Article

Not peer-reviewed version

Evaluating Usual Intake Estimation Methods

Xiaoshu Zhu * , Christine Borger * , Jill DeMatteis , Brenda Sun

Posted Date: 4 November 2025

doi: 10.20944/preprints202511.0233.v1

Keywords: Usual intake (UI) estimation; National Cancer Institute (NCI) methods; Markov Chain Monte Carlo (MCMC

Article

Evaluating Usual Intake Estimation Methods

Xiaoshu Zhu [1,*], Christine Borger [1,*], Jill DeMatteis [2] and Brenda Sun [1]

[1] Westat, USA
[2] ForsMarsh, USA
* Correspondence: xiaoshuzhu@westat.com (X.Z.); christineborger@westat.com (C.B.)

Abstract

Objectives: This paper illustrates the use of the National Cancer Institute (NCI) Markov Chain Monte Carlo (MCMC) method for usual intake (UI) analyses of dietary intake data, comparing results with three other methods: a single-day dietary recall, the average of 2 days of recalls, and univariate or bivariate NCI models. **Methods:** Using 24-hour dietary recall data from the Infant and Toddler Feeding Practices Study-2 (ITFPS-2), researchers estimated children's UIs of sodium, added sugar, and whole grains individually or jointly using each of four methods. They compared results for two types of analyses: (1) population distributions of single dietary constituents or ratios of constituents, and (2) regression models evaluating the association between dietary intake and covariates. **Results:** Except for episodically consumed whole grains, the MCMC method and the univariate and bivariate NCI models yielded comparable estimates for both the population distribution and regression analyses. The distribution estimated by the univariate and bivariate NCI methods was closer to normal than the non-NCI estimation methods. Even though the methods yielded regression models that disagreed on the significance of several covariates, none of the differences were significant across the four methods. **Conclusions:** The MCMC method simultaneously models the UIs of multiple dietary constituents. This study demonstrates its applicability as an alternative for a range of diet quality and intake evaluations. Compared with the other NCI methods examined in this study, the MCMC method outperforms by delivering a shorter computing time, requiring smaller storage, and experiencing fewer convergence failures.

Keywords: Usual intake (UI) estimation; National Cancer Institute (NCI) methods; Markov Chain Monte Carlo (MCMC)

1. Introduction

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) is a public health nutrition program that safeguards the health of women, infants, and children in low-income households who are at nutritional risk (U.S. Department of Agriculture [USDA] & Food and Nutrition Service, 2022). The program currently serves about 7 million women, infants, and children (Kline et al., 2022). Between 2011 and 2025, USDA funded a nationally representative, longitudinal study called the Infant and Toddler Feeding Practices Study (ITFPS-2). The study, now known as WIC ITFPS-2, followed a prospective cohort from birth through age 9, focusing on caregiver feeding practices and children's diet-related health outcomes in the context of WIC service provision (Harrison et al., 2014). To address the study's research objectives (Borger et al., 2025), researchers needed an analytic method that not only accurately reflects children's diet quality but also captures the relationship between children's dietary outcomes and family and environmental factors.

Dietary analysis is essential for understanding the relationship between nutrition and human health. Assessment of diet quality requires following people's long-term consumption to determine usual intake (UI), which accounts for day-to-day variation. UI reflects the interactions among multiple foods and beverages reported in estimates for nutrients, energy, and other dietary components. Even though self-report measures commonly involve extensive measurement error

doi:10.20944/preprints202511.0233.v1

(Kirkpatrick & Collins, 2016), intake estimates tend to be less biased when using short-term instruments, such as 24-hour dietary recalls and food records or diaries, than long-term instruments like food frequency questionnaires, which measure the frequency of consumption directly (Tooze, 2020). Over decades, various statistical models have been proposed and developed to better understand and address measurement error in dietary intake estimation from 24-hour dietary recalls (Dodd et al., 2006; Nusser et al., 1995; Tooze et al., 2006; Zhang et al., 2011a, 2011b). Some models attempt to tackle the issues of within-person variation and episodically consumed foods, including the methods developed by the National Cancer Institute (NCI).

