implementing the NCI Method with each of the age categorizations. Analysts are encouraged to choose the age categorization that best matches the goals of their analysis.

[Appendix II](#) details the steps that are taken before and after creating BRR

weights to check that poststratification was successful. [Appendix III](#) provides syntax to create poststratified BRR weights expressed as integers. The syntax to create BRR weights in SAS is straightforward.

SAS macros for the NCI Method

As stated earlier, this report focuses on the estimation of single dietary components consumed daily (i.e., calcium, total fat, and total energy); a single dietary component consumed episodically (i.e., fruit); and a ratio of two dietary components consumed daily (i.e., percentage of energy from fat, also referred to as a bivariate model).

[Appendix IV](#) summarizes the macros used for each nutrient and food being estimated. All macros use BRR variance estimation to calculate standard errors and confidence intervals. Below is a brief description of each macro. For more detailed information, see the User's Guide (version 2.1) available from NCI ([32](#)). [Appendix V](#) provides sample code for the usual intake estimation of calcium as an example of univariate estimation, and [Appendix VI](#) provides sample code for the usual intake estimation of the percentage of energy from fat as an example of bivariate estimation.

Estimating mean usual intake

Given that NHANES is designed so that a single 24-hour recall can be used to estimate the mean consumption of a population or subgroup ([4–8](#)), the need to use more sophisticated modeling methods to estimate mean usual intake might seem unnecessary. However, the NCI macros build on each other, and it is necessary to estimate mean usual intake first, so that it is possible to next estimate a nutrient usual intake distribution. For a single dietary component, the **MIXTRAN Macro** fits nonlinear mixed effects models, using the SAS **NLMIXED** procedure to obtain parameter estimates and allow for the evaluation of covariate effects. The **MIXTRAN Macro** has the flexibility of running three different models based on the user-defined consumption pattern of the food or nutrient being modeled: (a) ubiquitously consumed nutrients, (b)

episodically consumed nutrients where the probability of consumption and the amount consumed are not correlated, and (c) episodically consumed nutrients where the probability of consumption and the amount consumed are correlated.

- For a *single dietary component consumed daily* (i.e., ubiquitously consumed), **MIXTRAN** runs a one-part or an amount-only model (**AMOUNT**). A commonly applied rule of thumb to determine whether a given dietary component is consumed ubiquitously or episodically can be determined as follows: If more than 5% of the analytic population reports zero intake of a given dietary component (i.e., nonconsumption on a given day), then that dietary component would be considered episodically consumed ([53](#)). In cases where less than 5% of the analytic population report zero intake, analysts may proceed with the amount-only models. However, it is standard to replace zero intake values with one-half the minimum amount consumed before running the amount-only model. This calculation is performed within the **MIXTRAN Macro**. Additionally, other cutoffs for zero intake, such as 10%, have been used to distinguish between ubiquitous or episodic consumption of a given dietary component. For dietary components that fall within the range of 5%–10% of the population reporting zero intake, analysts may wish to run both the one-part model or amount-only model and the two-part model and see if the results are materially different. Unfortunately, because of the complex survey design, it is not appropriate to try to select the best-fitting model based on typical output parameters, such as the Akaike information criterion (AIC) or $-2 \log\text{-likelihood}$, where lower values are indicative of better fit ([53](#)). Prior analyses have suggested that in the majority of these cases, the two-part model is preferred ([52](#)).
- For a *single dietary component that is consumed episodically*,

MIXTRAN runs either a correlated (**CORR**) or noncorrelated (**NOCORR**) two-part model.

MIXTRAN first implements a mixed-effects logistic model to estimate the propensity to consume (e.g., the outcome is binary, modeling consumption or nonconsumption of a given dietary component on sampled days as realizations of Bernoulli random variables with success probabilities that differ across individuals). The second part of the model is a mixed-effects linear model to estimate the amount of the dietary component consumed. Users can specify whether the probability to consume is correlated with the amount consumed (**CORR** model) or not (**NOCORR** model). The correlation is between the person-specific random effects in both parts of the mixed-effects models: the probability part and the amount part. The probability to consume and the amount consumed may also be linked by the inclusion of a shared set of covariates in each model (53). Conceptually, the model equates usual intake to the product of an individual's propensity to consume and the usual amount consumed by the individual on consumption days. In other words, people who eat a certain type of food tend to eat more of it because they like it or believe consumption is good for them. For example, people who eat whole grains are likely to eat larger amounts when they eat them. The same is true for less frequently consumed foods and small consumption amounts. If the model estimates a direct relationship between consumption amount and probability, then the estimated usual intake distribution will have more extreme higher values (i.e., a longer right tail) than if the model estimates no relationship, because in the former case, high probabilities will be more likely to be multiplied by large amounts and small probabilities by small amounts.

In deciding whether to run a correlated or uncorrelated two-part model

for episodically consumed nutrients or foods, analysts might consider several factors. First, it may be clear due to existing literature that frequent consumers of a given dietary component tend to consume greater amounts of that nutrient or food (5). Second, in cases where it may not be known whether the probability to consume and the amount consumed are correlated, analysts may wish to have a consistent method or standard to apply across all groups and dietary components. One such example would be if there are many nutrients and foods being estimated across multiple demographic groups or strata (52) where the correlation may only be evident in some subgroups or for some dietary components. In this case, it would be more appropriate to allow for the potential correlation between the probability to consume and the amount consumed parts of the models across all of the components and groups, than to assume they are not correlated. However, the correlated models are more computationally intensive, and will thus require more time to execute than the uncorrelated models. For details regarding an empirical approach to choosing a correlated or uncorrelated two-part model for episodically consumed foods, see [Appendix VII](#).

Estimating usual intake distributions and DRIs

After the implementation of the **MIXTRAN Macro**, the **DISTRIB Macro** uses parameter estimates from **MIXTRAN** and a Monte Carlo method to estimate the distribution of usual intake for a given food or nutrient. Cut points can be specified to estimate the proportion of the population consuming less than or greater than a prespecified value, such as EAR or UL. Additional calculations must be performed to obtain standard errors and confidence intervals for the mean and percentiles from the usual intake distribution obtained from **DISTRIB**. [Appendix V](#) includes the macro **BRR_PVALUE_CI Macro**, which demonstrates how to use the estimates obtained from each BRR replication to produce estimated standard errors and confidence intervals.

The **DISTRIB Macro** requires various parameters, unrelated to the variables that are generated and used in the modeling procedures, defined by the user, and can therefore vary from user to user. As a result, differences can occur in the estimation of usual intake of the same data. For example, differences in the number of repetitions or the seed number for the random number generator used in the Monte Carlo simulation can result in different estimates for the same data set. Analysts wishing to replicate their results should set a seed, or implement a reproducible process for varying seeds, for all of their analyses for consistency. See the NCI Method User's Guide (version 2.1) for additional details about the **DISTRIB Macro** (32).

Ratios

When calculating the percentage of energy from fat, there are two ways to calculate the ratio. Estimates may not differ considerably from one another, but the choice of one over another depends on the question of interest. Although the ratio of usual intakes is typically preferred, as it more closely reflects the long-term intake of two dietary components, the estimation of the usual intake of ratios is simpler, as it is based on the univariate distribution of the ratio of daily intakes (54). In practice, both estimates tend to be very close. Freedman *et al.* (54) describes differences of less than one percentage point between the ratio of usual intakes and the usual intake of ratio estimates, examining the percentage of energy from saturated fat, and the percentage of energy from total fat. Freedman *et al.* (54) also provides guidance for quantifying the expected difference between the two methods based on the within-person correlations of each of the two dietary components and the within-person correlation between the two dietary components.

1. The **mean ratio** is determined by calculating the ratio for each individual (e.g., fat kilocalories/total kilocalories) and then taking the mean of the ratios using standard survey data analysis software, such as SAS proc SURVEYMEANS or SUDAAN proc DESCRIPT.

2. The population ratio or **ratio of means** is where the mean fat and mean energy are calculated for the population, and then mean fat is divided by mean energy (55).

For a ratio of usual intakes of two dietary components consumed daily (e.g., percentage of energy from fat), a set of SAS macros are available to examine the bivariate distribution. First, the **NLMIXED UNIVARIATE Macro** is used to fit a univariate measurement error model. This macro is implemented twice, once for the numerator (e.g., kilocalories from total fat) and once for the denominator (e.g., total kilocalories) of the ratio. Initial parameter estimates obtained from each of these two runs of the **NLMIXED UNIVARIATE Macro** are then used in the **NLMIXED BIVARIATE Macro**, which fits a bivariate measurement error model for both of the dietary components simultaneously.

Similar to the process for estimating the distribution of a single dietary component described above, the parameter estimates from the **NLMIXED BIVARIATE Macro** are then used to create a Monte Carlo distribution of the usual intakes of two dietary components using the **DISTRIB BIVARIATE Macro**. This distribution can then be used to estimate the mean and percentiles of the population distribution of the ratio of usual intake using the **PERCENTILES_SURVEY Macro**. The **PERCENTILES_SURVEY Macro** allows users to specify a cut point (e.g., a DRI value expressed as a proportion) and will return the probability that the ratio of usual intakes is less than or equal to the specified DRI. For example, using this macro, analysts can estimate the proportion of the population that consumes 30% or less than their total kilocalories in the form of fat by specifying the cut point as 0.30 in the implementation of the **PERCENTILES_SURVEY Macro**. SAS code for implementing the above macros are available in [Appendix V](#).

At this time, methods to examine the ratio of two episodically consumed dietary components, or between a ubiquitously consumed and episodically

consumed dietary component, have not yet been published. If the usual intake ratio is the estimate of interest rather than the ratio of usual intakes, it can be calculated at the individual-day level and then modeled using **MIXTRAN** and **DISTRIB**, as described above.

An additional consideration to weigh in choosing which value to report is that the usual intake of a ratio (equivalent to mean ratio) is more straightforward to estimate by using the macros **MIXTRAN** and **DISTRIB**, which are appropriate for estimating univariate distributions. In contrast, the ratio of usual intakes requires estimating the univariate measurement error model (**NLMIXED UNIVARIATE**) for each of the nutrients (e.g., fat kilocalories and total kilocalories) and subsequently estimating the bivariate measurement error model (**NLMIXED BIVARIATE**) for the two dietary components simultaneously (54).

Additional consideration when using SAS macros:

Transformations

Because food or nutrient intake data are typically skewed, an additional macro, **BOXCOX_SURVEY Macro**, transforms a variable to approximate normality by applying a Box–Cox transformation. The **MIXTRAN Macro** described above applies a Box–Cox transformation within the macro itself, so this extra step is not entirely necessary. However, analysts may find it helpful to examine the transformation parameters from the **BOXCOX_SURVEY Macro** to determine if results may be sensitive to the particular transformation applied. Specifically, prior simulation studies have shown that various methods for estimating usual intake, including early versions of the NCI Method, may not perform well (i.e., may have larger mean bias) under certain scenarios where the lambda value obtained from the Box–Cox transformation is less than approximately 0.15 (8).

Convergence

Depending upon the complexity of the models, the distribution of the data and the size of the subgroups, along with

other factors, there are instances where one or more BRR replicates will fail to converge. There are several steps that analysts can take in these instances.

Assuming the base run (replicate zero) is successful, and only a few replicate runs failed to converge, those replicates can be excluded from the **BRR_PVALUE_CI Macro**, and the degrees of freedom can be adjusted downward by subtracting the number of failed replicates. Another option is to adjust the starting values used in the **MIXTRAN**, **DISTRIB**, or the **NLMIXED Macros** or to change some of the default **NLMIXED** options, such as the maximum number of iterations or quadrature points used in the estimation procedures. See [Appendix V](#) for examples. Analysts should also consult the appropriate SAS manuals, documentation, and other SAS-support resources relating to mixed models (56, 57).

Application of NCI Method to estimate usual intake

For usual intake estimation, version 2.1 of the **MIXTRAN** and **DISTRIB Macros** were used, with poststratified BRR weights based on the Day 1 dietary weight, to account for day of the week, nonresponse, noncoverage, and unequal probabilities of selection, as well as survey design variables.

For calcium, energy, fat, and the usual intake of ratios for the percentage of energy from fat, one-part **AMOUNT-only** models were used, as less than 1% of the population had no reported consumption over the 2 days of recall. Because of the episodic nature of fruit consumption (roughly 25% of respondents did not report fruit consumption on both days of recall), a two-part (**PROBABILITY** and **AMOUNT**) correlated (**CORR**) model was chosen, which allows for the probability to consume an item to be correlated with the amount consumed. However, convergence failures with the correlated model necessitated the use of the simpler uncorrelated two-part model (**NOCORR**). Usual intake with stratified models (by age and sex) were estimated to allow the macros to find a best fit transformation of the data.

Therefore, no covariates were included in these analyses. For replicability of results, the following parameters were defined for the **DISTRIB Macro**: seed = 5454768 and nsim_mc = 1000. However, for general analytic purposes, where exact replication is not strictly necessary, the seed for each **DISTRIB** run should change. For a detailed discussion about the parameters included in **MIXTRAN** and **DISTRIB**, see the User's Guide (32).

Statistical analyses were conducted using SAS version 9.3 (SAS Institute Inc.) for usual intake analysis and SUDAAN version 11 for all other analyses. For the current analysis, population means, percentiles, and standard errors calculated three ways are reported, using: (a) given day, (b) WPM, and (c) the NCI Method. Estimates are presented by age and sex. For given day and WPM calculations, Day 1 dietary weights were used to account for day of the week, nonresponse, noncoverage, and unequal probabilities of selection. Standard errors for given day and WPM estimates were estimated with Taylor series linearization to account for the complex, stratified sampling design.

Results

Table 2 presents given day, WPM, and the NCI Method estimates for calcium. The estimated means are within 1 standard error of each other for all three methods, except for males aged 12–19. Percentiles at the extremes of estimated intake vary and there is a tendency for estimated intake based on given day or WPM to be lower than the NCI Method at percentiles below the 50th percentile (median) and greater than the NCI Method for percentiles above the median. There is also a tendency for the difference between estimated intakes based on the NCI Method and given day to be larger than the difference between the NCI Method and WPM. For illustrative purposes, **Figure 1** shows this for men aged 40–59. A similar pattern can be seen when comparing estimates of the mean and percentiles for total fat (**Table 3**) and energy (**Table 4**).

There was also variation in assessment of DRIs according to method. For example, the percentage of women aged 20–39 identified as consuming below EAR for calcium by given day, WPM, and the NCI Method were 51%, 48%, and 42%, respectively (**Figure 2** and **Table 2**). Similar patterns can be observed among the percentage of the population above UL, where given day and WPM overestimate the percentage above UL compared with the NCI Method. For example, among men aged 40–59, 5%, 4%, and 1% were identified as consuming above UL for calcium using given day, WPM, and the NCI Method, respectively (**Table 2**). The patterns in relation to these DRIs were not as consistent by age and sex as patterns of the distribution of intake.

Table 5 presents estimates of the proportion of energy from total fat by five methods, because as noted in the **Methods** section, there are two ways to calculate ratios—mean ratio or a ratio of means—and three methods of estimation—given day, WPM, and the NCI Method. As for the other nutrients, the means are very similar according to method. However, unlike the single nutrients, there was somewhat more variability in estimated intake among the

ratio methods, though only at percentiles below the median. For percentiles above the median, estimated intake was more similar according to method. As shown with the single nutrients, the NCI Method reduces the spread of the distribution, thus the proportion below the bottom threshold or above the top threshold is reduced. This is apparent when examining the ratio of the percentage of energy from fat within AMDR, where intake according to nonusual intake methods is generally underestimated. For example, among boys aged 6–11 years, the estimates of the proportion within AMDR are 49%, 58%, and 52% for given day, ratio of mean WPM, and WPM mean ratio, respectively, and 70% and 78% for ratio of mean NCI Method and NCI Method mean ratio, respectively (**Table 5**).

Table 6 presents estimates of fruit consumption for NHANES respondents aged 2 years and over from 2009 through 2012. As with the nutrients examined above, the means for all three methods are similar. However, SUDAAN, similar to SAS survey procedures, is unable to estimate any quantile that is less than or equal to the percentage of data accounted for by the 0 values, thus SUDAAN was unable to calculate the 5th, 10th,

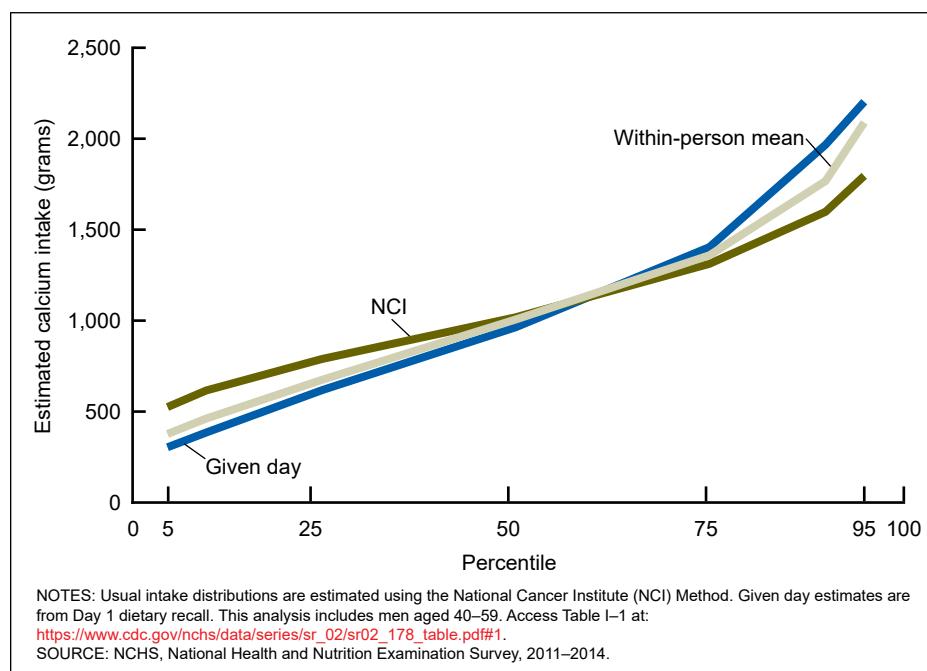


Figure 1. Estimated distributions of calcium using three different methods for men aged 40–59: National Health and Nutrition Examination Survey, 2011–2014