The purpose of this study was to illustrate the use of the NCI's Markov Chain Monte Carlo (MCMC) method for various types of UI analyses, comparing MCMC UI estimates with results from other NCI methods. The comparison between the MCMC method and other NCI methods is compelling because it informs the possibility of using the MCMC method to support more general dietary intake estimation and assessment. Moreover, NCI recently decided that it would no longer update the univariate and bivariate approaches, and it is important to understand the implications of this decision.

1.1. The NCI Approach to Estimating UI

The NCI approach accounts for measurement error in dietary intake by separating within-person variation from between-person variation. The NCI approach requires that at least a subset of respondents provide two or more recalls to enable estimation of usual dietary intake patterns. It accommodates episodically or rarely consumed foods using a two-part model in which the probability of consumption is estimated in addition to the amount consumed. The technique allows researchers to incorporate covariates, facilitating the use of covariates in generating UI estimates for subpopulations (Tooze et al., 2010), and to assess the effect of UI on health outcomes (Kipnis et al., 2009).

The UI estimation starts by fitting the NCI measurement error model (Brassard, 2022; NCI, 2013) to the intake data to obtain parameter estimates. Subsequently, the approach generates a pseudo population by drawing a random sample from the distribution using the parameter estimates from the first step. Lastly, it involves calculating desired statistics or running models using the pseudo population. For data obtained from a complex sample survey such as WIC ITFPS-2 or the National Health and Nutrition Examination Survey (NHANES), these steps incorporate sampling weights and are repeated across replicates created for variance estimation so that estimates accurately represent the population.

NCI has created different SAS macros for different types of dietary analyses. Selecting a macro for a given analysis depends on the number of dietary constituents involved in UI estimation (i.e., single, two, multiple) and how frequently they are consumed (i.e., regularly, episodically, rarely). NCI discusses which macros to select for specific types of analyses (NCI, n.d., 2024).

For multivariate analyses (i.e., estimating >2 dietary components), NCI extended its methodology to develop the MCMC method (Zhang et al., 2011a, 2011b). It simultaneously models multiple constituents, which could be a mixture of daily, episodically, or rarely consumed foods and beverages, accounting for correlations among them. In the MCMC method, researchers fit a multivariate measurement error model using the MCMC algorithm. Then, they use the resulting parameter estimates to generate a multivariate Monte Carlo draw of UI estimates for the multiple dietary constituents. This method can capture the complexity and dynamics of diet behaviors and variability in food and nutrition. One promising application of the MCMC method is estimating the Healthy Eating Index (HEI) scores, which are functions of multiple dietary components, including total energy (Gelfand & Tangney, 2021; Hutchinson et al., 2023; Wambogo et al., 2020).

1.2. The MCMC Method for HEI Calculation

The HEI is a widely used tool for assessing diet quality in relation to the recommendations in the Dietary Guidelines for Americans (DGA) (USDA & U.S. Department of Health & Human

Services, 2020). The HEI is updated every 5 years, and the latest version is the HEI-2020, which provides HEI scores for individuals aged 2 years or older (National Institutes of Health et al., 2025).[1] Shortly after the most recent update to HEI, a new HEI-Toddlers-2020 for children under age 2 was released. The overall HEI-2020 and HEI-Toddlers-2020 scores are the sum of 13 component scores. Nine of 13 components are categorized as adequacy components, which means that there is no limit on the amount recommended for consumption, and people are encouraged to eat more for good health. Four of the 13 components are categorized as moderation components, which means that limits are recommended for good health. All but one component is density based (i.e., intake amounts per 1,000 kcal). The exception is the score for fatty acids, for which the HEI component is expressed as a ratio of poly- and monounsaturated to saturated fatty acids. The process of calculating HEI scores involves estimating the intake of multiple dietary constituents consumed either daily or episodically. Given the multivariate nature of HEI, the NCI recommends the MCMC method because it can jointly account for relationships among dietary constituents, including energy.