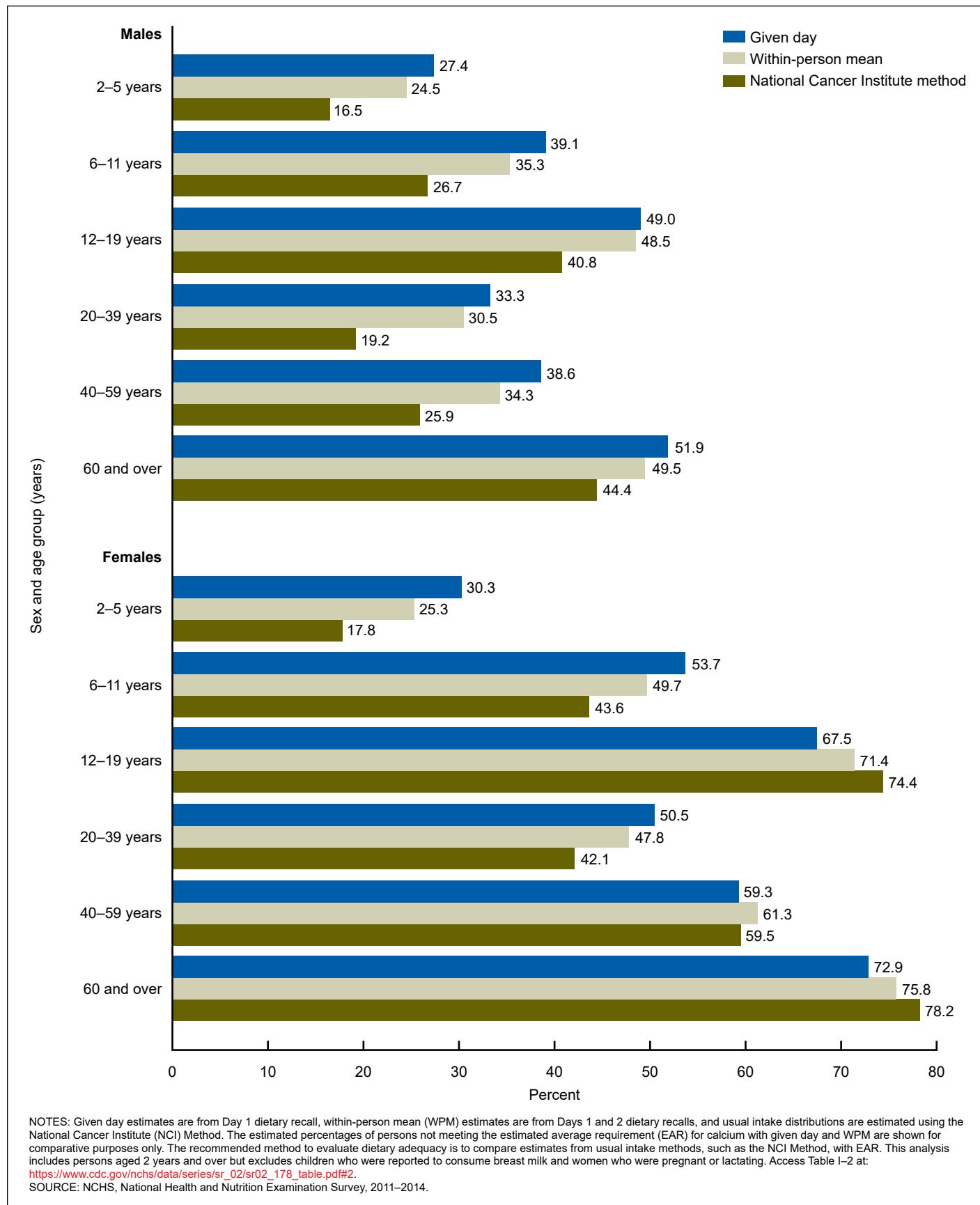


Figure 2. Percentage of U.S. population with calcium intake below estimated average requirement using different estimation methods, by age and sex: National Health and Nutrition Examination Survey, 2011–2014

and sometimes the 25th percentile of intake for fruit consumption for some age groups. Where comparison between methods was possible, usual intake estimates for percentiles approximately below the 50th percentile were greater than given day and WPM estimates, while usual intake estimates for percentiles approximately above the 50th percentile tended to be smaller than given day or WPM estimates (Figure 3). This indicates compression of the tails of the distribution toward the mean, as seen with calcium, fat, energy, and percentage of energy from fat. For example, among boys aged 6–11 years, the estimated usual intake of fruit at the 25th percentile was 0.1, 0.4, and 0.8 cup equivalents using given day, WPM, and the NCI Method, respectively. Among girls aged 2–5 years, estimates at the 90th percentile were 2.9, 3.1, and 2.4 fruit cup equivalents using given day, WPM, and the NCI Method, respectively (Table 6).

In addition to differences observed in the percentile estimates, the estimated percentage below the recommended minimum of consumption and above the recommended maximum consumption varied according to usual intake or nonusual intake methods. Except for the youngest age group (2–5 years), the percentage consuming less than the recommended intake was underestimated using nonusual intake methods compared with the NCI Method (Figure 4). In general, the NCI Method compresses the distribution toward the mean. Therefore, by comparison, nonusual intake methods will tend to overestimate the percentage of the population below the lower threshold and above the upper threshold. A different pattern was observed for usual fruit consumption. For example, for females aged 12–19 years, given-day methods estimated 45% consumed less than the recommended intake, WPM estimated 64%, and the NCI Method estimated 91%. Nonusual intake methods underestimated the percentage below the recommended minimum consumption level. As previously observed with other nutrients, the percentage of persons consuming above the recommended range tended to be overestimated when using nonusual intake methods (e.g.,

among women aged 60 and over, 18%, 17%, and 11% were found to consume fruit above the recommended level, using given day, WPM, and the NCI Method, respectively) (Table 6).

Discussion

This report supplements existing information about the use of the NCI Method by providing guidance to analysts seeking to estimate the usual intake of dietary components using NHANES data. There are various analytic and programming considerations that an analyst may encounter when applying the NCI Method to NHANES dietary data, such as how to create and use BRR weights to account for the complex survey design. This report provides a description of these considerations as well as syntax and instructions to help analysts.

When implementing the NCI Method with NHANES dietary data, there are many analytic decisions and considerations to make that are specific to the question of interest and population under examination. For example, different populations have different consumption patterns, so running an

amount model (**AMOUNT**) in one population (e.g., adults) may require an episodic model (**CORR** or **NOCORR**) in another (e.g., children). The probability to consume and amount consumed might be correlated for some foods in some populations, but they may be uncorrelated in others. Variation in dietary intake patterns within a population could also necessitate different models and transformations by age or sex. For example, the lambda value of the Box-Cox transformation for a single nutrient may vary by age and sex. Rather than selecting a single fixed transformation parameter, it is better to conduct a stratified analysis, allowing for different transformation parameters by strata. This will result in better approximations of normality for each subgroup or stratum of interest, which is important because highly skewed distributions may result in convergence problems and increase bias (8).

The shape of the distribution is critical for determining the proportion of a population at risk for deficiency or excess. Using the NHANES dietary data more simplistically (e.g., by using distributions from given day and WPM methods) to estimate usual intake can lead to erroneous conclusions about the magnitude of a public health problem,

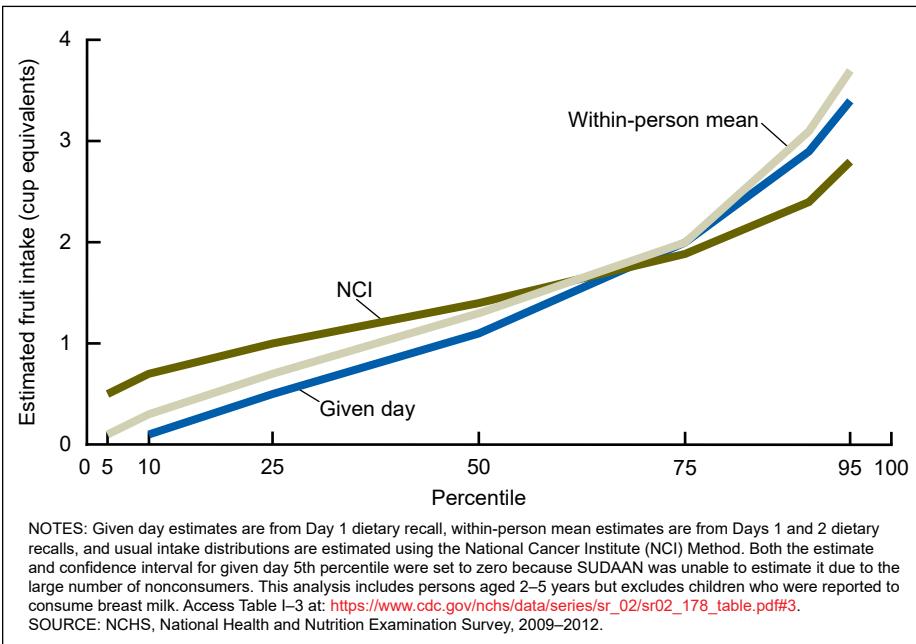


Figure 3. Estimated distributions of fruit intake using three different methods for girls aged 2–5 years: National Health and Nutrition Examination Survey, 2009–2012

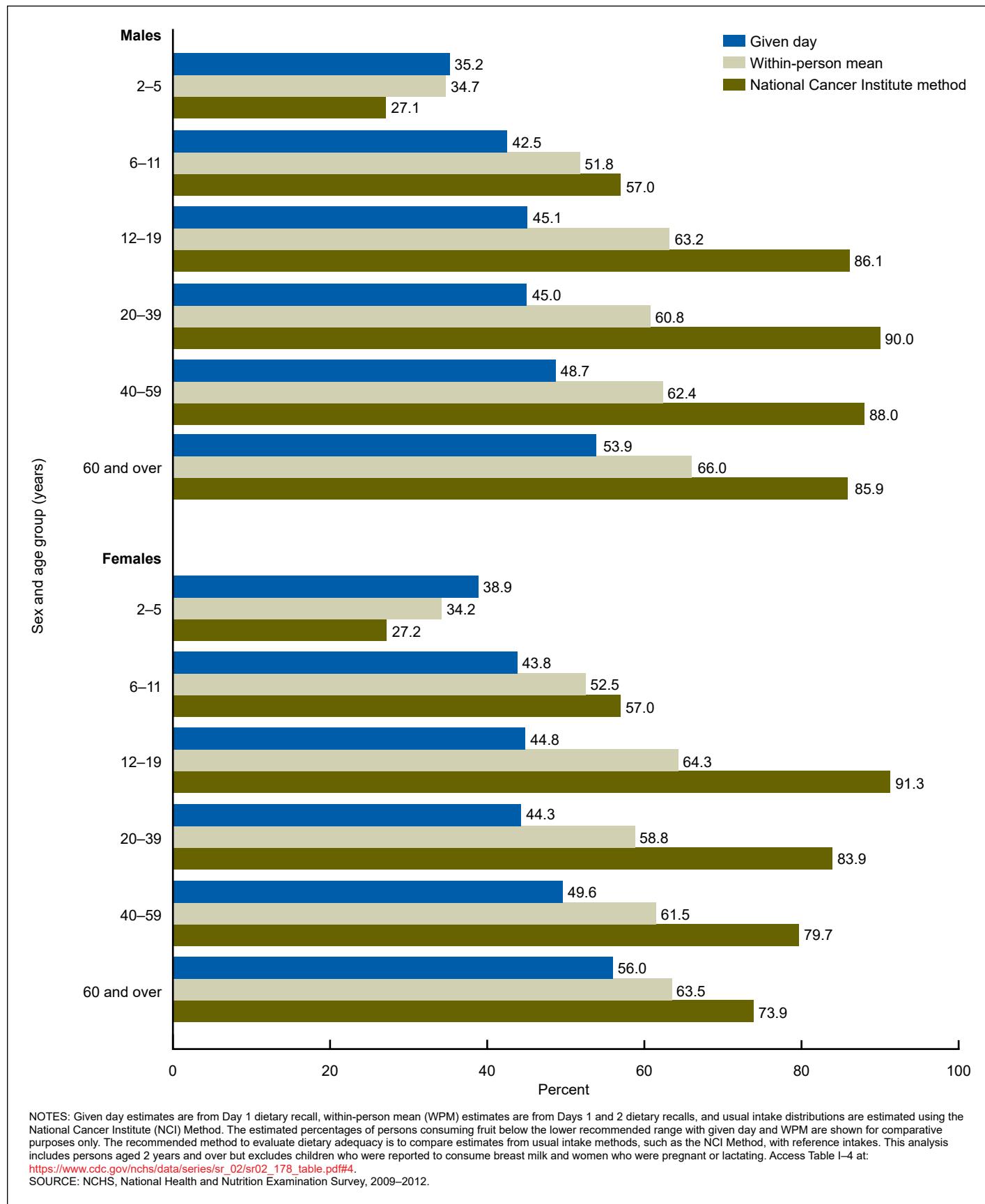


Figure 4. Percentage of U.S. population with fruit intake below the lower recommended range, by age and sex: National Health and Nutrition Examination Survey, 2009–2012

such as inadequate or excessive intake of certain key nutrients. Using usual intake methods, such as the NCI Method with NHANES data, can lead to more appropriate conclusions about population dietary intake patterns, which could facilitate program and intervention planning by providing more accurate measurement of population dietary intake relative to DRIs.

Similar to previous analyses comparing various methods for estimating dietary intake (5,8,9,28,31), this report demonstrates that overall, the means were approximately the same, irrespective of the method used to estimate intake. This is not surprising, since the population mean estimated using the NCI Method was designed to be consistent with estimates using a single 24-hour recall per individual. Similar to other reports (5,9,31,54,58), this report also shows that the NCI Method compresses the distribution toward the mean. For ubiquitously consumed items, such as calcium, energy, fat, and the percentage of energy from fat, the compression of the distribution toward the mean was most apparent when examining the proportion of the population below EAR or above UL. In this situation, methods that did not account for intra-individual variability, namely given day or WPM, tended to overestimate the percentage with the most extreme values of the distribution for ubiquitously consumed nutrients (5,8,9,23,28). Analysts seeking to evaluate the adequacy of diet (i.e., food and nutrient intake) in relation to reference standards using NHANES dietary data should use usual intake methods. The estimates based on single day and WPM in relation to DRI show how the magnitude of dietary insufficiency or excess could be overestimated if usual intake methods are not used (16–19).

This report also examined the impact of estimating an episodically consumed food (e.g., fruit in cup equivalents) using 2009–2012 NHANES FPED files. Similar to the findings for ubiquitously consumed nutrients, estimates of means were similar with all three methods of estimation. Additionally, this report

showed that the NCI Method compressed the distribution of consumption toward the mean for fruit, an episodically consumed food (4,8,9,12) but only for the right side of the distribution. The left side of the distribution is inflated by the large number of nonconsumers, and this changes the distribution such that the percentage below the lower threshold is underestimated using nonusual intake methods. This report only evaluated one example of an episodically consumed food in the U.S. population. Different patterns between results for the usual intake, given day, and WPM methods could occur for episodically consumed nutrients in different populations that have a different distribution than the example used here. The inflation of the left side of the distribution may also be due to the fact that practical considerations prevented the implementation of the correlated NCI model. If there truly was a significant positive correlation between probability and amount, the single-day and WPM distributions would reflect the correlation somewhat, while the uncorrelated NCI Method would not estimate enough people with more extreme higher values of intake (i.e., a long enough right tail of the distribution). The proportion of the population who reported no consumption of fruit over the 2 days of recall may also reflect a large number of nonconsumers of fruit. A similar left inflation in the distribution of estimated intake for a food with a large number of nonconsumers has been reported elsewhere (8). The expectation that given day and WPM methods tend to overestimate the extreme values of the distribution does not necessarily apply for episodically consumed foods where there is a large percentage of nonconsumers.

NHANES uses specific age, sex, and race and Hispanic-origin groupings as part of the standard sampling frame. The age categories in the sampling frame differ from those used by the National Academies of Science, Engineering and Medicine (formerly the Institute of Medicine) to establish DRIs (17). The difference between the two age categorizations would likely have minimal impact on estimates of

usual intake using the NCI Method. However, analysts should be aware of the difference. While it is often standard to report results specific to the NHANES sampling frame-based age groups (e.g., 2–5, 6–11, and 12–19 years), the use of the DRI age and sex subgroups is more straightforward if the goal of the analysis is to calculate population proportions at risk of deficiency or excess.

Limitations

These analyses were limited to nutrients (calcium, energy, fat, and the percentage of energy from fat) that were consumed by nearly all NHANES participants every day and a single food item, fruit, for which 25% of the population reported no consumption. These two types of analyses could be described as the most straightforward situations for estimating usual intake, because their distributions were approximately normal or could be transformed to normality. Foods and nutrients that are infrequently consumed, such as nuts or specific fruits and vegetables, can present problems during analysis, because the distribution is skewed with a relatively larger proportion of nonconsumers or heavy consumers. If a Box–Cox transformation is not sufficient to transform the distribution of observed intakes to approximate normality (e.g., for multimodal or extremely heavy-tailed distributions), a key model assumption is violated, and the NCI Method may not be able to overcome this limitation.

Computing power and computational intensity is another consideration when using NHANES data to estimate usual intake with the NCI Method. As the complexity of the research question increases, so does the time it takes to run a model. Running these models on a single computer requires considerable processing power, and even with small age and sex groups, the time to return a distribution for a simple **AMOUNT** model was often measured in minutes (10–15 on average on a 32GB RAM machine) for two survey cycles of NHANES data, roughly 16,000 respondents. As models became more

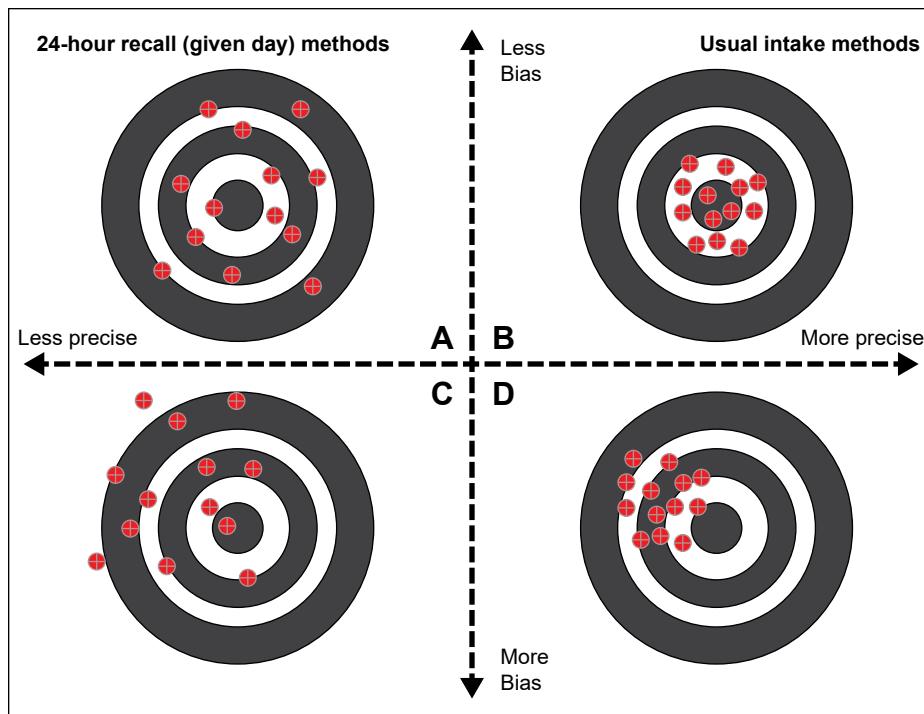


Figure 5. The relationship between bias and precision when comparing given day with usual intake methods for estimating intake

complex, for example, correlated fruit models for single age and sex groups, it was necessary to run models on a computer platform with increased processing power and speed. Despite this, it often took hours to run a single age and sex model. Analysts could explore other ways to reduce processing time depending on available statistical computing resources.

Another aspect that deserves mention is the need to scrutinize SAS logs to troubleshoot when the macros do not work as expected. It is not uncommon for some of the weight replicates to fail to converge. If this happens for a small number of replicates (e.g., roughly 3–4 replicates out of 32 for an NHANES data set that includes 4 years), it is reasonable to delete these replicates and estimate standard errors with the remaining replicates (reducing the degrees of freedom when calculating standard errors). If more than a small number fail to converge, it is necessary to look more closely at why convergence is a problem and explore alternative solutions.

Importantly, the NCI Method does not address systematic error or bias in self-reported dietary intake. Estimates of

certain dietary components, energy intake most prominently, have long been shown to be misreported in 24-hour recalls. The degree of misreporting varies by dietary component and population characteristics (12). Thus, it is well established that the unbiasedness assumption underlying the NCI Method (and other methods of estimating dietary intake) is likely violated in many circumstances. While usual intake estimates are more precise than estimates based on a single day of recall, residual systematic error may not lead to a more accurate or less biased measure (59) (Figure 5). Ideally, methods to estimate usual or habitual dietary intake would produce results characterized by panel B in Figure 5. Realistically, panel D shows the results of applying these methods: an increase in precision, but bias is not addressed. Results of any analyses of dietary intake should be interpreted cautiously and appropriately in the context of known limitations related to the data and methods (12).

Finally, this report demonstrates two of the most common and straightforward examples of using NHANES data to estimate usual intake of a given dietary

component with the NCI Method—describing food and nutrient distributions in a population and identifying the proportion of a population consuming above or below a threshold. There are other applications of the NCI methodology that are beyond the scope of this report, such as modeling intake of more than two potentially correlated dietary components through multivariate measurement error models and outputting conditional mean intakes for use in modeling associations between dietary intake and health outcomes. Further discussion of these other applications and related methods is provided elsewhere (12,24,34,36,60).

Summary

This report describes how to adapt the NCI Method to estimate usual intake distributions for foods and nutrients using NHANES dietary data. Detailed SAS coding and corresponding output tables are provided to data users to replicate these analyses and implement the NCI Method with NHANES data.

This report shows that the means for calcium, fat, energy, percentage of energy from fat, and fruit using NHANES dietary data were similar across all three approaches. For this reason, population estimates of mean intake using NHANES can reasonably be estimated using given-day methods. In this situation, usual intake methods offer no advantages over the simpler method, while increasing complexity. In contrast, the distributions estimated by given day and WPM methods were wider than those estimated by the NCI Method. When using NHANES data to estimate the proportion of the population above or below a certain limit for adequate or excess nutrient or food intake, if the assumptions of the NCI Method are satisfied, the wider distributions of the WPM or given-day methods can overestimate or underestimate these proportions depending on whether the nutrient or food is ubiquitously or episodically consumed. This leads to biased estimates of the proportion of the population at risk of insufficiency or

excess. For distributions and comparisons with DRIs, usual intake methods are necessary to reduce within-person variation and produce a more precise estimate when using NHANES dietary data. Beyond analytical applications, program and intervention planning that use NHANES data based on methods that do not use usual intake estimates may be based on imprecise estimates of the magnitude of various population dietary shortfalls or excesses.

References

1. Ahluwalia N, Dwyer J, Terry A, Moshfegh A, Johnson C. Update on NHANES dietary data: Focus on collection, release, analytical considerations, and uses to inform public policy. *Adv Nutr* 7(1):121–34. 2016.
2. Zipf G, Chiappa M, Porter KS, Ostchega Y, Lewis BG, Dostal J. National Health and Nutrition Examination Survey: Plan and operations, 1999–2010. National Center for Health Statistics. *Vital Health Stat* 1(56). 2013.
3. National Academies of Sciences, Engineering, and Medicine. Review of WIC food packages: Improving balance and choice: Final report. Washington, DC: National Academies Press. 2017.
4. Gibson RS. Principles of nutritional assessment. 2nd ed. New York, NY: Oxford University Press. 2005.
5. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc* 106(10):1575–87. 2006.
6. Willett W. Nutritional epidemiology. New York, NY: Oxford University Press. 1998.
7. Tooze JA, Grunwald GK, Jones RH. Analysis of repeated measures data with clumping at zero. *Stat Methods Med Res* 11(4):341–55. 2002.
8. Souverein OW, Dekkers AL, Geelen A, Haubrock J, de Vries JH, Ocké MC, et al. Comparing four methods to estimate usual intake distributions. *Eur J Clin Nutr* 65 Suppl 1:S92–101. 2011.
9. Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: A review of the theory. *Journal Am Diet Assoc* 106(10):1640–50. 2006.
10. National Cancer Institute. Measurement error webinar series: Webinar 2, estimating usual intake distributions for dietary components consumed daily by nearly all persons. Available from: <https://epi.grants.cancer.gov/events/measurement-error/#session>.
11. Kirkpatrick SI, Collins CE. Assessment of nutrient intakes: Introduction to the special issue. *Nutrients* 8(4):184. 2016.
12. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhofer ML, et al. Addressing current criticism regarding the value of self-report dietary data. *J Nutr* 145(12):2639–45. 2015.
13. Satija A, Yu E, Willett WC, Hu FB. Understanding nutritional epidemiology and its role in policy. *Adv Nutr* 6(1):5–18. 2015.
14. Hébert JR, Hurley TG, Steck SE, Miller DR, Tabung FK, Peterson KE, et al. Considering the value of dietary assessment data in informing nutrition-related health policy. *Adv Nutr* 5(4):447–55. 2014.
15. U.S. Department of Health and Human Services and U.S. Department of Agriculture. Dietary guidelines for Americans 2015–2020. 8th ed. 2015.
16. Barr SI, Murphy SP, Poos MI. Interpreting and using the dietary reference intakes in dietary assessment of individuals and groups. *J Am Diet Assoc* 102(6):780–8. 2002.
17. Institute of Medicine of the National Academies. Dietary reference intakes: The essential guide to nutrient requirements. Washington, DC: The National Academies Press. 2006.
18. Murphy SP, Barr SI, Poos MI. Using the new dietary reference intakes to assess diets: A map to the maze. *Nutr Rev* 60(9):267–75. 2002.
19. Murphy SP, Guenther PM, Kretsch MJ. Using the dietary reference intakes to assess intakes of groups: Pitfalls to avoid. *J Am Diet Assoc* 106(10):1550–3. 2006.
20. Hoffmann K, Boeing H, Dufour A, Volatier JL, Telman J, Virtanen M, et al. Estimating the distribution of usual dietary intake by short-term measurements. *Eur J Clin Nutr* 56 Suppl 2:S53–62. 2002.
21. Freedman LS, Midthune D, Carroll RJ, Krebs-Smith S, Subar AF, Troiano RP, et al. Adjustments to improve the estimation of usual dietary intake distributions in the population. *J Nutr* 134(7):1836–43. 2004.
22. Subar AF, Dodd KW, Guenther PM, Kipnis V, Midthune D, McDowell M, et al. The food propensity questionnaire: Concept, development, and validation for use as a covariate in a model to estimate usual food intake. *J Am Diet Assoc* 106(10):1556–63. 2006.
23. Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, et al. Modeling data with excess zeros and measurement error: Application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics* 65(4):1003–10. 2009.
24. Freedman LS, Kipnis V, Schatzkin A, Tasevska N, Potischman N. Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? *Epidemiol Perspect Innov* 7(1):2. 2010.
25. Dekkers AL, Verkaik-Kloosterman J, van Rossum CT, Ocké MC. SPADE, a new statistical program to estimate habitual dietary intake from multiple food sources and dietary supplements. *J Nutr* 144(12):2083–91. 2014.

26. Dodd KW. A technical guide to C-SIDE (software for intake distribution estimation). Dietary Assessment Research Series Report 9. CARD Technical Reports. 1996.
27. Nusser SM, Carriquiry AL, Dodd K, Fuller WA. A semiparametric transformation approach to estimating usual daily intake distributions. *J Am Stat Assoc* 91(436):9. 1996.
28. Tooze JA, Kipnis V, Buckman DW, Carroll RJ, Freedman LS, Guenther PM, et al. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI Method. *Stat Med* 29(27):2857–68. 2010.
29. Zhang S, Krebs-Smith SM, Midthune D, Perez A, Buckman DW, Kipnis V, et al. Fitting a bivariate measurement error model for episodically consumed dietary components. *Int J Biostat* 7(1):1. 2011.
30. Zhang S, Midthune D, Guenther PM, Krebs-Smith SM, Kipnis V, Dodd KW, et al. A new multivariate measurement error model with zero-inflated dietary data, and its application to dietary assessment. *Ann Appl Stat* 5(2B):1456–87. 2011.
31. Laureano GH, Torman VB, Crispim SP, Dekkers AL, Camey SA. Comparison of the ISU, NCI, MSM, and SPADE methods for estimating usual intake: A simulation study of nutrients consumed daily. *Nutrients* 8(3):166. 2016.
32. National Cancer Institute. User's guide for analysis of usual intakes, version 2.1. Available from: https://epi.grants.cancer.gov/diet/usualintakes/Users_Guide_v2.1.pdf.
33. NHANES Dietary Web Tutorial. Advanced dietary analyses. Available from: <https://www.cdc.gov/nchs/tutorials/dietary/advanced/index.htm>.
34. National Cancer Institute. Usual dietary intakes: SAS macros for the NCI Method. Available from: <https://epi.grants.cancer.gov/diet/usualintakes/macros.html>.
35. National Cancer Institute. Usual dietary intakes: SAS macros for analysis of a single dietary component. Available from: https://epi.grants.cancer.gov/diet/usualintakes/macros_single.html.
36. National Cancer Institute. Usual dietary intakes: SAS macros for estimating ratios of two dietary components that are consumed nearly every day. Available from: https://epi.grants.cancer.gov/diet/usualintakes/macros_double.html.
37. CDC. National Health and Nutrition Examination Survey, 2013–2014 data documentation, codebook, and frequencies. Dietary interview—total nutrient intakes, first day (DR1TOT_H). Available from: https://www.cdc.gov/Nchs/Nhanes/2013-2014/DR1TOT_H.htm.
38. Moshfegh AJ, Rhodes DG, Baer DJ, Murayi T, Clemens JC, Rumpler WV, et al. The US Department of Agriculture automated multiple-pass method reduces bias in the collection of energy intakes. *Am J Clin Nutr* 88(2):324–32. 2008.
39. Herrick KA, Rossen LM, Nielsen SJ, Branum AM, Ogden CL. Fruit consumption by youth in the United States. *Pediatrics* 136(4):664–71. 2015.
40. National Center for Health Statistics. NHANES response rates and population totals. Available from: https://www.cdc.gov/nchs/nhanes/response_rates_cps.htm.
41. Curtain LR, Mohadjer LK, Dohrmann SM, Kruszon-Moran D, Mirel LB, Carroll MD, et al. National Health and Nutrition Examination Survey: Sample design, 2007–2010. National Center for Health Statistics. *Vital Health Stat* 2(160). 2013.
42. CDC. National Health and Nutrition Examination Survey, 2011–2012 data documentation, codebook, and frequencies. Dietary interview—total nutrient intakes, first day (DR1TOT_G). Available from: https://www.cdc.gov/Nchs/Nhanes/2011-2012/DR1TOT_G.htm.
43. CDC. National Health and Nutrition Examination Survey, 2011–2012 data documentation, codebook, and frequencies. Dietary interview—total nutrient intakes, second day (DR2TOT_G). Available from: https://www.cdc.gov/Nchs/Nhanes/2011-2012/DR2TOT_G.htm.
44. CDC. National Health and Nutrition Examination Survey, 2013–2014 data documentation, codebook, and frequencies. Dietary interview—total nutrient intakes, second day (DR2TOT_H). Available from: https://www.cdc.gov/Nchs/Nhanes/2013-2014/DR2TOT_H.htm.
45. United States Department of Agriculture. Food patterns equivalents database: Databases and SAS data sets. Available from: <https://www.ars.usda.gov/northeast-area/beltsville-md/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fped-databases/>.
46. Institute of Medicine. Dietary reference intakes: Applications in dietary assessment. Washington, DC: National Academies Press. 2000.
47. Institute of Medicine of the National Academies. Dietary reference intakes for calcium and vitamin D. Washington, DC: National Academies Press. 2011.
48. United States Department of Agriculture. Dietary reference intakes (DRIs). Washington, DC: The National Academies Press. 2000.
49. Institute of Medicine. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein and amino acids. Washington, DC: The National Academies. 2005.
50. CDC. National Health and Nutrition Examination Survey. Dietary Interview—total nutrient intakes, 2009–2010. Available from: https://www.cdc.gov/Nchs/Nhanes/2009-2010/DR1TOT_F.htm#WTDRD1.

51. Moshfegh A, Goldman J, Ahuja J, Rhodes D, LaComb R. What we eat in America, NHANES 2005–2006: Usual nutrient intakes from food and water compared to 1997 dietary reference intakes for vitamin D, calcium, phosphorus, and magnesium. U.S. Department of Agriculture, Agricultural Research Service. 2009. Available from: https://www.ars.usda.gov/ARSUserFiles/80400530/pdf/0506/usual_nutrient_intake_vitD_ca_phos_mg_2005-06.pdf.
52. National Cancer Institute. Usual dietary intakes: Food intakes, U.S. population, 2007–10. Available from: <https://epi.grants.cancer.gov/diet/usualintakes/pop/2007-10/#methods>.
53. National Cancer Institute. Measurement error webinar series: Webinar 3, estimating usual intake distributions for dietary components consumed episodically. Available from: <https://epi.grants.cancer.gov/events/measurement-error/#session>.
54. Freedman LS, Guenther PM, Dodd KW, Krebs-Smith SM, Midthune D. The population distribution of ratios of usual intakes of dietary components that are consumed every day can be estimated from repeated 24-hour recalls. *J Nutr* 140(1):111–6. 2010.
55. Krebs-Smith SM, Kott PS, Guenther PM. Mean proportion and population proportion: Two answers to the same question? *J Am Diet Assoc* 89(5):671–6. 1989.
56. Kiernan K, Tao J, Gibbs P. Tips and strategies for mixed modeling with SAS/STAT procedures. 2012. Available from: <http://support.sas.com/resources/papers/proceedings12/332-2012.pdf>.
57. SAS/STAT(R) 9.2 User's guide. PROC NL MIXED PROCEDURE. Available from: https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_nlmixed_sect009.htm.
58. Carriquiry AL. Estimation of usual intake distributions of nutrients and foods. *J Nutr* 133(2):601S–8S. 2003.
59. National Cancer Institute. Measurement error webinar series: Webinar 1, introduction to the problem of measurement error in dietary intake data. Available from: <https://epi.grants.cancer.gov/events/measurement-error/#session>.
60. Prentice RL, Tinker LF, Huang Y, Neuhouser ML. Calibration of self-reported dietary measures using biomarkers: An approach to enhancing nutritional epidemiology reliability. *Curr Atheroscler Rep* 15(9):353. 2013.
61. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol* 57(1):289–300. 1995.

Table 1. SAS macros developed by National Cancer Institute to run nonlinear mixed models used in this report

Type of dietary component to be modeled	Macro	Function of macro
Single dietary component consumed daily or episodically	MIXTRAN	Fits model to obtain parameter estimates and allows for evaluation of covariate effects.
	DISTRIB	Uses parameter estimates from MIXTRAN and Monte Carlo method to estimate distribution of usual intake for a food or nutrient.
Ratio of two dietary components where desired estimate is ratio of usual intakes	NLMIXED UNIVARIATE	Fits a univariate measurement error model for a single dietary component. Primary purpose is to provide initial parameter estimates for the macro NLMIXED BIVARIATE.
	NLMIXED BIVARIATE	Fits a bivariate measurement error model for two dietary components simultaneously.
	DISTRIB BIVARIATE	Uses parameter estimates from NLMIXED BIVARIATE Macro to estimate distribution of the ratio of usual intakes of two dietary components.
	PERCENTILES_SURVEY	Uses distribution generated by DISTRIB BIVARIATE Macro to calculate mean and percentiles of population distribution of usual intake.
Supplementary macros		
Single dietary component consumed daily or episodically	BOXCOX_SURVEY	Uses Box–Cox transformation to transform variable to approximate normality.
Single dietary component or ratio of two dietary components (ratio of usual intakes)	BRR_PVALUE_CI	Performs balanced repeated replication variance estimation and calculates standard errors, confidence intervals, and <i>p</i> values for user-specified parameters of interest using estimates from DISTRIB macros (DISTRIB and DISTRIB BIVARIATE).

SOURCE: Usual Dietary Intakes: SAS macros for the National Cancer Institute Method (see reference 34 in the report).