Because the MCMC method does not allow overlap among dietary constituents, the HEI components must be disaggregated into mutually exclusive constituents for UI estimation and then added together to form components before scoring. For example, legumes contribute to five adequacy components and are modeled as individual constituents in the macro. Researchers must identify and separate any dietary constituents consumed daily or episodically. They can also include covariates such as demographic characteristics and feeding practices to model the probability of consumption and the consumption quantity. The detailed steps and SAS macros they used include the following:

1. Obtain the Box-Cox transformation parameters for each dietary constituent using the BOXCOX_SURVEY macro with the full-sample weights
2. Run the STD_COV_BOXCOX24HR_CONDAY_MINAMT macro to prepare the input file for MULTIVAR_MCMC
3. Using the file outputted in step 2, run the MULTIVAR_MCMC macro with the adjusted full-sample weights (i.e., original weight * total number of observed individuals/sum of original weights) to fit the multivariate measurement error model (NCI, 2013), which corrects for error in dietary intake data
4. Run the MULTIVAR_DISTRIB macro using the parameter estimates from step 3 to generate a pseudo population with 100 pseudo persons per 1 observed individual
5. Aggregate HEI components and calculate HEI scores for each individual in the pseudo population
6. Calculate the distribution estimates and run regression models with adjusted weights (i.e., the weights outputted from MULTIVAR_MCMC divided by the number of pseudo persons per observed individual [100]), and store the point estimates
7. Repeat steps 3 to 6 in each replicate (using the replicate weight provided in the input data file in place of the full-sample weight)
8. Calculate standard errors (SEs) using the estimates from the full-sample run and the 40 replicate estimates

For documentation on the SAS macros, refer to the MCMC macro user guide (NCI, 2013).

At step 4, the output file contains the intake estimates for every dietary constituent later used to calculate the HEI component and total scores for each pseudo person generated from the MCMC method. It is unknown how using those estimates for other types of UI analysis (besides HEI calculation) compares with using UI estimates with the bivariate or univariate UI macros. For example, if the goal is to estimate the population distribution of a single dietary constituent or a ratio of two constituents or explore the relationship between dietary constituents and/or covariate information in a regression model, how do the results based on the MCMC method compare with those based on the bivariate or univariate UI modeling approaches?

---

[1] The standards for the HEI-2020 are the same as in the HEI-2015 as a result of insignificant changes in USDA Dietary Patterns.

2. Methods

WIC ITFPS-2 is a prospective cohort study that follows children from around the time of birth through age 6, with a follow-up at age 9. Harrison and colleagues (2014) describe the initial study design; Borger and colleagues (2022) present the study design after multiple study extensions. This study used 24-hour dietary recall data collected when the children in WIC ITFPS-2 were 5 years old and survey information from the longitudinal sample collected in multiple interviews through age 5.

2.1. Analytic Sample

WIC ITFPS-2 recruited study participants from study-eligible clinics during fall 2013 and followed them through the child's sixth year, with a follow-up at age 9. Eligible participants were caregivers enrolling in WIC at a study-eligible clinic for the first time for their current pregnancy or their newborn; caregivers were at least 16 years old and spoke English or Spanish. Siegfried and colleagues (2023) describe the sampling and recruitment processes in detail.

The longitudinal sample used in this study comprises individuals who responded to every postnatal interview beginning with the 1- or 3-month interview (depending on the child's age at WIC enrollment) through the fifth-year interview. Additionally, the current study used 24-hour dietary recall information from 1,030 children in the longitudinal sample, 11.8% of whom had second-day recalls; these dietary recalls were collected when the study children were aged 5 years.

Table 1 presents the sample characteristics. Data are weighted by statistical survey weights (full-sample weights and 40 replicate weights) that account for unequal selection probabilities and nonresponse to make inferences for the study-eligible population. Cases with missing values were excluded from the sample for consistency across all the methods discussed.

Table 1. Descriptive statistics of the analytic sample (N = 1,030).