Table 2. Calcium (grams per day): Mean, percentile, and percentage of persons meeting dietary reference intakes, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014

Sex, age group (years), and method	Mean (standard error)	5th (standard error)	10th (standard error)	25th (standard error)	50th (standard error)
Males					
2–5					
Given day	1,005.5 (36.4)	257.0 (30.1)	394.8 (34.4)	647.4 (31.0)	899.2 (31.4)
Within-person mean	987.0 (33.0)	379.6 (24.0)	471.1 (26.4)	656.4 (23.5)	904.9 (33.4)
National Cancer Institute Method	982.6 (29.8)	518.2 (20.7)	600.0 (19.6)	754.4 (20.9)	952.1 (28.4)
6–11					
Given day	1,156.2 (25.5)	387.7 (30.4)	523.9 (29.5)	747.3 (25.3)	1,065.5 (31.3)
Within-person mean	1,170.8 (21.4)	483.0 (27.6)	596.4 (20.3)	849.5 (22.7)	1,122.0 (31.3)
National Cancer Institute Method	1,168.6 (21.0)	683.0 (43.0)	774.4 (36.9)	938.2 (25.8)	1,143.1 (19.9)
12–19					
Given day	1,248.5 (31.4)	331.1 (25.2)	447.7 (30.8)	688.0 (31.7)	1,088.5 (24.2)
Within-person mean	1,195.4 (25.9)	423.2 (25.8)	530.1 (21.1)	775.8 (22.1)	1,090.5 (32.7)
National Cancer Institute Method	1,195.7 (27.1)	601.0 (30.0)	703.2 (27.8)	896.4 (24.3)	1,149.1 (24.9)
20–39					
Given day	1,202.4 (29.9)	339.1 (23.1)	455.1 (15.3)	674.3 (19.6)	1,029.1 (25.6)
Within-person mean	1,185.8 (25.5)	412.7 (14.7)	516.3 (15.7)	721.9 (20.2)	1,065.5 (21.0)
National Cancer Institute Method	1,180.7 (22.6)	571.9 (22.9)	672.3 (20.6)	866.3 (16.8)	1,124.9 (19.2)
40–59					
Given day	1,080.1 (24.0)	303.7 (21.4)	384.9 (16.9)	622.1 (23.8)	964.3 (34.0)
Within-person mean	1,071.3 (21.5)	377.0 (16.3)	461.5 (18.7)	671.0 (20.5)	988.6 (19.3)
National Cancer Institute Method	1,075.6 (19.7)	524.8 (23.3)	615.9 (20.0)	791.5 (15.0)	1,026.3 (17.2)
60 and over					
Given day	974.9 (18.6)	263.9 (18.3)	360.1 (15.6)	581.0 (16.7)	858.5 (23.2)
Within-person mean	972.3 (15.8)	355.9 (14.8)	434.0 (16.8)	623.4 (12.7)	879.8 (18.1)
National Cancer Institute Method	974.4 (14.7)	457.0 (21.5)	540.3 (19.1)	704.0 (13.9)	924.9 (11.5)
Females					
2–5					
Given day	937.6 (25.6)	330.8 (26.1)	431.8 (21.2)	632.6 (24.5)	853.4 (30.0)
Within-person mean	935.4 (21.0)	383.4 (22.9)	470.2 (23.8)	634.1 (26.8)	895.0 (25.2)
National Cancer Institute Method	934.0 (24.7)	530.1 (25.3)	601.4 (24.1)	734.6 (22.6)	906.1 (23.7)
6–11					
Given day	982.3 (23.1)	304.1 (24.1)	440.7 (24.8)	624.3 (27.7)	907.6 (31.0)
Within-person mean	1,004.4 (17.7)	430.2 (22.5)	527.3 (21.1)	706.4 (22.9)	941.6 (28.3)
National Cancer Institute Method	1,004.9 (15.3)	661.7 (35.2)	727.2 (30.6)	843.9 (22.1)	988.6 (15.4)
12–19					
Given day	889.9 (22.7)	227.5 (20.3)	298.4 (17.3)	503.5 (27.4)	819.3 (34.5)
Within-person mean	866.6 (20.6)	273.5 (28.3)	358.0 (15.5)	569.1 (18.1)	834.4 (20.3)
National Cancer Institute Method	869.2 (19.8)	424.2 (25.3)	501.1 (23.4)	646.4 (19.9)	836.1 (19.0)
20–39					
Given day	892.1 (20.7)	257.5 (9.60)	333.7 (11.3)	543.3 (21.9)	788.3 (26.6)
Within-person mean	876.2 (16.1)	317.1 (14.5)	406.4 (13.7)	570.6 (13.3)	820.0 (16.4)
National Cancer Institute Method	877.1 (15.4)	499.3 (27.8)	565.4 (24.9)	689.9 (19.2)	850.3 (15.2)
40–59					
Given day	853.1 (13.1)	232.6 (17.5)	329.9 (13.6)	499.1 (14.8)	773.0 (18.9)
Within-person mean	852.0 (12.0)	309.0 (20.7)	397.5 (12.3)	560.4 (9.30)	772.1 (18.3)
National Cancer Institute Method	850.6 (11.4)	430.0 (18.9)	500.0 (17.4)	634.5 (14.0)	813.1 (10.6)
60 and over					
Given day	810.2 (20.6)	245.1 (17.8)	321.8 (19.0)	496.2 (21.4)	724.9 (18.0)
Within-person mean	811.3 (15.5)	319.4 (16.5)	397.1 (14.5)	563.1 (15.6)	750.1 (16.9)
National Cancer Institute Method	810.9 (13.3)	433.1 (15.1)	498.0 (13.8)	621.7 (11.4)	781.7 (11.6)

Table 2. Calcium (grams per day): Mean, percentile, and percentage of persons meeting dietary reference intakes, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014—Con.

Sex, age group (years), and method	75th (standard error)	90th (standard error)	95th (standard error)	Percent less than EAR (standard error)	Percent more than UL (standard error)
Males					
2–5					
Given day	1,271.9 (48.6)	1,743.7 (95.2)	1,992.1 (111.9)	27.4 (2.9)	1.7 (0.7)
Within-person mean	1,226.6 (34.7)	1,614.3 (62.7)	1,794.0 (95.7)	24.5 (2.8)	0.6 (0.3)
National Cancer Institute Method	1,177.7 (41.2)	1,402.9 (56.7)	1,551.8 (67.6)	16.5 (1.7)	0.0 (0.0)
6–11					
Given day	1,462.4 (42.2)	1,867.5 (47.6)	2,178.1 (99.1)	39.1 (1.9)	1.0 (0.4)
Within-person mean	1,420.6 (34.6)	1,785.9 (57.6)	2,005.7 (82.3)	35.3 (2.2)	1.3 (0.6)
National Cancer Institute Method	1,371.1 (36.0)	1,596.4 (61.9)	1,742.3 (80.9)	26.7 (2.4)	0.0 (0.1)
12–19					
Given day	1,632.6 (59.4)	2,275.8 (75.4)	2,677.2 (118.1)	49.0 (1.9)	4.6 (1.2)
Within-person mean	1,481.1 (39.2)	1,958.2 (83.7)	2,467.9 (128.7)	48.5 (2.2)	1.7 (0.9)
National Cancer Institute Method	1,444.5 (37.6)	1,749.4 (58.8)	1,949.3 (75.0)	40.8 (2.9)	0.1 (0.1)
20–39					
Given day	1,538.1 (33.5)	2,124.6 (70.7)	2,441.5 (125.8)	33.3 (1.5)	4.8 (0.9)
Within-person mean	1,462.3 (28.6)	1,952.7 (78.0)	2,342.4 (107.6)	30.5 (1.6)	3.6 (0.9)
National Cancer Institute Method	1,434.0 (34.8)	1,760.3 (58.4)	1,980.2 (76.7)	19.2 (1.5)	0.8 (0.4)
40–59					
Given day	1,400.9 (34.2)	1,966.8 (51.1)	2,203.1 (67.3)	38.6 (2.0)	5.2 (0.8)
Within-person mean	1,363.7 (37.3)	1,767.2 (57.3)	2,089.1 (78.4)	34.3 (2.0)	3.6 (0.8)
National Cancer Institute Method	1,305.9 (31.5)	1,598.1 (50.6)	1,795.8 (64.3)	25.9 (1.5)	1.2 (0.4)
60 and over					
Given day	1,254.7 (24.5)	1,698.7 (59.7)	2,053.3 (58.0)	51.9 (1.8)	6.0 (0.8)
Within-person mean	1,211.3 (22.3)	1,579.9 (40.2)	1,880.4 (90.2)	49.5 (1.9)	4.3 (0.7)
National Cancer Institute Method	1,190.4 (25.5)	1,472.4 (47.7)	1,662.9 (65.4)	44.4 (1.3)	1.3 (0.5)
Females					
2–5					
Given day	1,158.4 (35.8)	1,533.8 (90.3)	1,813.7 (81.1)	30.3 (2.6)	0.9 (0.4)
Within-person mean	1,145.3 (43.3)	1,480.0 (46.9)	1,648.8 (42.9)	25.3 (2.2)	0.2 (0.1)
National Cancer Institute Method	1,102.8 (32.1)	1,301.7 (47.3)	1,434.4 (59.6)	17.8 (2.5)	0.0 (0.0)
6–11					
Given day	1,259.4 (42.5)	1,646.3 (58.4)	1,839.4 (78.9)	53.7 (2.4)	0.9 (0.6)
Within-person mean	1,262.6 (25.9)	1,564.4 (43.4)	1,820.1 (51.0)	49.7 (2.4)	0.0 (0.0)
National Cancer Institute Method	1,147.3 (22.0)	1,304.8 (37.8)	1,405.3 (49.4)	43.6 (2.6)	0.0 (0.0)
12–19					
Given day	1,216.1 (45.6)	1,519.7 (32.9)	1,724.1 (60.6)	67.5 (2.1)	0.2 (0.2)
Within-person mean	1,108.4 (31.9)	1,395.2 (42.7)	1,616.9 (59.6)	71.4 (2.3)	0.2 (0.2)
National Cancer Institute Method	1,055.3 (27.4)	1,280.0 (42.6)	1,427.3 (54.2)	74.4 (2.2)	0.0 (0.0)
20–39					
Given day	1,166.9 (33.1)	1,507.5 (25.7)	1,797.4 (84.7)	50.5 (1.9)	1.1 (0.4)
Within-person mean	1,104.1 (25.6)	1,408.4 (26.1)	1,713.7 (72.1)	47.8 (1.7)	0.2 (0.1)
National Cancer Institute Method	1,034.8 (21.6)	1,223.1 (35.9)	1,347.5 (46.9)	42.1 (2.5)	0.0 (0.0)
40–59					
Given day	1,097.6 (27.4)	1,478.2 (43.4)	1,757.7 (40.7)	59.3 (1.5)	1.7 (0.5)
Within-person mean	1,073.5 (17.4)	1,376.5 (39.9)	1,641.9 (31.9)	61.3 (1.6)	1.1 (0.3)
National Cancer Institute Method	1,025.8 (15.2)	1,249.7 (29.3)	1,400.1 (41.3)	59.5 (1.4)	0.1 (0.1)
60 and over					
Given day	1,044.6 (33.1)	1,372.3 (37.1)	1,643.2 (71.4)	72.9 (1.6)	1.8 (0.6)
Within-person mean	989.5 (23.6)	1,290.1 (29.2)	1,519.1 (37.1)	75.8 (1.8)	0.7 (0.2)
National Cancer Institute Method	968.5 (19.0)	1,160.9 (31.1)	1,288.6 (40.0)	78.2 (2.0)	0.1 (0.0)

0.0 Quantity more than zero but less than 0.05.

NOTES: EAR is estimated average requirement and UL is tolerable upper limit. Given day estimates are from Day 1 dietary recall, within-person mean estimates are from Days 1 and 2 dietary recalls, and usual intake distributions are estimated using the National Cancer Institute Method.

The recommended way to evaluate the adequacy of dietary intake in relation to reference standards is by using Methods that estimate usual intake and therefore account for within-person variation.

Comparing estimates based on given day and within-person mean with reference standards is not recommended. They are presented to illustrate the inaccuracy associated with using methods that do not account for within-person variation (see references 12–15 in the report). Children who were reported to consume breast milk and women who were pregnant or lactating were excluded from the data analysis.

SOURCE: NCHS, National Health and Nutrition Examination Survey, 2011–2014.

Table 3. Total fat (grams per day): Mean and percentile, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014

Sex, age group (years), and method	Mean (standard error)	5th (standard error)	10th (standard error)	25th (standard error)	50th (standard error)	75th (standard error)	90th (standard error)	95th (standard error)
Males								
2–5								
Given day	57.5 (1.4)	27.9 (1.1)	31.1 (0.9)	40.4 (0.8)	52.6 (0.9)	70.5 (2.1)	86.2 (3.2)	100.1 (4.9)
Within-person mean	56.2 (1.2)	26.7 (1.7)	32.8 (1.2)	40.3 (1.0)	52.2 (1.1)	67.0 (1.2)	81.9 (3.2)	94.7 (5.0)
National Cancer Institute Method	56.4 (1.2)	33.2 (1.7)	37.4 (1.4)	45.3 (1.1)	55.1 (1.0)	66.1 (1.9)	77.0 (3.3)	84.1 (4.3)
6–11								
Given day	76.3 (1.6)	32.4 (1.6)	37.9 (1.8)	53.1 (1.2)	68.9 (1.6)	92.4 (3.0)	123.1 (3.6)	143.9 (6.9)
Within-person mean	74.6 (1.4)	35.1 (2.1)	42.7 (1.4)	55.0 (1.5)	68.3 (1.6)	89.0 (2.3)	114.0 (4.1)	127.2 (5.2)
National Cancer Institute Method	74.7 (1.4)	47.0 (2.4)	52.2 (2.0)	61.5 (1.5)	73.2 (1.3)	86.2 (2.1)	99.2 (3.4)	107.6 (4.3)
12–19								
Given day	92.3 (2.1)	31.7 (1.7)	40.5 (1.2)	56.6 (1.8)	80.8 (2.2)	115.0 (4.9)	158.2 (6.1)	188.6 (12.7)
Within-person mean	87.4 (2.2)	35.7 (1.6)	42.5 (1.8)	59.9 (1.8)	80.1 (1.7)	107.9 (4.0)	141.1 (7.0)	159.7 (10.1)
National Cancer Institute Method	87.5 (2.3)	48.1 (2.9)	55.0 (2.6)	67.8 (2.0)	84.5 (2.0)	104.0 (3.5)	124.1 (5.7)	137.4 (7.3)
20–39								
Given day	101.2 (1.3)	32.2 (1.5)	43.6 (1.3)	65.2 (1.1)	94.7 (1.6)	127.4 (2.2)	165.0 (3.2)	193.0 (4.7)
Within-person mean	97.3 (1.3)	39.3 (1.4)	50.0 (1.6)	67.3 (1.0)	89.9 (1.4)	116.9 (2.0)	155.4 (2.7)	181.3 (8.0)
National Cancer Institute Method	97.1 (1.2)	55.6 (2.0)	63.2 (1.7)	77.2 (1.3)	94.8 (1.1)	114.4 (2.0)	134.0 (3.3)	146.6 (4.3)
40–59								
Given day	97.4 (1.9)	30.9 (1.6)	41.0 (2.1)	60.5 (1.1)	88.3 (2.6)	123.3 (3.4)	168.3 (4.9)	199.3 (7.2)
Within-person mean	93.7 (1.5)	37.6 (1.2)	46.6 (2.1)	65.3 (1.3)	88.5 (2.0)	114.8 (2.7)	149.6 (3.3)	170.4 (4.4)
National Cancer Institute Method	94.2 (1.4)	53.0 (1.9)	60.5 (1.8)	74.2 (1.6)	91.7 (1.4)	111.4 (1.7)	131.2 (2.7)	144.1 (3.6)
60 and over								
Given day	84.2 (1.3)	28.4 (1.3)	36.6 (1.0)	54.7 (1.3)	79.6 (1.8)	105.9 (2.7)	138.1 (4.3)	158.6 (5.2)
Within-person mean	82.7 (1.2)	32.5 (1.5)	42.0 (1.7)	57.3 (1.3)	78.6 (1.2)	100.6 (1.6)	132.2 (3.2)	147.9 (4.1)
National Cancer Institute Method	83.0 (1.2)	45.0 (1.4)	51.8 (1.3)	64.6 (1.0)	80.7 (1.1)	98.9 (1.6)	117.1 (2.5)	128.9 (3.2)
Females								
2–5								
Given day	53.6 (1.2)	19.7 (1.1)	26.2 (1.2)	34.7 (1.4)	49.1 (1.3)	65.7 (1.8)	89.4 (4.0)	99.6 (3.3)
Within-person mean	52.7 (1.1)	25.5 (1.2)	28.9 (0.9)	37.9 (1.2)	49.5 (1.2)	64.9 (2.6)	79.7 (2.3)	92.8 (4.4)
National Cancer Institute Method	52.5 (1.3)	32.6 (1.4)	36.3 (1.2)	43.0 (0.9)	51.4 (1.2)	60.8 (2.0)	69.9 (3.0)	75.9 (3.8)
6–11								
Given day	68.8 (1.2)	27.5 (1.5)	34.6 (1.6)	47.7 (1.5)	66.2 (1.5)	83.9 (2.1)	106.8 (2.6)	123.4 (3.3)
Within-person mean	67.9 (1.1)	32.4 (2.0)	40.4 (1.2)	50.4 (1.4)	65.8 (1.2)	80.4 (1.6)	101.6 (4.1)	113.3 (2.1)
National Cancer Institute Method	67.7 (1.0)	49.3 (3.3)	52.9 (2.8)	59.3 (1.9)	67.1 (1.0)	75.4 (1.7)	83.5 (3.2)	88.6 (4.2)
12–19								
Given day	65.9 (1.8)	16.0 (2.8)	25.7 (2.4)	39.3 (1.9)	61.3 (1.6)	84.1 (2.1)	114.0 (2.9)	131.7 (4.6)
Within-person mean	63.9 (1.2)	23.7 (3.0)	30.2 (1.6)	44.3 (1.7)	61.4 (1.8)	81.0 (2.0)	97.3 (1.8)	115.3 (4.4)
National Cancer Institute Method	64.1 (1.1)	38.1 (3.9)	43.0 (3.4)	52.0 (2.4)	62.9 (1.3)	75.0 (1.6)	86.7 (3.1)	94.2 (4.3)
20–39								
Given day	74.1 (1.0)	23.2 (1.4)	31.7 (1.0)	46.5 (1.2)	66.5 (1.4)	95.4 (2.0)	126.3 (2.4)	146.4 (4.5)
Within-person mean	70.3 (0.9)	26.5 (1.7)	33.3 (1.5)	47.2 (1.1)	65.3 (1.1)	87.3 (1.6)	111.1 (2.5)	126.8 (4.0)
National Cancer Institute Method	70.3 (0.9)	41.6 (2.3)	46.8 (2.1)	56.4 (1.5)	68.6 (0.9)	82.3 (1.2)	96.0 (2.3)	104.9 (3.1)
40–59								
Given day	70.3 (0.8)	20.9 (1.3)	27.8 (1.5)	43.3 (1.3)	65.3 (0.8)	89.1 (1.4)	118.7 (2.9)	139.3 (3.7)
Within-person mean	69.5 (0.9)	25.5 (1.5)	33.8 (1.6)	48.2 (0.9)	65.4 (0.9)	87.0 (1.7)	106.3 (3.0)	126.9 (5.2)
National Cancer Institute Method	69.6 (1.0)	36.8 (1.6)	42.7 (1.4)	53.6 (1.1)	67.5 (1.0)	83.3 (1.4)	99.2 (2.2)	109.5 (2.9)
60 and over								
Given day	63.9 (1.3)	18.7 (1.0)	26.8 (1.2)	40.1 (1.1)	57.7 (1.6)	80.8 (2.0)	107.6 (3.0)	127.4 (4.1)
Within-person mean	63.4 (1.0)	24.7 (0.9)	30.5 (1.1)	42.8 (1.1)	60.2 (0.9)	79.3 (1.3)	100.0 (2.5)	114.3 (4.4)
National Cancer Institute Method	63.6 (0.9)	33.5 (1.5)	38.7 (1.3)	48.6 (1.0)	61.4 (0.8)	76.1 (1.4)	91.2 (2.3)	101.2 (3.1)

NOTES: Given day estimates are from Day 1 dietary recall, within-person mean estimates are from Days 1 and 2 dietary recalls, and usual intake distributions are estimated using the National Cancer Institute Method. Children who were reported to consume breast milk and women who were pregnant or lactating were excluded from the data analysis.

SOURCE: NCHS, National Health and Nutrition Examination Survey, 2011–2014.

Table 4. Energy (kilocalories per day): Mean and percentile, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014

Sex, age group (years), and method	Mean (standard error)	5th (standard error)	10th (standard error)	25th (standard error)	50th (standard error)	75th (standard error)	90th (standard error)	95th (standard error)
Males								
2–5								
Given day	1,613 (25)	859 (24)	986 (31)	1,251 (30)	1,542 (35)	1,877 (30)	2,229 (52)	2,543 (109)
Within-person mean	1,589 (21)	926 (40)	1,055 (25)	1,270 (28)	1,529 (25)	1,857 (22)	2,148 (47)	2,383 (46)
National Cancer Institute Method	1,590 (19)	1,055 (37)	1,158 (33)	1,342 (25)	1,567 (19)	1,813 (25)	2,050 (42)	2,202 (55)
6–11								
Given day	2,061 (31)	1,082 (51)	1,264 (32)	1,589 (37)	1,982 (37)	2,404 (38)	2,983 (79)	3,280 (81)
Within-person mean	2,033 (26)	1,207 (30)	1,368 (35)	1,643 (24)	1,986 (37)	2,354 (43)	2,758 (71)	3,095 (77)
National Cancer Institute Method	2,036 (25)	1,463 (49)	1,578 (42)	1,776 (30)	2,016 (25)	2,273 (40)	2,521 (63)	2,678 (79)
12–19								
Given day	2,453 (36)	1,038 (46)	1,256 (46)	1,702 (49)	2,211 (37)	2,965 (54)	3,960 (137)	4,641 (270)
Within-person mean	2,327 (44)	1,124 (52)	1,290 (58)	1,732 (44)	2,190 (29)	2,744 (76)	3,483 (170)	4,074 (169)
National Cancer Institute Method	2,326 (48)	1,450 (78)	1,612 (66)	1,906 (47)	2,276 (43)	2,692 (78)	3,108 (129)	3,376 (165)
20–39								
Given day	2,716 (30)	1,225 (36)	1,454 (44)	1,914 (34)	2,570 (33)	3,320 (61)	4,147 (44)	4,756 (129)
Within-person mean	2,608 (30)	1,290 (33)	1,518 (33)	1,917 (29)	2,491 (29)	3,058 (43)	3,799 (95)	4,471 (138)
National Cancer Institute Method	2,602 (29)	1,582 (41)	1,775 (36)	2,124 (27)	2,553 (27)	3,026 (46)	3,492 (76)	3,791 (97)
40–59								
Given day	2,560 (45)	1,099 (57)	1,392 (33)	1,798 (41)	2,400 (36)	3,116 (65)	3,921 (86)	4,434 (182)
Within-person mean	2,455 (38)	1,258 (34)	1,483 (29)	1,865 (37)	2,360 (35)	2,916 (62)	3,631 (98)	4,000 (84)
National Cancer Institute Method	2,463 (39)	1,522 (43)	1,696 (39)	2,014 (33)	2,412 (35)	2,856 (53)	3,295 (83)	3,581 (106)
60 and over								
Given day	2,167 (27)	980 (31)	1,186 (35)	1,629 (32)	2,077 (31)	2,631 (48)	3,196 (84)	3,637 (99)
Within-person mean	2,140 (25)	1,111 (39)	1,318 (40)	1,671 (21)	2,093 (26)	2,564 (41)	3,035 (54)	3,428 (84)
National Cancer Institute Method	2,147 (25)	1,344 (41)	1,497 (34)	1,775 (25)	2,113 (22)	2,482 (37)	2,841 (57)	3,069 (72)
Females								
2–5								
Given day	1,459 (24)	745 (27)	842 (33)	1,096 (23)	1,415 (28)	1,771 (37)	2,142 (82)	2,378 (72)
Within-person mean	1,453 (22)	819 (21)	921 (25)	1,150 (23)	1,401 (22)	1,708 (43)	1,992 (61)	2,271 (80)
National Cancer Institute Method	1,446 (26)	1,011 (30)	1,096 (26)	1,247 (22)	1,429 (25)	1,628 (37)	1,818 (53)	1,940 (65)
6–11								
Given day	1,833 (26)	958 (44)	1,133 (21)	1,406 (33)	1,759 (37)	2,192 (44)	2,649 (83)	2,890 (78)
Within-person mean	1,816 (25)	1,078 (33)	1,205 (33)	1,460 (36)	1,786 (20)	2,046 (35)	2,526 (75)	2,807 (79)
National Cancer Institute Method	1,815 (22)	1,366 (66)	1,457 (57)	1,613 (39)	1,801 (23)	2,000 (31)	2,192 (57)	2,313 (75)
12–19								
Given day	1,749 (31)	712 (40)	923 (42)	1,220 (28)	1,647 (41)	2,196 (47)	2,671 (55)	3,061 (83)
Within-person mean	1,704 (25)	802 (46)	974 (37)	1,257 (36)	1,671 (36)	2,110 (43)	2,439 (42)	2,725 (60)
National Cancer Institute Method	1,706 (24)	1,097 (44)	1,218 (38)	1,430 (30)	1,686 (24)	1,959 (28)	2,221 (41)	2,386 (51)
20–39								
Given day	1,966 (22)	830 (38)	1,041 (31)	1,398 (25)	1,830 (31)	2,404 (24)	3,054 (56)	3,510 (106)
Within-person mean	1,880 (20)	906 (35)	1,097 (41)	1,427 (22)	1,809 (27)	2,238 (29)	2,805 (46)	3,097 (45)
National Cancer Institute Method	1,883 (21)	1,207 (47)	1,333 (41)	1,564 (31)	1,849 (22)	2,164 (28)	2,475 (46)	2,675 (60)
40–59								
Given day	1,824 (18)	805 (22)	962 (25)	1,317 (31)	1,749 (27)	2,200 (34)	2,722 (28)	3,014 (84)
Within-person mean	1,796 (18)	893 (36)	1,050 (29)	1,391 (22)	1,727 (25)	2,137 (32)	2,526 (48)	2,872 (51)
National Cancer Institute Method	1,797 (18)	1,097 (28)	1,232 (25)	1,473 (20)	1,767 (18)	2,089 (25)	2,403 (39)	2,602 (50)
60 and over								
Given day	1,637 (24)	720 (40)	902 (25)	1,208 (19)	1,562 (26)	1,961 (39)	2,448 (62)	2,789 (56)
Within-person mean	1,628 (19)	821 (26)	975 (22)	1,255 (21)	1,597 (23)	1,936 (30)	2,299 (36)	2,619 (57)
National Cancer Institute Method	1,631 (17)	982 (23)	1,103 (21)	1,324 (18)	1,597 (17)	1,902 (24)	2,204 (36)	2,399 (46)

NOTES: Given day estimates are from Day 1 dietary recall, within-person mean estimates are from Days 1 and 2 dietary recalls, and usual intake distributions are estimated using the National Cancer Institute Method. The limitations of self-reported measures of energy and the tendency of these to underestimate true energy intake are well documented (see references 6,15–17 in the report). Usual intake methods do not address the systematic bias associated with underreporting in dietary recalls (see reference 9 in the report). Children who were reported to consume breast milk and women who were pregnant or lactating were excluded from the data analysis.

SOURCE: NCHS, National Health and Nutrition Examination Survey, 2011–2014.

Table 5. Percentage of energy intake from total fat (per day): Mean, percentile, and percentage of persons meeting dietary reference intakes, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014

Sex, age group (years), and method	Mean (standard error)	5th (standard error)	10th (standard error)	25th (standard error)	50th (standard error)	75th (standard error)	90th (standard error)	95th (standard error)	Percent within AMDR (standard error)
Males									
2–5									
Given day	31.7 (0.41)	21.0 (0.63)	22.8 (0.49)	26.2 (0.47)	31.7 (0.53)	36.7 (0.41)	40.6 (0.64)	43.2 (1.48)	48.8 (3.33)
Ratio of within-person mean	31.5 (0.37)	21.8 (0.44)	23.6 (0.51)	26.8 (0.47)	31.5 (0.47)	35.8 (0.53)	40.0 (0.65)	41.5 (0.65)	53.0 (2.63)
Within-person mean mean ratio	28.9 (0.38)	12.7 (0.96)	16.8 (1.13)	24.0 (0.63)	30.0 (0.44)	34.5 (0.50)	38.9 (0.76)	40.9 (0.60)	45.8 (3.60)
Ratio of National Cancer Institute Method	32.0 (0.47)	25.3 (0.38)	26.7 (0.28)	29.2 (0.27)	31.9 (0.45)	34.8 (0.75)	37.2 (1.01)	39.1 (1.12)	72.7 ---
National Cancer Institute Method mean ratio	31.5 (0.36)	24.6 (0.49)	26.1 (0.37)	28.7 (0.28)	31.5 (0.38)	34.3 (0.58)	36.8 (0.77)	38.3 (0.90)	67.6 (2.47)
6–11									
Given day	32.7 (0.26)	20.8 (0.73)	23.5 (0.59)	28.4 (0.51)	32.9 (0.37)	37.3 (0.58)	41.6 (0.67)	44.2 (0.66)	49.4 (2.20)
Ratio of within-person mean	32.5 (0.26)	23.5 (0.52)	25.3 (0.42)	28.7 (0.36)	32.5 (0.41)	36.3 (0.44)	40.0 (0.36)	42.2 (0.70)	57.6 (3.03)
Within-person mean mean ratio	30.1 (0.34)	14.7 (0.73)	18.1 (1.06)	26.0 (0.57)	31.2 (0.40)	35.4 (0.45)	38.6 (0.46)	40.5 (0.52)	52.4 (2.57)
Ratio of National Cancer Institute Method	33.0 (0.27)	27.4 (0.80)	28.8 (0.61)	30.8 (0.42)	33.0 (0.28)	35.3 (0.37)	37.2 (0.59)	38.4 (0.74)	70.2 ---
National Cancer Institute Method mean ratio	32.4 (0.27)	27.0 (0.81)	28.2 (0.66)	30.2 (0.42)	32.4 (0.27)	34.6 (0.38)	36.5 (0.58)	37.7 (0.73)	77.7 (4.37)
12–19									
Given day	33.2 (0.37)	20.6 (0.59)	23.3 (0.75)	27.8 (0.54)	33.7 (0.51)	38.2 (0.55)	42.9 (0.51)	45.0 (0.96)	44.2 (2.98)
Ratio of within-person mean	33.3 (0.30)	23.0 (0.74)	25.8 (0.39)	29.5 (0.39)	33.4 (0.40)	37.6 (0.30)	40.7 (0.26)	42.3 (0.44)	52.3 (2.52)
Within-person mean mean ratio	31.1 (0.34)	15.9 (1.07)	19.1 (0.86)	27.1 (0.52)	32.3 (0.50)	36.5 (0.52)	40.2 (0.41)	41.5 (0.31)	47.8 (2.45)
Ratio of National Cancer Institute Method	33.5 (0.42)	27.9 (0.77)	29.2 (0.59)	31.2 (0.45)	33.5 (0.42)	35.8 (0.57)	37.7 (0.77)	38.9 (0.99)	66.4 ---
National Cancer Institute Method mean ratio	33.2 (0.27)	27.8 (0.81)	29.0 (0.64)	31.0 (0.40)	33.2 (0.27)	35.4 (0.42)	37.3 (0.65)	38.5 (0.79)	70.9 (4.45)
20–39									
Given day	33.1 (0.30)	18.4 (0.76)	22.0 (0.62)	27.6 (0.32)	33.4 (0.35)	38.4 (0.32)	43.8 (0.45)	48.1 (0.58)	50.9 (1.55)
Ratio of within-person mean	33.3 (0.27)	21.3 (0.71)	24.4 (0.54)	28.7 (0.28)	33.2 (0.39)	37.9 (0.28)	42.6 (0.27)	45.4 (0.79)	55.9 (1.42)
Within-person mean mean ratio	30.1 (0.41)	13.8 (0.95)	16.7 (0.65)	24.4 (0.74)	31.0 (0.41)	36.4 (0.41)	41.3 (0.53)	44.0 (0.55)	51.6 (1.67)
Ratio of National Cancer Institute Method	33.7 (0.30)	26.8 (0.65)	28.3 (0.57)	30.8 (0.44)	33.7 (0.33)	36.5 (0.40)	39.1 (0.58)	40.8 (0.78)	62.9 ---
National Cancer Institute Method mean ratio	33.1 (0.24)	25.8 (0.69)	27.4 (0.54)	30.1 (0.31)	33.1 (0.25)	36.0 (0.46)	38.6 (0.72)	40.2 (0.87)	66.8 (3.30)
40–59									
Given day	33.7 (0.32)	19.1 (0.68)	21.8 (0.49)	28.0 (0.48)	33.9 (0.32)	39.7 (0.43)	44.3 (0.57)	48.4 (0.78)	49.8 (1.51)
Ratio of within-person mean	34.0 (0.27)	21.2 (0.62)	24.1 (0.44)	29.6 (0.35)	34.3 (0.32)	38.7 (0.23)	42.9 (0.53)	46.6 (0.78)	50.2 (1.60)
Within-person mean mean ratio	32.4 (0.33)	16.4 (0.58)	19.9 (0.60)	27.6 (0.52)	33.1 (0.37)	38.0 (0.29)	42.3 (0.54)	46.3 (0.99)	49.7 (1.67)
Ratio of National Cancer Institute Method	34.2 (0.46)	27.2 (0.62)	28.7 (0.52)	31.3 (0.42)	34.2 (0.45)	37.1 (0.65)	39.7 (0.99)	41.7 (1.24)	59.0 ---
National Cancer Institute Method mean ratio	33.8 (0.23)	26.4 (0.57)	28.1 (0.45)	30.8 (0.28)	33.9 (0.23)	36.9 (0.37)	39.6 (0.55)	41.2 (0.67)	59.9 (2.39)

See footnotes at end of table.

Table 5. Percentage of energy intake from total fat (per day): Mean, percentile, and percentage of persons meeting dietary reference intakes, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014—Con.

Sex, age group (years), and method	Mean (standard error)	5th (standard error)	10th (standard error)	25th (standard error)	50th (standard error)	75th (standard error)	90th (standard error)	95th (standard error)	Percent within AMDR (standard error)
Males—Con.									
60 and over									
Given day	34.3 (0.34)	18.6 (1.03)	23.7 (0.69)	28.7 (0.26)	34.1 (0.46)	39.4 (0.36)	45.9 (0.75)	49.7 (0.58)	48.4 (1.88)
Ratio of within-person mean	34.3 (0.29)	21.9 (0.58)	25.3 (0.59)	29.2 (0.31)	34.3 (0.43)	39.0 (0.37)	43.4 (0.59)	47.0 (0.71)	50.4 (2.35)
Within-person mean mean ratio	32.8 (0.32)	17.4 (0.81)	21.0 (0.69)	27.8 (0.37)	33.5 (0.49)	38.2 (0.35)	42.5 (0.60)	45.6 (0.84)	51.2 (2.36)
Ratio of National Cancer Institute Method	34.8 (0.32)	26.1 (0.57)	28.1 (0.48)	31.2 (0.41)	34.8 (0.33)	38.3 (0.36)	41.4 (0.44)	43.9 (0.55)	51.7 ---
National Cancer Institute Method mean ratio	34.0 (0.28)	25.2 (0.50)	27.2 (0.41)	30.5 (0.31)	34.1 (0.28)	37.7 (0.37)	40.8 (0.50)	42.7 (0.60)	56.2 (2.20)
Females									
2–5									
Given day	32.4 (0.33)	20.9 (1.05)	24.0 (0.50)	27.8 (0.39)	32.5 (0.35)	37.5 (0.44)	40.9 (0.45)	42.8 (0.58)	49.7 (2.41)
Ratio of within-person mean	32.0 (0.28)	22.3 (0.49)	24.5 (0.41)	28.5 (0.47)	32.3 (0.37)	36.0 (0.38)	39.1 (0.44)	41.4 (0.72)	57.3 (2.96)
Within-person mean mean ratio	29.5 (0.37)	14.4 (0.90)	17.9 (1.03)	25.3 (0.90)	30.9 (0.47)	34.7 (0.34)	38.3 (0.46)	40.0 (0.71)	51.0 (3.23)
Ratio of National Cancer Institute Method	32.8 (0.34)	27.1 (0.87)	28.4 (0.68)	30.5 (0.48)	32.9 (0.35)	35.1 (0.45)	37.1 (0.69)	38.6 (0.83)	77.7 ---
National Cancer Institute Method mean ratio	32.0 (0.27)	26.5 (0.92)	27.7 (0.72)	29.8 (0.41)	32.0 (0.27)	34.2 (0.49)	36.1 (0.76)	37.3 (0.94)	76.0 (4.74)
6–11									
Given day	33.1 (0.29)	20.7 (0.60)	23.6 (0.50)	28.6 (0.32)	33.5 (0.49)	37.6 (0.42)	41.8 (0.47)	44.2 (0.64)	44.2 (2.23)
Ratio of within-person mean	33.3 (0.23)	23.9 (0.48)	26.4 (0.29)	29.7 (0.30)	33.2 (0.34)	36.9 (0.34)	40.5 (0.60)	42.8 (0.47)	55.4 (2.54)
Within-person mean mean ratio	31.3 (0.36)	17.4 (0.75)	21.1 (1.28)	27.4 (0.56)	32.0 (0.40)	36.1 (0.24)	39.6 (0.47)	41.9 (0.58)	52.3 (2.43)
Ratio of National Cancer Institute Method	33.8 (0.29)	29.4 (0.71)	30.4 (0.53)	32.0 (0.33)	33.8 (0.29)	35.5 (0.44)	37.2 (0.61)	38.2 (0.86)	67.2 ---
National Cancer Institute Method mean ratio	33.0 (0.26)	28.3 (0.62)	29.4 (0.48)	31.1 (0.29)	33.0 (0.26)	34.9 (0.43)	36.6 (0.64)	37.6 (0.78)	75.8 (5.19)
12–19									
Given day	33.0 (0.50)	18.0 (2.60)	22.5 (1.09)	28.3 (0.58)	33.3 (0.46)	38.7 (0.49)	43.4 (0.61)	45.8 (0.78)	42.5 (2.60)
Ratio of within-person mean	33.2 (0.28)	21.2 (1.18)	24.8 (0.74)	29.4 (0.45)	33.7 (0.40)	37.5 (0.26)	41.4 (0.41)	43.4 (0.69)	48.7 (2.48)
Within-person mean mean ratio	31.3 (0.38)	14.8 (0.97)	18.8 (0.98)	27.3 (0.72)	32.5 (0.43)	36.9 (0.40)	40.7 (0.51)	43.1 (0.58)	45.4 (2.18)
Ratio of National Cancer Institute Method	33.8 (0.51)	28.5 (1.70)	29.7 (1.28)	31.7 (0.76)	33.9 (0.47)	35.9 (0.66)	37.7 (0.97)	38.8 (1.21)	63.9 ---
National Cancer Institute Method mean ratio	32.9 (0.34)	26.5 (1.54)	28.0 (1.19)	30.4 (0.68)	33.0 (0.33)	35.5 (0.62)	37.6 (1.05)	38.9 (1.31)	68.9 (6.78)
20–39									
Given day	33.4 (0.27)	17.6 (0.70)	21.5 (0.41)	28.0 (0.35)	33.4 (0.30)	39.3 (0.44)	43.8 (0.44)	47.7 (1.04)	50.6 (1.74)
Ratio of within-person mean	33.2 (0.24)	20.6 (0.55)	23.8 (0.39)	28.5 (0.22)	33.0 (0.24)	37.8 (0.27)	42.6 (0.48)	45.2 (0.82)	57.1 (1.09)
Within-person mean mean ratio	30.8 (0.33)	15.4 (0.66)	18.3 (0.66)	25.7 (0.41)	31.3 (0.33)	36.6 (0.28)	41.4 (0.44)	44.1 (0.69)	54.8 (1.22)
Ratio of National Cancer Institute Method	34.0 (0.24)	26.7 (0.67)	28.3 (0.55)	31.0 (0.38)	34.0 (0.25)	36.9 (0.33)	39.7 (0.51)	41.8 (0.74)	60.2 ---
National Cancer Institute Method mean ratio	32.9 (0.24)	25.7 (0.60)	27.3 (0.49)	29.9 (0.33)	32.8 (0.24)	35.8 (0.32)	38.4 (0.49)	40.0 (0.60)	68.8 (2.34)

See footnotes at end of table.