Continuous characteristic		**Mean (SD)**	**Min**	**Max**
Mother's age when giving birth		27.1 (5.7)	16	47
Age of the infant (in days) when the mother stopped breastfeeding		131.4 (148.2)	0	410
Number of snacks during the day		1.8 (1.1)	0	7
Categorical characteristics	**Level**	**N**	**%**	
	Demographic			
Baby's sex	Male	534	51.84	
	Female	496	48.16	
Caregiver's race	African American	278	26.99	
	Other	601	58.35	
	White	151	14.66	
Caregiver's ethnicity	Hispanic or Latino	400	38.83	
	Non-Hispanic or non-Latino	630	61.17	
Caregiver's education level	High school or less	528	51.26	
	More than high school	502	48.74	
Marital status	Married	438	42.52	
	Not married	592	57.48	
Parity	Firstborn	396	38.45	

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2025 doi:10.20944/preprints202511.0233.v1

	Second born	301	29.22
	Third or subsequent born	333	32.33
Currently using regular childcare	Yes	654	63.5
	No	376	36.5
WIC/SNAP participation			
SNAP participation status	Yes	406	39.42
	No	624	60.58
WIC and SNAP participation status	On WIC and SNAP	307	29.81
	On WIC only	264	25.63
	On SNAP only	136	13.2
	On neither	323	31.36
Pattern of WIC participation	1 year or less	95	9.22
	2–3 years	198	19.22
	4–5 years	153	14.85
	Consistently	443	43.01
	Intermittently	141	13.69
Feeding practices			
When solid foods were introduced	Before 4 months	280	27.18
	After 4 months	750	72.82
When sweet beverages were introduced	In child's first year	625	60.68
	In child's second year	226	21.94
	Not in child's first 2 years	179	17.38
Categorical characteristics	**Level**	**N**	**%**
When sweets were introduced	In child's first year	760	73.79
	In child's second year	173	16.8
	Not in child's first 2 years	97	9.42
When salty snacks were introduced	In child's first year	891	86.5
	In child's second year	81	7.86
	Not in child's first 2 years	58	5.63
TV on while eating	Most or sometimes	540	52.43
	Never or rarely	490	47.57
Family eats together per week	0–4 times	401	38.93
	5 or more times	629	61.07
Usual number of hours child sleeps	Less than 10 hours	186	18.06
	At least 10 hours	844	81.94

SD = standard deviation; SNAP = Supplemental Nutrition Assistance Program.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2025 doi:10.20944/preprints202511.0233.v1

2.2. Dietary Recall Data

WIC ITFPS-2 collected 24-hour dietary recall information using USDA's Automated Multiple-Pass Method (AMPM). Raper and colleagues (2004) detail the AMPM's five-step approach. WIC ITFPS-2 used USDA's Food and Nutrient Database for Dietary Studies 5.0 (FNDDS5) as the source of the nutrient values of food reported (Ahuja et al., 2012).

2.3. Analysis Design

The use of the MCMC method focused on two types of analyses: (1) estimating the distribution of intake in the population overall and among subpopulations, and (2) exploring the association between dietary intake and individual characteristics. Individual characteristics of particular interest for WIC ITFPS-2 included the child's WIC participation, assessed either at a point in time or over time. Borger and colleagues (2020, 2022) and Au and colleagues (2025) indicate that most children consume the foods in the WIC food package, suggesting that nutrition outcomes would be associated with WIC participation. WIC participation over time is reflected in the child's pattern of WIC participation, a variable that is included in models assessed. This study, therefore, included WIC participation as a focal individual characteristic.