Table 5. Percentage of energy intake from total fat (per day): Mean, percentile, and percentage of persons meeting dietary reference intakes, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2011–2014—Con.

Sex, age group (years), and method	Mean (standard error)	5th (standard error)	10th (standard error)	25th (standard error)	50th (standard error)	75th (standard error)	90th (standard error)	95th (standard error)	Percent within AMDR (standard error)
Females—Con.									
40–59									
Given day	33.6 (0.28)	17.1 (0.96)	21.9 (0.63)	27.7 (0.34)	33.8 (0.51)	40.1 (0.44)	45.0 (0.50)	48.4 (0.64)	46.9 (2.16)
Ratio of within-person mean	34.1 (0.24)	21.8 (0.55)	24.7 (0.46)	29.3 (0.43)	34.5 (0.32)	38.9 (0.36)	42.9 (0.38)	45.4 (0.32)	50.6 (1.81)
Within-person mean mean ratio	32.8 (0.30)	17.5 (0.61)	21.7 (0.72)	27.5 (0.36)	33.5 (0.33)	38.4 (0.44)	42.5 (0.35)	44.7 (0.29)	51.0 (1.85)
Ratio of National Cancer Institute Method	34.4 (0.24)	26.8 (0.55)	28.5 (0.44)	31.2 (0.35)	34.3 (0.24)	37.5 (0.30)	40.0 (0.43)	42.0 (0.56)	54.9 ---
National Cancer Institute Method mean ratio	33.8 (0.25)	26.0 (0.58)	27.7 (0.48)	30.6 (0.33)	33.8 (0.25)	37.0 (0.31)	39.8 (0.45)	41.5 (0.54)	59.7 (2.14)
60 and over									
Given day	34.1 (0.37)	18.8 (0.59)	22.3 (0.40)	28.3 (0.45)	34.2 (0.45)	40.4 (0.46)	45.4 (0.57)	49.3 (0.99)	47.4 (2.44)
Ratio of within-person mean	34.3 (0.27)	21.8 (0.52)	24.6 (0.38)	29.6 (0.31)	34.4 (0.37)	39.4 (0.38)	43.8 (0.40)	46.9 (0.55)	50.2 (1.97)
Within-person mean mean ratio	33.2 (0.34)	18.3 (0.81)	22.3 (0.62)	28.3 (0.44)	33.7 (0.37)	38.4 (0.41)	43.2 (0.48)	46.2 (0.56)	51.2 (1.77)
Ratio of National Cancer Institute Method	34.9 (0.30)	26.9 (0.85)	28.7 (0.66)	31.6 (0.44)	34.9 (0.30)	38.2 (0.47)	41.0 (0.72)	43.4 (0.93)	51.0 ---
National Cancer Institute Method mean ratio	34.1 (0.24)	26.2 (0.72)	27.9 (0.56)	30.8 (0.32)	34.0 (0.24)	37.3 (0.46)	40.3 (0.72)	42.1 (0.89)	57.9 (2.44)

--- Data not available for bivariate macros (ratio of usual intake). The ratio of usual intake is run with bivariate usual intake macros, not the mixtran and distrib for usual ratio. The bivariate macros can only calculate the percentage below the cutoff (i.e., the percentage below the lower cut point and the percentage below the upper cut point). To find the percentage within acceptable macronutrient distribution range (AMDR), simple subtraction is used, which is a different calculation method from usual intake ratio, where the percentage within AMDR is actually estimated in the macro.

NOTES: AMDR is acceptable macronutrient distribution range. Given day estimates are from Day 1 dietary recall, within-person mean estimates are from Days 1 and 2 dietary recalls, and usual intake distributions are estimated using the National Cancer Institute Method. The recommended way to evaluate the adequacy of dietary intake in relation to reference standards is by using methods that estimate usual intake and therefore account for within-person variation. Comparing estimates based on given day and within-person mean with reference standards is not recommended. They are presented to illustrate the inaccuracy associated with using methods that do not account for within-person variation (see references 12–15 in the report). Children who were reported to consume breast milk and women who were pregnant or lactating were excluded from the data analysis.

SOURCE: NCHS, National Health and Nutrition Examination Survey, 2011–2014.

Table 6. Total fruit (cup equivalents per day): Mean, percentile, and the percentage of persons below and above the 2015–2020 Dietary Guidelines recommended consumption range, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2009–2012

Sex, age group (years), and method	Mean (standard error)	5th (standard error)	10th (standard error)	25th (standard error)	50th (standard error)
Males					
2–5					
Given day	1.5 (0.06)	---	0.0 (0.02)	0.5 (0.07)	1.2 (0.09)
Within-person mean	1.5 (0.05)	0.1 (0.02)	0.2 (0.03)	0.7 (0.06)	1.3 (0.06)
National Cancer Institute Method	1.5 (0.06)	0.5 (0.06)	0.7 (0.06)	1.0 (0.06)	1.4 (0.06)
6–11					
Given day	1.2 (0.06)	---	---	0.1 (0.04)	0.8 (0.08)
Within-person mean	1.2 (0.05)	---	0.0 (0.02)	0.4 (0.05)	1.0 (0.04)
National Cancer Institute Method	1.2 (0.05)	0.4 (0.07)	0.6 (0.07)	0.8 (0.07)	1.1 (0.06)
12–19					
Given day	1.1 (0.11)	---	---	---	0.5 (0.14)
Within-person mean	1.1 (0.08)	---	---	0.1 (0.03)	0.6 (0.07)
National Cancer Institute Method	1.0 (0.07)	0.2 (0.03)	0.2 (0.04)	0.5 (0.05)	0.8 (0.06)
20–39					
Given day	1.0 (0.05)	---	---	---	0.3 (0.05)
Within-person mean	1.0 (0.05)	---	---	---	0.6 (0.05)
National Cancer Institute Method	1.0 (0.04)	0.1 (0.02)	0.2 (0.03)	0.4 (0.04)	0.8 (0.04)
40–59					
Given day	1.0 (0.04)	---	---	---	0.5 (0.08)
Within-person mean	1.1 (0.04)	---	---	0.0 (0.00)	0.7 (0.10)
National Cancer Institute Method	1.1 (0.04)	0.1 (0.03)	0.2 (0.03)	0.5 (0.04)	0.9 (0.05)
60 and over					
Given day	1.2 (0.06)	---	---	---	0.8 (0.05)
Within-person mean	1.2 (0.05)	---	---	0.3 (0.04)	0.9 (0.05)
National Cancer Institute Method	1.2 (0.05)	0.2 (0.03)	0.3 (0.03)	0.6 (0.03)	1.0 (0.05)
Females					
2–5					
Given day	1.4 (0.07)	---	0.1 (0.03)	0.5 (0.05)	1.1 (0.07)
Within-person mean	1.5 (0.07)	0.1 (0.05)	0.3 (0.06)	0.7 (0.06)	1.3 (0.07)
National Cancer Institute Method	1.5 (0.08)	0.5 (0.07)	0.7 (0.06)	1.0 (0.06)	1.4 (0.07)
6–11					
Given day	1.2 (0.06)	---	---	0.2 (0.08)	0.9 (0.10)
Within-person mean	1.2 (0.07)	---	0.1 (0.03)	0.5 (0.04)	1.0 (0.07)
National Cancer Institute Method	1.2 (0.07)	0.5 (0.08)	0.6 (0.07)	0.9 (0.07)	1.1 (0.07)
12–19					
Given day	0.8 (0.05)	---	---	---	0.3 (0.07)
Within-person mean	0.8 (0.04)	---	---	0.1 (0.02)	0.6 (0.06)
National Cancer Institute Method	0.8 (0.04)	0.2 (0.04)	0.3 (0.04)	0.4 (0.04)	0.7 (0.04)
20–39					
Given day	0.9 (0.05)	---	---	---	0.4 (0.07)
Within-person mean	0.9 (0.04)	---	---	0.1 (0.02)	0.6 (0.04)
National Cancer Institute Method	0.9 (0.04)	0.2 (0.03)	0.3 (0.03)	0.5 (0.04)	0.8 (0.04)
40–59					
Given day	1.0 (0.04)	---	---	---	0.6 (0.05)
Within-person mean	1.0 (0.04)	---	---	0.2 (0.04)	0.7 (0.05)
National Cancer Institute Method	1.0 (0.04)	0.2 (0.02)	0.3 (0.03)	0.5 (0.03)	0.9 (0.03)
60 and over					
Given day	1.1 (0.04)	---	---	0.1 (0.03)	0.8 (0.03)
Within-person mean	1.2 (0.03)	---	0.0 (0.01)	0.4 (0.03)	1.0 (0.03)
National Cancer Institute Method	1.2 (0.04)	0.3 (0.03)	0.4 (0.03)	0.7 (0.03)	1.0 (0.04)

Table 6. Total fruit (cup equivalents per day): Mean, percentile, and the percentage of persons below and above the 2015–2020 Dietary Guidelines recommended consumption range, calculated by given day, within-person mean, and the National Cancer Institute Method in the U.S. population aged 2 years and over: National Health and Nutrition Examination Survey, 2009–2012—Con.

Sex, age group (years), and method	75th (standard error)	90th (standard error)	95th (standard error)	Percent below range (standard error)	Percent above range (standard error)
Males					
2–5					
Given day	2.1 (0.14)	3.3 (0.16)	4.2 (0.22)	35.2 (3.15)	33.9 (2.76)
Within-person mean	2.2 (0.13)	3.0 (0.13)	3.6 (0.20)	34.7 (2.40)	34.3 (3.20)
National Cancer Institute Method	1.9 (0.09)	2.5 (0.14)	2.8 (0.18)	27.1 (3.43)	33.2 (3.43)
6–11					
Given day	1.8 (0.08)	2.8 (0.10)	3.5 (0.18)	42.5 (2.31)	20.5 (1.79)
Within-person mean	1.8 (0.09)	2.5 (0.12)	3.2 (0.15)	51.8 (2.20)	18.9 (1.90)
National Cancer Institute Method	1.5 (0.06)	1.9 (0.10)	2.2 (0.14)	57.0 (3.71)	8.9 (2.10)
12–19					
Given day	1.7 (0.15)	3.0 (0.26)	3.9 (0.33)	45.1 (2.61)	15.2 (2.17)
Within-person mean	1.6 (0.11)	2.6 (0.28)	3.7 (0.20)	63.2 (1.90)	12.6 (1.90)
National Cancer Institute Method	1.4 (0.09)	2.1 (0.15)	2.6 (0.20)	86.1 (2.26)	7.3 (1.62)
20–39					
Given day	1.5 (0.09)	3.1 (0.18)	4.1 (0.18)	45.0 (1.72)	13.8 (1.12)
Within-person mean	1.5 (0.08)	2.6 (0.15)	3.4 (0.24)	60.8 (1.90)	10.7 (1.20)
National Cancer Institute Method	1.3 (0.06)	2.0 (0.10)	2.5 (0.14)	90.0 (1.38)	5.0 (0.99)
40–59					
Given day	1.6 (0.09)	2.7 (0.15)	3.8 (0.21)	48.7 (2.36)	11.3 (1.16)
Within-person mean	1.7 (0.08)	2.7 (0.09)	3.4 (0.18)	62.4 (2.00)	12.3 (1.10)
National Cancer Institute Method	1.4 (0.05)	2.1 (0.08)	2.7 (0.13)	88.0 (1.18)	6.2 (0.94)
60 and over					
Given day	1.8 (0.09)	3.0 (0.12)	3.6 (0.18)	53.9 (1.94)	14.1 (1.43)
Within-person mean	1.7 (0.09)	2.7 (0.10)	3.3 (0.13)	66.0 (2.20)	13.4 (1.40)
National Cancer Institute Method	1.6 (0.08)	2.2 (0.12)	2.7 (0.15)	85.9 (2.08)	6.8 (1.41)
Females					
2–5					
Given day	2.0 (0.12)	2.9 (0.12)	3.4 (0.26)	38.9 (2.09)	39.4 (1.92)
Within-person mean	2.0 (0.11)	3.1 (0.21)	3.7 (0.28)	34.2 (2.70)	41.3 (2.50)
National Cancer Institute Method	1.9 (0.10)	2.4 (0.15)	2.8 (0.19)	27.2 (3.65)	43.3 (4.08)
6–11					
Given day	1.8 (0.13)	2.8 (0.08)	3.2 (0.13)	43.8 (2.09)	27.5 (2.35)
Within-person mean	1.8 (0.10)	2.5 (0.13)	3.1 (0.24)	52.5 (2.90)	25.8 (3.00)
National Cancer Institute Method	1.5 (0.09)	1.9 (0.12)	2.1 (0.16)	57.0 (5.29)	15.3 (3.88)
12–19					
Given day	1.4 (0.08)	2.2 (0.16)	2.7 (0.18)	44.8 (3.43)	12.1 (1.74)
Within-person mean	1.2 (0.07)	1.8 (0.09)	2.3 (0.13)	64.3 (2.70)	8.5 (1.10)
National Cancer Institute Method	1.0 (0.05)	1.4 (0.10)	1.7 (0.14)	91.3 (2.30)	2.4 (1.22)
20–39					
Given day	1.5 (0.07)	2.5 (0.14)	3.4 (0.21)	44.3 (1.20)	16.4 (1.26)
Within-person mean	1.4 (0.07)	2.4 (0.11)	3.2 (0.15)	58.8 (1.60)	14.5 (1.30)
National Cancer Institute Method	1.2 (0.05)	1.8 (0.08)	2.1 (0.11)	83.9 (1.67)	6.4 (1.20)
40–59					
Given day	1.5 (0.09)	2.5 (0.15)	3.3 (0.23)	49.6 (1.85)	16.3 (1.23)
Within-person mean	1.5 (0.06)	2.4 (0.06)	3.1 (0.13)	61.5 (1.60)	15.5 (1.00)
National Cancer Institute Method	1.4 (0.05)	1.9 (0.10)	2.4 (0.14)	79.7 (1.90)	9.2 (1.51)
60 and over					
Given day	1.7 (0.05)	2.6 (0.09)	3.1 (0.12)	56.0 (1.48)	18.3 (1.16)
Within-person mean	1.7 (0.06)	2.4 (0.09)	3.1 (0.09)	63.5 (2.00)	16.5 (1.20)
National Cancer Institute Method	1.5 (0.05)	2.1 (0.08)	2.4 (0.10)	73.9 (2.14)	11.0 (1.48)

0.0 Quantity more than zero but less than 0.05.

--- SUDAAN is unable to estimate any quantile less than or equal to the percentage of data accounted for by the 0 values.

NOTES: Given day estimates are from Day 1 dietary recall, within-person mean estimates are from Days 1 and 2 dietary recalls, and usual intake distributions are estimated using the National Cancer Institute Method. The recommended way to evaluate the adequacy of dietary intake in relation to reference standards is by using methods that estimate usual intake and therefore account for within-person variation. Comparing estimates based on given day and within-person mean with reference standards is not recommended. They are presented to illustrate the inaccuracy associated with using methods that do not account for within-person variation (see references 12–15 in the report). Children who were reported to consume breast milk and women who were pregnant or lactating were excluded from the data analysis.

SOURCE: NCHS, National Health and Nutrition Examination Survey, 2009–2012.

Appendix I. Nutrient Reference Values Used in This Report

Table I. Dietary reference intake for calcium (estimated average requirement and tolerable upper limit), percentage of energy from fat (acceptable macronutrient distribution range), and recommended intake range of fruit, by the Institute of Medicine Life Stage Group

Sex and life stage age group (years) ¹	Estimated average requirement ¹ for calcium (milligrams per day)	Tolerable upper limit ¹ for calcium (milligrams per day)	Acceptable macronutrient distribution range ¹ for percent of energy from fat	Recommended intake of fruit ² (cup equivalents per day)
Males				
1–3	700	2,500	30–40	1.0–1.5
4–8	1,000	2,500	25–35	1.0–2.0
9–13	1,300	3,000	25–35	1.5–2.0
14–18	1,300	3,000	25–35	2.0–2.5
19–30	1,000	2,500	20–35	2.0–2.5
31–50	1,000	2,500	20–35	2.0–2.5
51–70	1,000	2,500	20–35	2.0–2.5
70 and over	1,200	2,000	20–35	2.0–2.5
Females				
1–3	700	2,500	30–40	1.0–1.5
4–8	1,000	2,500	25–35	1.0–1.5
9–13	1,300	3,000	25–35	1.5–2.0
14–18	1,300	3,000	25–35	1.5–2.0
19–30	1,000	2,500	20–35	1.5–2.0
31–50	1,000	2,500	20–35	1.5–2.0
51–70	1,200	2,000	20–35	1.5–2.0
70 and over	1,200	2,000	20–35	1.5–2.0

¹Institute of Medicine. Dietary reference intakes: The essential guide to nutrient requirements. Washington, DC: The National Academies Press. 2006.

²U.S. Department of Health and Human Services and U.S. Department of Agriculture. Dietary guidelines for Americans 2015–2020. 8th ed. 2015.

Appendix II. Steps for Creating Poststratified Balanced Repeated Replication Weights

Key Notes

- Must use full sample
- Use PROC SURVEYFREQ to calculate balanced repeated replication (BRR) weights
- Poststratify BRR weights to match original weighting

Steps

1. Import data set into SAS where all respondents have a weight greater than zero (Day 1 data only).
2. Define the poststratification cells:
 - a. Typically sex x race x age
 - b. Use age groups defined by Dietary Reference Intakes (DRI)
 - i. Example: age groups from **Table I**
3. Calculate the weighted count for all possible combinations of sex x race x age for the poststratification cells (PScell). THESE ARE YOUR CONTROL TOTALS.
4. Merge in Day 2 data.
5. Create a data set with a maximum of two lines of data per person (Days 1 and 2 diet recalls).
6. Create BRR weights in SAS.
7. Run macro for poststratification:
 - a. Sort data with BRR weights by PScell variable
 - b. Output weighted frequency counts of the sample data set, once for each BRR weight
 - c. Merge weighted frequency counts of the sample data set with original control data set, by PScell
 - d. Calculate the ratio of the weighted control frequency counts to weighted frequency counts of the sample data set for each BRR weight (ctrl/COUNT) (by RepWT_i, where i = 1 to the number of BRR weights created, dependent on the number of NHANES survey cycles merged)
 - i. This is referred to as the expansion factor
8. Check to ensure that all poststratified BRR weights are equivalent to original weighted frequency counts (control).
- e. Merge new expansion factor variable into the original data set with BRR weights (by PScell)
- f. Calculate NEW rounded poststratified weight (PSWEIGHT_i) by multiplying the BRR weight by the expansion factor (or ratio)
- g. Repeat steps b–f for each RepWt
- h. Run steps b–f for original weight variable (wtdrd1)

Appendix III. SAS Syntax Used to Create Poststratified Adjusted Balanced Repeated Replication Weights for Use in Usual Intake Estimation

*****;

Syntax used to create poststratified adjusted balanced repeated replication (BRR) weights for use in Usual Intake Estimation by

the NCI Method

ALL US persons, 0–150 YRS 2009–2010

Purpose of syntax: BRR WEIGHT DERIVATION FOR A SINGLE SURVEY OF NHANES, IN THIS EXAMPLE 2009–2010

SUBPOP: ALL US PERSONS (BRR WEIGHTS WILL ONLY BE CREATED IF SP HAS A WEIGHT VALUE FOR WTDRD1)

NOTES:

To adapt this syntax to additional survey cycles of NHANES analysis, merge in the relevant years and create the appropriate number of BRR weights according to the number of cycles of NHANES.