The three dietary constituents selected for univariate analyses were added sugar, sodium, and whole grains. Among 5-year-old children, the first two are consumed daily, whereas the latter is episodically consumed. Each of the dietary constituents contributes to one of the HEI component scores (the first two are moderation components, and the latter is an adequacy component). Moreover, over- or underconsumption of these three dietary constituents has been associated with different health outcomes. For instance, strong evidence supports the association between added sugar consumption and dental caries in children (Chi & Scott, 2019), and high sugar intake is associated with reduced cognitive function in children (Gillespie et al., 2024). Evidence also shows that high sodium intake in childhood is associated with elevated blood pressure (Leyvraz et al., 2018). With increased consumption of whole grains, diet quality and nutrient intake significantly improve in children (O'Neil et al., 2011). As the denominator in all HEI component scores, dietary energy (calories) was also included in the univariate analysis.

The study assessed results from four modeling efforts using NCI and non-NCI methods:

- For univariate analysis, a linear regression model was fit to assess sodium intake as a function of the child's sex, parity, the caregiver's demographic characteristics, timing of food introduced to the child, and WIC or Supplemental Nutrition Assistance Program (SNAP) participation status.
- For bivariate analysis, the sodium model was extended by adding energy as a control of total dietary intake.
- A second model involving bivariate analysis was a logistic regression with a binary outcome determined by whether energy from added sugar was below (value = 1) or above (value = 0) 10% of total energy as recommended by the DGA, with the same set of covariates used in the linear regression model for sodium.
- Multivariate analysis focused on the total HEI score, its distribution in the population and subpopulations, and its association with the covariates.

Table 2 summarizes the intake methods considered for each type of analysis. All statistical analyses were conducted in SAS 9.4.

Table 2. Intake methods for the three main types of analysis.

	Analysis type		Intake method			
	Population distribution overall and by subpopulation	**Regression modeling association between intake and covariates**	**Single-day recall**	**Average of 2 days of recalls**	**Multivariate MCMC**	**Other NCI method**
Univariate	Added sugar, whole grains, sodium, and energy	Linear regression model with sodium as a function of covariates	✓	✓	✓	Univariate macros
Bivariate	Ratio of added sugar, sodium, and whole grains to energy	Linear regression model with sodium as a function of covariates, after controlling for energy	✓	✓	✓	Bivariate macros
		Logistic regression model with a binary indicator of whether intake met DGA recommendation for added sugars				
Multivariate	HEI scores	Linear regression model with HEI as a function of covariates	✓	✓	✓	N/A

3. Results

3.1. Population Distribution Overall and by Subpopulation

3.1.1. Univariate Distribution

Table 3 presents the mean and quartiles of intake estimates for the three dietary constituents—added sugars, whole grains, and sodium—as well as for energy. The results show differences between the two non-NCI methods (single-day recall and average of 2 days' recall) and the two NCI methods (the univariate NCI method and the MCMC method). The NCI methods result in less variability in the distribution of intakes than the non-NCI methods. Median intakes are closer among the four methods, but the mean intakes show the smallest difference. Surprisingly, whole grains are an exception. Estimated whole grain intake is much higher for the univariate NCI method than for the MCMC method and the two non-NCI methods across the mean and all three quartiles. The two NCI methods typically produce slightly larger values for the SEs than the non-NCI methods. The difference is particularly striking given how close the SEs are for the means. However, the differences are within one SE of each other for the four methods regarding added sugar, sodium, and energy.

Table 3. Mean and quartiles of intake estimation for four dietary constituents by estimation methods.

Dietary constituents	Intake estimation method	Mean (SE)	First quartile (SE)	Median (SE)	Third quartile (SE)
Added sugar (tsp equivalent [eq])	1 day	9.86 (0.31)	4.38 (0.21)	8.05 (0.23)	13.20 (0.43)
	Average of 2 days	9.87 (0.30)	4.40 (0.20)	7.97 (0.28)	13.37 (0.44)