*****;

CREATING THE DATA SETS

DATA:

NHANES 2009–2010 (0910);

*paste the path name to the folder where the data files are stored;

libname f “datafile\path\name\here”;

*paste the path name to the folder where the BRR output is to be saved;

libname DataOut “outputBRRweights\path\name\here “;

*-----;

* Create a data set from the permanent data sets ;

*-----;

*DEMOGRAPHICS;

data WORK.demo ;

set
 F.demo_f (**KEEP**=SEQN SDDSRVYR RIDSTATR RIAGENDR RIDAGEYR RIDRETH1
 SDMVPSU SDMVSTRA);

RUN;

PROC SORT DATA=WORK.DEMO;

BY SEQN;

RUN;

```

*DIET DATA - Day one;
data WORK.DR1TOT (KEEP=SEQN WTDRD1 DR1DRSTZ DR1TCALC DR1DAY) ;
   set F.DR1TOT_F ;
RUN;

PROC SORT DATA=WORK.DR1TOT;
BY SEQN;
RUN;

DATA WORK.DR1TOTPLUS ;
  MERGE      WORK.DEMO
               WORK.DR1TOT
  BY SEQN;

```

*create all possible combinations of poststratification cells by gender x race x age;
 *the age groups are based on Dietary Reference Intakes (DRI) age groups;
 *must collapse some race groups to get stable estimates;
 * USE ALL ages;

```

if riagendr=1 and ridreth1 in (1,2,5) and .<ridgeyr<1 then PScell=1;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=1 and ridgeyr<4 then PScell=2;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=4 and ridgeyr<9 then PScell=3;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=9 and ridgeyr<14 then PScell=4;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=14 and ridgeyr<19 then PScell=5;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=19 and ridgeyr<31 then PScell=6;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=31 and ridgeyr<51 then PScell=8;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=51 and ridgeyr<70 then PScell=8;
if riagendr=1 and ridreth1 in (1,2,5) and ridgeyr>=70 then PScell=9;

if riagendr=2 and ridreth1 in (1,2,5) and .<ridgeyr<1 then PScell=10;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=1 and ridgeyr<4 then PScell=11;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=4 and ridgeyr<9 then PScell=12;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=9 and ridgeyr<14 then PScell=13;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=14 and ridgeyr<19 then PScell=14;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=19 and ridgeyr<31 then PScell=15;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=31 and ridgeyr<51 then PScell=16;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=51 and ridgeyr<70 then PScell=17;
if riagendr=2 and ridreth1 in (1,2,5) and ridgeyr>=70 then PScell=18;

```

```

if riagendr=1 and ridreth1 in (3) and .<ridgeyr<1 then PScell=19;
if riagendr=1 and ridreth1 in (3) and ridgeyr>=1 and ridgeyr<4 then PScell=20;
if riagendr=1 and ridreth1 in (3) and ridgeyr>=4 and ridgeyr<9 then PScell=21;
if riagendr=1 and ridreth1 in (3) and ridgeyr>=9 and ridgeyr<14 then PScell=22;

if riagendr=1 and ridreth1 in (3) and ridgeyr>=14 and ridgeyr<19 then PScell=23;
if riagendr=1 and ridreth1 in (3) and ridgeyr>=19 and ridgeyr<31 then PScell=24;
if riagendr=1 and ridreth1 in (3) and ridgeyr>=31 and ridgeyr<51 then PScell=25;
if riagendr=1 and ridreth1 in (3) and ridgeyr>=51 and ridgeyr<70 then PScell=26;
if riagendr=1 and ridreth1 in (3) and ridgeyr>=70 then PScell=27;

if riagendr=2 and ridreth1 in (3) and .<ridgeyr<1 then PScell=28;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=1 and ridgeyr<4 then PScell=29;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=4 and ridgeyr<9 then PScell=30;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=9 and ridgeyr<14 then PScell=31;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=14 and ridgeyr<19 then PScell=32;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=19 and ridgeyr<31 then PScell=33;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=31 and ridgeyr<51 then PScell=34;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=51 and ridgeyr<70 then PScell=35;
if riagendr=2 and ridreth1 in (3) and ridgeyr>=70 then PScell=36;

if riagendr=1 and ridreth1 in (4) and .<ridgeyr<1 then PScell=37;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=1 and ridgeyr<4 then PScell=38;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=4 and ridgeyr<9 then PScell=39;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=9 and ridgeyr<14 then PScell=40;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=14 and ridgeyr<19 then PScell=41;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=19 and ridgeyr<31 then PScell=42;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=31 and ridgeyr<51 then PScell=43;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=51 and ridgeyr<70 then PScell=44;
if riagendr=1 and ridreth1 in (4) and ridgeyr>=70 then PScell=45;

if riagendr=2 and ridreth1 in (4) and .<ridgeyr<1 then PScell=46;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=1 and ridgeyr<4 then PScell=47;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=4 and ridgeyr<9 then PScell=48;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=9 and ridgeyr<14 then PScell=49;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=14 and ridgeyr<19 then PScell=50;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=19 and ridgeyr<31 then PScell=51;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=31 and ridgeyr<51 then PScell=52;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=51 and ridgeyr<70 then PScell=53;
if riagendr=2 and ridreth1 in (4) and ridgeyr>=70 then PScell=54;
RUN;

```

```

PROC FREQ data=dr1totplus;
table pscell;
title "";
run; *no missing;

```

```

proc freq data=work.dr1totplus;
  where ridstatr=2;
  table dr1drstz;
run;

proc means data=work.dr1totplus;
  where ridstatr=2 and dr1drstz in (1, 4);
  var seqn;
  title "confirming filters for USDA weight replication";
run; *match N=9754;

*-----;
* create an output data set, called ctrl, that saves the weighted counts ;
* for each of the poststratification cells;
* the weighted counts are saved in a variable named 'ctrl';
*-----;

proc freq data= DR1TOTPLUS ; *NOW it is okay to restrict the data to the subpopulation of interest;
tables pscell/ out=ctrl(drop=percent rename=(count=ctrl));
weight wtdrd1;
run;

proc print data=ctrl;
run;

*-----;
* CHECK FOR stratum with more than 2 PSUs      ;
*-----;

proc freq data=work.DR1TOTPLUS;
table SDMVPSU SDMVSTRA;
title "categories of PSU AND STATUM";
run;

PROC FREQ DATA=WORK.DR1TOTPLUS;
WHERE SDMVPSU=3;
TABLE SDMVSTRA;
title ">2 psu's per stratum check";
RUN; *STRATUM 86;

proc sort data=work.DR1TOTPLUS;
by seqn;
run;

```

```
data work.dietUI; set work.DR1TOTPLUS;
if SDMVSTRA=86 and SDMVPSU=3 then SDMVPSU=1; *reassign PSU of stratum 86 to PSU=1;
by seqn;
if first.seqn and pscell ne . ; *keep first obs per person;
run;

proc print data=work.dr1totplus (obs=20);
where dr1drstz=1;
var dr1drstz ridstatr WTDRD1;
title "check for weights";
run;

proc print data=work.dr1totplus (obs=20);
where dr1drstz=2;
var dr1drstz ridstatr WTDRD1;
title "check for weights";
run;

proc print data=work.dr1totplus (obs=20);
where dr1drstz=4;
var dr1drstz ridstatr WTDRD1;
title "check for weights";
run;

proc print data=work.dr1totplus (obs=20);
where dr1drstz=5;
var dr1drstz ridstatr WTDRD1;
title "check for weights";
run;

proc print data=work.dr1totplus (obs=20);
where ridstatr=1;
var dr1drstz ridstatr WTDRD1;
title "check for weights";
run;
```

*****;

* Create BRR weights using SAS;

* NOTE: paper by USDA uses 0.3 for FAY method;

* 0.3 correlates with a perturbation factor of 70%: $100(1-k)$, where $k=0.3$;

* USDA paper = WWEIA, NHANES 2005–2006. Usual Nutrient Intakes from Food and Water;

* Compared to 1997 Dietary Reference Intakes for Vitamin D, Calcium, Phosphorus and Magnesium;

* to create BRR weights appropriate for additional survey cycles – ;

* CHANGE the value of REPS= ;

*****;

* This syntax is appropriate for a single survey cycle of NHANES;

* Each NHANES survey cycle typically has 15 Strata;

* The number of replicates (R) that are needed for BRR is the smallest integer that is ;

* divisible by 4 and is greater than or equal to S. For 2 years of NHANES data (a single ;

* survey cycle), this number is 16;

*****;

```
proc surveyfreq data=WORK.DIETUI
  varmethod=BRR(reps=16 outweights=brr_wts FAY=0.3); *paper by USDA uses 0.3 for FAY;
  strata SDMVSTRA;
  cluster SDMVPSU;
  weight wtdrd1;
  tables pscell;
```

run;

```
*****
*      Poststratification Macro (pstrat);
*Definitions: ;
*numwts = number of BRR weights created (in this example, 16)
* basename = name of the base weight used to create BRR weights, (WTDRD1 in this example)
* repname = replicate weight variables (RepWt_ is the name that SAS creates)
* controldata= data set that has the weighted counts of the full population (created above)
* psvar = variable that defines the categories of the poststratification combinations (created above)
* controlvar= variable that has the weighted counts per category of the poststratification combinations
(created above) ;
* data = name of the data set that has the BRR weights created by PROC SURVEYFREQ (user defined)
;
*****;

%macro pstrat(numwts=, basename=, repname=RepWt_, controldata=ctrl, psvar=pscell, controlvar=ctrl,
data=brr_wts);
proc sort data=&data;
  by &psvar;
  run;
%do i=1 %to &numwts;

  proc freq data=&data noprint;
  tables &psvar/out=smpl(drop=percent);
  weight &repname.&i;
  run;
  data smpl;
  merge smpl &controldata;
  by &psvar;
  ratio=&controlvar / count;
  run;
/* Updated with cumulate and round code */
  data &data;
  retain _c _r -1;
  merge &data smpl(drop=&controlvar count);
  by &psvar;
  psweight_&i=&repname.&i * ratio;
  if first.&psvar then do;
    _c=0;_r=0;
  end;
  _c = _c + psweight_&i;
  psweight_&i=round(_c) - _r;
  _r = _r + psweight_&i;
  drop ratio _c _r;
  run;

%end;
```

```

proc freq data=&data noprint;
tables &psvar/out=smpl(drop=percent);
weight &basename;
run;
data smpl;
merge smpl &controldata;
by &psvar;
ratio=&controlvar / count;
run;
data &data;
retain _c _r -1;
merge &data smpl(drop=&controlvar count);
by &psvar;
psweight_0=&basename * ratio;
if first.&psvar then do;
_c=0;_r=0;
end;
_c = _c + psweight_0;
psweight_0=round(_c) - _r;
_r = _r + psweight_0;
drop ratio _c _r;
run;

%mend;
%pstrat(numwts=16, basename=wtdrd1);

proc freq data=brr_wts;
tables pscell;
weight psweight_0;
run;

proc freq data=brr_wts;
tables pscell;
weight psweight_16;
run;

data WORK.brr_0910NEW;
set brr_wts(keep=SEQN RIDSTATR DR1DRSTZ psweight_0 - psweight_16);
run;

PROC MEANS DATA=WORK.brr_0910NEW n NMISS MEAN MIN MAX SUM; RUN;

PROC SORT DATA=WORK.brr_0910NEW ;
BY SEQN;
RUN;

```

```
data BRR.brr_0910NEW;
set WORK.brr_0910NEW;(keep=SEQN RIDSTATR DR1DRSTZ psweight_0 - psweight_16);
run;

*****
* Syntax above has cumulate and round code;
*****
* END ;
```

Appendix IV. National Cancer Institute Macros Used to Estimate Usual Intake

Table II. National Cancer Institute macros used to estimate usual intake

Nutrient and food estimated	Method	Macros used
Calcium	Ubiquitously consumed— AMOUNT	MIXTRAN, DISTRIB, BRR_PVALUE_CI
Total energy	Ubiquitously consumed— AMOUNT	MIXTRAN, DISTRIB, BRR_PVALUE_CI
Total fat	Ubiquitously consumed— AMOUNT	MIXTRAN, DISTRIB, BRR_PVALUE_CI
Fruit	Episodically consumed— NOCORR	MIXTRAN, DISTRIB, BRR_PVALUE_CI
Percent of energy from fat	Usual intake of ratios Ratio of usual intakes	MIXTRAN, DISTRIB, BRR_PVALUE_CI NLMIXED UNIVARIATE, NLMIXED BIVARIATE, DISTRIB BIVARIATE, PERCENTILES_SURVEY

Appendix V. Usual Intake Estimation—Univariate—Calcium Example

```

/************************/
*****LIBRARY*****
/****************************/
libname b24 "\path\tolibrary";
libname outlib "\path\tolibrary_where_files_are_to_be_saved";

/****************************/
*****CALL MACROS*****
/****************************/

/*Macros from NCI*/
%include "\pathname\mixtran_macro_v2.1.sas";
%include "\pathname\distrib_macro_v2.1.sas";
%include "\pathname\brr_pvalue_ci_macro_v1.0.sas";
%include "\pathname\percentiles_survey.macro.v1.1.sas";

*****;
*****;
*****;

/*2 macros built to implement UI, run mixtran, distrib, and estimate SEs: nut_UI and brrcalcs*/

%macro nut_UI (data, id, outcome, replicfirst, repliclast, replicvar, repeat, cpts, nctpts, outcome_ear, recco,
outcome_tul,subgroup);

%do replicnum = &replicfirst %to &repliclast;

%PUT "UI &outcome &replicnum";

%mixtran(data=&data, response=&outcome, foodtype=&outcome, subject=&id, repeat=&repeat,
covars_prob=, covars_amt= ,
outlib=outlib, modeltype=amount, lambda=, replicate_var=&replicvar&replicnum, seq=,
weekend=weekend, vargroup=, numvargroups=, start_val1=, start_val2=,
start_val3=, vcontrol=, nloptions=, titles=, printlevel=2, subgroup=&subgroup );
run;

data outlib._pred_unc_&outcome._&&replicnum;
set outlib._pred_unc_&outcome;
rep="&replicnum";
run;

```

```

data outlib._param_unc_&outcome._&&replicnum;
set outlib._param_unc_&outcome;
rep="&replicnum";
run;

data outlib.etas_&outcome._&&replicnum;
set outlib.etas_&outcome;
rep="&replicnum";
run;

data outlib._parmsf2_&outcome._&&replicnum;
set outlib._parmsf2_&outcome;
rep="&replicnum";
run;

%Distrib (seed=5454768, nsim_mc=1000, modeltype=amount, add_da=&data,
      pred=outlib._pred_unc_&&outcome._&replicnum, param=outlib._param_
      unc_&&outcome._&replicnum, outlib=outlib,
      cutpoints=&ctpts, recamt=&outcome_ear, recamt_co=&recco, recamt_hi=&outcome_tul,
      ncutpnt=&nctpts, byvar=, subject=seqn, titles=4, food=&outcome, subgroup=&subgroup);
run;

data outlib.brr_&&outcome._&replicvar&replicnum;
set outlib.descript_&&outcome._&replicvar&replicnum;
run=&replicnum;
run;

data outlib.brr_&outcome;
set outlib.brr_&&outcome._&replicvar&replicfirst-outlib.brr_&&outcome._&replicvar&repliclast;
run;

%end;

%MEND nut_UI;

*****;
*****;
*****;

%macro brrcalcs(outcome, paramnames);
%let param_names1= &paramnames;

data origwt_paramest_&outcome brr_paramest_&outcome;
  set outlib.brr_&outcome;
  if run=0 then output origwt_paramest_&outcome;
  else output brr_paramest_&outcome;
run;

```

title4 "Mixtran and Distrib BRR Means and SEs - Dietary Component:&outcome";

```
%brr_pvalue_ci(data_orig_estimates_1rec = origwt_paramest_&outcome,
  data_brr_estimates_brecs = brr_paramest_&outcome,
  param_estimate_names = &param_names1,
  set_f_method_fay = 0.3,
  set_confidence_level = 0.95,
  set_degrees_freedom = 31,
  print = y,
  titles = 4
);
```

```
%mend brrcalcs;
```

```
*****;
*****;
*****;
```

```
/*process data to generate 'data file' here*/
```

```
*
```

```
;
```

```
/*run mixtran, distrib and replicate*/
```

```
%nut_UI (data=datafile,outcome=calcium,id=seqn, repeat=daycode,
  replicfirst=0,repliclast=32,replicvar=psweight_ctpts=,
  nctpts=, outcome_ear=cal_hi,recco=GT,outcome_tul=,subgroup=);
```

```
*****call BRR macro to get SEs;
```

```
%brrcalcs(outcome=calcium, paramnames=mean_mc_t percent_rec_amt tpercentile5 tpercentile25 tpercentile50 tpercentile75 tpercentile95);
```

Appendix VI. Usual Intake Estimation—Ratio and Bivariate–Percentage of Energy From Fat Example

```

*****LIBRARY*****
*****CALL MACROS*****
/*Macros from NCI*/
%include "\pathname\boxcox_survey.macro.v1.1.sas";
%include "\pathname\std_cov_boxcox24hr_conday_minamt_macro_v1.0.sas";
%include "\pathname\nlmixed_bivariate.macro.v1.1.sas";
%include "\pathname\nlmixed_univariate.macro.v1.1.sas";
%include "\pathname\distrib_bivariate.macro.v1.1.sas";
%include "\pathname\brr_pvalue_ci_macro_v1.0.sas";
%include "\pathname\percentiles_survey.macro.v1.1.sas";

*****;
*****;
*****;

/*5 macros built to implement UI, run mixtran, distrib, and estimate SEs: nut_UI and brrcalcs*/
/*These 5 macros can be saved in separate SAS programs and called as NCI macros above*/

/*1: rununi runs the univariate models for the numerator and denominator of a ratio */

%macro rununi (data,foodvar,modeltype,nloptions,outlib,run);
/* get nhanes data and rename variables */

data biv;
set &data;
id      = seqn;
repeat  = daycode;
dayofweek = intake_day;

recall_food = &foodvar;

run;

proc sort data=biv; by id repeat; run;