 doi:10.20944/preprints202511.0233.v1

	Univariate NCI macros	9.93 (0.30)	6.30 (0.41)	9.13 (0.28)	12.68 (0.49)
	MCMC	10.06 (0.30)	6.93 (0.68)	9.48 (0.37)	12.53 (0.56)
Whole grains (oz eq)	1 day	0.76 (0.04)	0.00 (0.05)	0.52 (0.04)	1.19 (0.09)
	Average of 2 days	0.75 (0.04)	0.00 (0.05)	0.53 (0.04)	1.14 (0.08)
	Univariate NCI macros	1.07 (0.04)	0.84 (0.09)	1.03 (0.05)	1.25 (0.06)
	MCMC	0.78 (0.04)	0.45 (0.07)	0.73 (0.05)	1.04 (0.07)
Sodium (mg)	1 day	2,566.49 (56.92)	1,774.19 (53.49)	2,433.96 (64.96)	3,158.12 (65.53)
	Average of 2 days	2,563.10 (55.72)	1,763.91 (53.52)	2,427.50 (69.15)	3,158.74 (60.78)
	Univariate NCI macros	2,572.47 (56.91)	2,135.13 (74.02)	2,520.23 (58.04)	2,949.04 (80.31)
	MCMC	2,583.39 (60.03)	2,200.63 (92.06)	2,544.68 (64.41)	2,925.16 (75.71)
Total energy (kcal)	1 day	1,586.71 (26.06)	1,172.97 (25.06)	1,527.09 (28.10)	1,932.23 (33.09)
	Average of 2 days	1,589.50 (25.99)	1,175.05 (25.76)	1,530.77 (27.98)	1,936.28 (35.10)
	Univariate NCI macros	1,594.69 (26.23)	1,323.99 (38.60)	1,561.62 (27.21)	1,828.58 (38.80)
	MCMC	1,597.94 (27.48)	1,362.16 (47.51)	1,576.70 (30.58)	1,809.58 (38.05)

3.1.2. Distribution of Ratios for HEI Component Scores

Because the univariate model cannot estimate ratios, Table 4 summarizes the mean and quantile estimates for the added sugars, whole grains, and sodium HEI components from four different approaches. Both the NCI and non-NCI methods yield similar estimates of the mean ratio scores, including both point estimates and SEs. The difference among the quartiles is consistent with the pattern observed in univariate analysis, whereby the bivariate NCI macro and the MCMC method produce higher estimates on the first quartile and median, but smaller estimates on the third quartile. The SEs from the two NCI methods are larger than those of the non-NCI methods for the mean and first quartile.

Table 4. Mean and quartiles of three ratios for the HEI component scores by intake estimation methods.

HEI component	Intake estimation method	Mean (SE)	First quartile (SE)	Median (SE)	Third quartile (SE)
Added sugars (tsp eq per 1,000 kcal)	1 day	9.74 (0.22)	4.92 (0.19)	8.37 (0.27)	13.34 (0.36)
	Average of 2 days	9.72 (0.21)	4.89 (0.19)	8.35 (0.25)	13.34 (0.32)
	Bivariate NCI macros	9.92 (0.25)	6.80 (0.36)	9.29 (0.23)	12.38 (0.36)
	MCMC	9.90 (0.24)	7.56 (0.68)	9.61 (0.28)	11.93 (0.55)
Whole grains (oz eq per 1,000 kcal)	1 day	0.53 (0.03)	0.00 (0.03)	0.32 (0.03)	0.83 (0.05)
	Average of 2 days	0.52 (0.02)	0.00 (0.03)	0.34 (0.03)	0.80 (0.05)
	Bivariate NCI macros	0.50 (0.03)	0.27 (0.06)	0.44 (0.03)	0.66 (0.05)
	MCMC	0.50 (0.04)	0.29 (0.05)	0.46 (0.03)	0.66 (0.06)
Sodium (mg per 1,000 kcal)	1 day	1.63 (0.02)	1.36 (0.01)	1.58 (0.02)	1.85 (0.02)
	Average of 2 days	1.63 (0.02)	1.36 (0.01)	1.58 (0.02)	1.83 (0.03)
	Bivariate NCI macros	1.62 (0.02)	1.50 (0.03)	1.61 (0.02)	1.74 (0.03)
	MCMC	1.63 (0.02)	1.48 (0.02)	1.62 (0.02)	1.76 (0.02)