```

```

/* create covariates for univariate food/nutrient model */

data biv;
  set biv;

if (repeat = 2) then repeat2 = 1;
else repeat2 = 0;
run;

proc sort data=biv; by id repeat; run;

%NLMixed_Univariate (data      = biv,
                      lambda    = ,
                      subject   = id,
                      repeat    = repeat,
                      response  = recall_food,
                      modeltype = &modeltype,
                      covars_prob = ,
                      covars_amt = repeat2 weekend ,
                      replicate_var = psweight_&run,
                      nloptions = &nloptions,
                      print     = Y,
                      ntitle   = 4);

/* save parameter estimates and predicted values. */
data outlib.parms_u_&foodvar._&run;
  set parms_u;
run;

data outlib.pred_u_&foodvar._&run;
  set pred_x_u;
run;

/* delete unneeded data sets

proc datasets nolist; delete biv b24_boxcox min_a parms_u pred_x_u ; run; quit;
*/
%mend rununi;

```

```
*****;
*****;
*****;
```

/*2: runbivar runs the NCI macro to estimate the bivariate model*/

```
%macro runbivar (data,foodvar1,foodvar2,modeltype,nloptions,outlib,run);
```

/* get nhanes data and rename variables */

```
data biv;
set &data;
id      = seqn;
repeat  = daycode;
dayofweek = intake_day;
recall_food1 = &foodvar1;
recall_food2 = &foodvar2;
run;
```

```
proc sort data=biv; by id repeat; run;
```

/* create covariates for univariate food/nutrient model */

```
data biv;
  set biv;

if (repeat = 2) then repeat2 = 1;
else repeat2 = 0;
run;
```

/* get minimum amount on consumption day. */

```
proc means data=biv noprint;
  where (recall_food1 > 0);
  var recall_food1;
  output out=min_a(keep=min_a) min=min_a;
run;
```

```
data biv;
  merge biv min_a;
  modeltype = upcase("“&modeltype”");
  if (modeltype = "ONEPART" & recall_food1 = 0) then recall_food1 = min_a / 2;
run;
```

```
data init_parms_f1;
  set outlib.parms_u_&foodvar1._&run;
```

```
/* rename univariate parameters for bivariate model. */
```

```
rename a_intercept      = A1_intercept  
      a_repeat2       = A1_repeat2  
      a_weekend        = A1_weekend  
      a_logsde         = A1_LogSDe  
      a_lambda          = A1_Lambda  
      logsdu2          = LogSDu2;  
run;
```

```
/* initial estimates for second food/nutrient. */
```

```
data init_parms_f2;  
set outlib parms_u_&foodvar2._&run;
```

```
/* rename univariate parameters for bivariate model. */
```

```
rename a_intercept      = A2_intercept  
      a_repeat2       = A2_repeat2  
      a_weekend        = A2_weekend  
      a_logsde         = A2_LogSDe  
      a_lambda          = A2_Lambda  
      logsdu2          = LogSDu3;  
run;
```

```
/* combine initial estimates for the two food/nutrients. */
```

```
data init_parms;  
merge init_parms_f1 init_parms_f2;  
  
keep A1_intercept A1_repeat2 A1_weekend A1_Lambda  
      A2_intercept A2_repeat2 A2_weekend A2_Lambda  
      A1_LogSDe A2_LogSDe  
      LogSDu2 LogSDu3;  
run;
```

```

/* create macro variables for the box-cox transformations for the two food/nutrients. */
data lambdas (keep=a1_lambda a2_lambda);
  merge init_parms_f1 init_parms_f2;
  run;

  call symput("lambda1",trim(left(put(a1_lambda,4.2)))); 
  call symput("lambda2",trim(left(put(a2_lambda,4.2)))); 
run;

proc print data=lambdas;
  var a1_lambda a2_lambda;
  title3 "Box-Cox Transformation Parameters";
  run;
title2;

%NLMixed_Bivariate (data      = biv,
                     init_parms = init_parms,
                     lambda1   = a1_lambda,
                     lambda2   = a2_lambda,
                     subject   = id,
                     repeat    = repeat,
                     response1 = recall_food1,
                     response2 = recall_food2,
                     modeltype = &modeltype,
                     covars_prob1 = ,
                     covars_amt1 = repeat2 weekend ,
                     covars_amt2 = repeat2 weekend ,
                     replicate_var = psweight_&run,
                     nloptions = &nloptions,
                     print     = Y,
                     ntitle    = 4);

/* save parameter estimates and predicted values. */
data outlib parms_b_&foodvar1._&foodvar2._&run;
  set parms_b;
  run;

data outlib pred_b_&foodvar1._&foodvar2._&run;
  set pred_x_b;
  run;

/* delete unneeded data sets.
proc datasets lib=work nolist;
  delete biv min_a1 min_a2 init_parms_f1 init_parms_f2 init_parms lambdas;
  run;
  quit;
*/
%mend runbivar;

```

```
*****;
*****;
*****;
```

/*3: rundist runs the bivariate distrib macro*/

```
%macro rundist(foodvar1, foodvar2, modeltype, seedv, cutpt1, cutpt2, outlib, run);
```

```
%global seed;
%let seed = &seedv;      /* seed for random number generator */
```

```
/* get parameter estimates and predicted values calculated by macro Food_Bivariate. */
```

```
data parms;
  set outlib.parms_b_&foodvar1._&foodvar2._&run;
run;
```

```
data pred;
  set outlib.pred_b_&foodvar1._&foodvar2._&run;
run;
```

```
proc sort data=pred; by id repeat; run;
```

```
/* for predicted values, create two records per subject, one for weekday and one for weekend. */
```

```
data pred;
  merge pred parms;
run;
```

```
data pred ;
  set pred;
  by id;
```

```
if (first.id); /* keep first record per subject. */
```

```
/* the following code assumes that "repeat2" is a covariate in the model:      */
/*  repeat2 = 0 for the first administered 24-hour dietary recall (drddaycd = 1).  */
/*  repeat2 = 1 for the second administered 24-hour dietary recall (drddaycd = 2). */
```

```
/* for each subject, calculate predicted value when repeat2 = 0. */
```

```
if (repeat2 = 1) then do;
  repeat = 1;
  repeat2 = 0;
  pred_x_a1 = pred_x_a1 - a1_repeat2;
  pred_x_a2 = pred_x_a2 - a2_repeat2;
end;
```

```

/* the following code assumes that "weekend" is a covariate in the model: */
/* weekend = 0 if the 24-hour dietary recall was for Monday-Thursday. */
/* weekend = 1 if the 24-hour dietary recall was for Friday-Sunday. */

/* create record for weekend = 0. */

if (weekend = 1) then do;
  weekend = 0;
  pred_x_a1 = pred_x_a1 - a1_weekend;
  pred_x_a2 = pred_x_a2 - a2_weekend;
  end;

day_wgt = 4/7;
output;

/* create record for weekend = 1. */

weekend = 1;
pred_x_a1 = pred_x_a1 + a1_weekend;
pred_x_a2 = pred_x_a2 + a2_weekend;
day_wgt = 3/7;
output;
run;

/* for each food/nutrient, create macro variable that equals half the minimum nonzero amount. */

data _null_;
  set pred;

min_a1 = min_a1 / 2;
call symput("min_a1",trim(left(put(min_a1,best12.))));

min_a2 = min_a2 / 2;
call symput("min_a2",trim(left(put(min_a2,best12.))));

run;

```

```
/* call macro Distrib_Bivariate to generate monte carlo distribution of usual intake. */
```

```
%Distrib_Bivariate (param      = parms,
                     predicted   = pred,
                     subject     = id,
                     modeltype   = &modeltype,
                     nsim_mc    = 1000,
                     day_wgt    = day_wgt,
                     min_a1     = &min_a1,
                     min_a2     = &min_a2,
                     print       = N,
                     ntitle     = 4);
```

```
options notes;
title2;
```

```
/* calculate usual nutrient density. */
```

```
data mcsim;
set _mcsim;
```

```
/* define nutrient density. */
```

```
*t_density = 1000 * (t1 / t2);      /* nutrient/food per 1000 kcal      */
*t_density = 100 * ((9 * t1) / t2); /* percent calories from saturated fat */
t_density=100*(t1/t2);
run;
```

```
/* call macro Percentiles_Survey to calculate percentiles of usual nutrient density. */
```

```
/* table 1: by age and gender */
```

```
title2 "Table 1: Percentiles by Age";
```

```
data mcsim2;
set mcsim;
run;
```

```
%Percentiles_Survey (data      = mesim2,
                     byvar    = ,
                     var      = t_density,
                     weight   = psweight_&run,
                     cutpoints = &cutpt1 &cutpt2,
                     print    = Y,
                     ntitle   = 4);
```

```
data pctl;
set _percentiles;
run;
```

```

/* print summary tables */

title2 "Estimated Mean and Percentiles";

proc print data=pctl label;
  id ;
  var Mean Pctile5 Pctile10 Pctile25 Pctile50 Pctile75 Pctile90 Pctile95;
  format Mean Pctile5 Pctile10 Pctile25 Pctile50 Pctile75 Pctile90 Pctile95 7.2;
run;

/* save means, percentiles and cut-point probabilities. */

data outlib.pctl_b_&foodvar1._&foodvar2._&run;
  set pctl;
  run=&run;
run;

/* delete unneeded data sets. */

proc datasets nolist;
  delete parms pred mcsim _mcsim mcsim2 pctl _percentiles;
  run;
quit;

%mend rundist;

*****;
*****;
*****;

/*4: bivrep repeats the univariate and bivariate runs for each replicate*/

%macro bivrep(data, run,outlib,foodvar1,foodvar2,cutpt1,cutpt2);
%do run=1 %to &run;
options nonotes;
%put BRR Run &run &foodvar1;

%rununi (data=&data,foodvar=&foodvar1,modeltype=ONEPART,nloptions=technique=tureg max-
func=1000 maxiter=1000,outlib=&outlib,
run=&run);
%rununi (data=&data,foodvar=&foodvar2,modeltype=ONEPART,nloptions=technique=tureg max-
func=1000 maxiter=1000,outlib=&outlib,
run=&run);

```

```
%runbivar (data=&data,foodvar1=&foodvar1, foodvar2=&foodvar2, modeltype=ONEPART, nloptions=tech-
nique=trureg maxfunc=10000 maxiter=1000,
outlib=&outlib,run=&run);

%rundist(foodvar1=&foodvar1, foodvar2=&foodvar2, modeltype=ONEPART, seedv=12345, outlib=&out-
lib,run=&run, cutpt1=&cutpt1, cutpt2=&cutpt2);

proc append base=brr_runs data=outlib.pctl_b_&foodvar1._&foodvar2._&run;

%end;

data outlib.brr_&foodvar1._&run;
set brr_runs;
rename mean=bmean pctile1-pctile99=bpctile1-bpctile99;
run;

%mend bivrep;
*****;
*****;
*****;

/*5: means_set appends the output from each replicate and uses the
brr_pvalue_ci macro to estimate SEs for the bivariate distribution*/



%macro means_set(foodvar1, foodvar2,replicfirst,repliclast, degfree,subgroup);
%put &foodvar1 ;

      data outlib.brr_biv_&foodvar1;
      set outlib.pctl_b_&foodvar1._&foodvar2._&replicfirst-outlib.pctl_b_&foodvar1._&foodvar2._&replic-
last;
      run;

data origwt_paramest_&foodvar1 brr_paramest_&foodvar1;
  set outlib.brr_biv_&foodvar1;
  if run=0 then output origwt_paramest_&foodvar1;
  else output brr_paramest_&foodvar1;
run;
```

```

*****;
*****;
*   BRR Calcs   *;
*****;
*****;

%brr_pvalue_ci(data_orig_estimates_1rec = origwt_paramest_&foodvar1,
      data_brr_estimates_brecs = brr_paramest_&foodvar1,
      param_estimate_names    = mean pctile5 pctile25 pctile50 pctile75 pctile95,
      set_f_method_fay       = 0.3,
      set_confidence_level   = 0.95,
      set_degrees_freedom    = &degfree,
      print                  = y,
      titles                 = 4
      );

%mend;

*****;
*****;
*****;

/*process data to generate ‘datafile’ here*/
;

/*run 2 univariate, bivariate, and bivariate distrib */

%rununi (data=datafile, foodvar=tfatkcal, modeltype=ONEPART, nloptions=technique=trureg maxfunc=10000
maxiter=1000, outlib=outlib,
run=0);

%rununi (data=datafile, foodvar=energy, modeltype=ONEPART, nloptions=technique=trureg maxfunc=10000
maxiter=1000, outlib=outlib,
run=0);

%runbivar (data=datafile, foodvar1=tfatkcal, foodvar2=energy, modeltype=ONEPART, nloptions=tech-
nique=trureg maxfunc=10000 maxiter=1000,
outlib=outlib, run=0);

%rundist (foodvar1=tfatkcal, foodvar2=energy, modeltype=ONEPART, seedv=12345, outlib=outlib, run=0);

*-----;
*-----      NOW, REPEAT THESE STEPS FOR THE 32 BRR RUNS AND GET SEs
-----;
*-----;

%bivrep(data=datafile, run=32, foodvar1=tfatkcal, foodvar2=energy, outlib=outlib, cutpt1=, cutpt2=);

%means_set(foodvar1=tfatkcal, foodvar2=energy, replicfirst=0, repliclast=32, degfree=31);

```

Appendix VII. Empirical Approach to Selecting a Correlated or Uncorrelated Model

Run both models and then test the parameter corresponding to the Fisher's Z-transformation of the correlation between random effects in the probability to consume and amount consumed parts of the for statistical significance using a *t* test where the standard error is computed via balanced repeated replication (BRR) (holding the lambda constant across BRR runs). If many dietary components are to be assessed this way, using the most parsimonious model for each component, the set of tests should be adjusted for multiple comparisons (e.g., using the methods described in a report by Benjamini & Hochberg, (1995) [64]). Briefly, this method uses a false discovery rate procedure to test the strength of the correlation between the two parts of the model (54) and allows analysts to proceed with the correlated model in cases where there is strong evidence of a significant correlation. If there is truly no correlation, using the correlated model will generally lead to more variable estimates. On the other hand, the simpler uncorrelated model will generally have less variability but more bias if the correlation is truly present (54). Implementing the Benjamini & Hochberg adjustment involves ranking the $K p$ values for the K individual tests from smallest to largest, selecting a value for the desired false Discovery Rate (e.g., $p = .10$), and then for $m = 1, \dots, K$, reject null hypothesis if m -th smallest p value is less than $(.10^*m)/K$.

Vital and Health Statistics

Series Descriptions

Active Series

Series 1. Programs and Collection Procedures

Reports describe the programs and data systems of the National Center for Health Statistics, and the data collection and survey methods used. Series 1 reports also include definitions, survey design, estimation, and other material necessary for understanding and analyzing the data.

Series 2. Data Evaluation and Methods Research

Reports present new statistical methodology including experimental tests of new survey methods, studies of vital and health statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. Reports also include comparison of U.S. methodology with those of other countries.

Series 3. Analytical and Epidemiological Studies

Reports present data analyses, epidemiological studies, and descriptive statistics based on national surveys and data systems. As of 2015, Series 3 includes reports that would have previously been published in Series 5, 10–15, and 20–23.

Discontinued Series

Series 4. Documents and Committee Reports

Reports contain findings of major committees concerned with vital and health statistics and documents. The last Series 4 report was published in 2002; these are now included in Series 2 or another appropriate series.

Series 5. International Vital and Health Statistics Reports

Reports present analytical and descriptive comparisons of U.S. vital and health statistics with those of other countries. The last Series 5 report was published in 2003; these are now included in Series 3 or another appropriate series.

Series 6. Cognition and Survey Measurement

Reports use methods of cognitive science to design, evaluate, and test survey instruments. The last Series 6 report was published in 1999; these are now included in Series 2.

Series 10. Data From the National Health Interview Survey

Reports present statistics on illness; accidental injuries; disability; use of hospital, medical, dental, and other services; and other health-related topics. As of 2015, these are included in Series 3.

Series 11. Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey

Reports present 1) estimates of the medically defined prevalence of specific diseases in the United States and the distribution of the population with respect to physical, physiological, and psychological characteristics and 2) analysis of relationships among the various measurements. As of 2015, these are included in Series 3.

Series 12. Data From the Institutionalized Population Surveys

The last Series 12 report was published in 1974; these reports were included in Series 13, and as of 2015 are in Series 3.

Series 13. Data From the National Health Care Survey

Reports present statistics on health resources and use of health care resources based on data collected from health care providers and provider records. As of 2015, these reports are included in Series 3.

Series 14. Data on Health Resources: Manpower and Facilities

The last Series 14 report was published in 1989; these reports were included in Series 13, and are now included in Series 3.

Series 15. Data From Special Surveys

Reports contain statistics on health and health-related topics from surveys that are not a part of the continuing data systems of the National Center for Health Statistics. The last Series 15 report was published in 2002; these reports are now included in Series 3.

Series 16. Compilations of Advance Data From Vital and Health Statistics

The last Series 16 report was published in 1996. All reports are available online; compilations are no longer needed.

Series 20. Data on Mortality

Reports include analyses by cause of death and demographic variables, and geographic and trend analyses. The last Series 20 report was published in 2007; these reports are now included in Series 3.

Series 21. Data on Natality, Marriage, and Divorce

Reports include analyses by health and demographic variables, and geographic and trend analyses. The last Series 21 report was published in 2006; these reports are now included in Series 3.

Series 22. Data From the National Mortality and Natality Surveys

The last Series 22 report was published in 1973. Reports from sample surveys of vital records were included in Series 20 or 21, and are now included in Series 3.

Series 23. Data From the National Survey of Family Growth

Reports contain statistics on factors that affect birth rates, factors affecting the formation and dissolution of families, and behavior related to the risk of HIV and other sexually transmitted diseases. The last Series 23 report was published in 2011; these reports are now included in Series 3.

Series 24. Compilations of Data on Natality, Mortality, Marriage, and Divorce

The last Series 24 report was published in 1996. All reports are available online; compilations are no longer needed.

For answers to questions about this report or for a list of reports published in these series, contact:

Information Dissemination Staff
National Center for Health Statistics
Centers for Disease Control and Prevention
3311 Toledo Road, Room 4551, MS P08
Hyattsville, MD 20782

Tel: 1–800–CDC–INFO (1–800–232–4636)
TTY: 1–888–232–6348

Internet: <https://www.cdc.gov/nchs>
Online request form: <https://www.cdc.gov/info>

For e-mail updates on NCHS publication releases, subscribe online at: <https://www.cdc.gov/nchs/govdelivery.htm>.

**U.S. DEPARTMENT OF
HEALTH & HUMAN SERVICES**

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road, Room 4551, MS P08
Hyattsville, MD 20782-2064

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300

FIRST CLASS MAIL POSTAGE & FEES PAID CDC/NCHS PERMIT NO. G-284