3.1.3. Percentage Meeting DGA Recommendation for Added Sugars

The methods examined in this study disagree on estimates of the portion of the study population with added sugars intake below the DGA recommendation of less than 10% of energy. The two NCI methods yield a smaller percentage (bivariate NCI macros, 56.7% [SE = 2.0 percentage points, or pp] and MCMC, 54.8% [SE = 2.3 pp]) than the two non-NCI methods (single-day recall, 61.7% [SE = 1.8 pp] and average of 2 days' recall, 61.4% [SE = 1.9 pp]). The percentage is consistently the smallest for the MCMC method, even though the difference is within one SE of the bivariate NCI method.

3.1.4. Distribution of HEI Total Scores

Because the univariate and bivariate NCI approaches are not sufficient for multiple dietary components, Table 5 summarizes the mean and quantile estimates of the total HEI scores by the three methods: using only 1 day of recall, using the average of 2 days of recall, and using the MCMC method. The point estimates and SEs in the MCMC method are consistently larger than those in the non-NCI methods for the mean and the three quartiles. The difference in the total HEI score between the MCMC method and the two non-NCI methods ranges from 1.91 on the third quartile to 6.48 on the first quartile.

Table 5. Mean and quartiles of HEI total scores by intake estimation methods.

Intake estimation method	Mean (SE)	First quartile (SE)	Median (SE)	Third quartile (SE)
1 day	55.23 (0.54)	46.56 (0.76)	55.33 (0.67)	63.61 (0.67)
Average of 2 days	55.43 (0.55)	46.84 (0.71)	55.55 (0.64)	63.61 (0.69)

 doi:10.20944/preprints202511.0233.v1

MCMC	59.25 (0.98)	53.04 (1.75)	59.30 (0.97)	65.52 (1.73)

The same patterns observed in the overall distributions are also seen in subpopulations, such as estimates by WIC participation pattern (see Supplementary Figure S4).

3.2. Modeling Association Between Intake and Covariates

3.2.1. Linear Model With Sodium Intake as an Outcome

The univariate regression analysis explores the association of sodium consumption with demographic characteristics and feeding practice variables. Table 6 presents the covariates with significant or opposite signs of regression coefficient estimates across the four methods. The SEs all show similar magnitude. All four methods reached the same results regarding significance tests of association between the covariates and sodium consumption. All significant effects point in the same direction across the methods, and no difference in parameter estimates between methods is more than one SE. There is no apparent pattern or method that produces consistently higher values across the covariates. Though not statistically significant, the two non-NCI methods show that children receiving both WIC and SNAP benefits have nominally lower mean sodium consumption versus children who participated in neither of the programs, whereas the two NCI methods demonstrate the opposite results.

Table 6. Coefficients with significant or opposite signs in the regression model, with sodium consumption as the outcome.

Variable		**1 day**	**Average of 2 days**	**Univariate NCI macros**	**MCMC**
		Point estimate (SE)			
Intercept		**2,607.86 (409.12)**	**2,472.29 (407.84)**	**2,319.28 (372.07)**	**2,375.59 (363.44)**
Baby's sex	Female	**−202.33 (64.42)**	**−243.97 (69.89)**	**−251.83 (69.10)**	**−234.84 (72.05)**
	Male (reference)	—	—	—	—
Caregiver's education level	High school or less	55.67 (101.41)	53.30 (102.92)	9.91 (105.13)	−1.87 (97.02)
	More than high school (reference)	—	—	—	—
Marital status	Not married	**229.11 (82.37)**	**238.88 (83.09)**	**205.12 (81.78)**	**196.42 (81.23)**
	Married (reference)	—	—	—	—
WIC and SNAP participation status	On WIC and SNAP	−23.06 (98.04)	−6.35 (97.48)	37.83 (100.94)	22.98 (104.78)
	On WIC only	−160.79 (115.62)	−99.02 (120.32)	−14.30 (117.95)	−36.72 (121.28)
	On SNAP only	−160.88 (159.65)	−165.33 (156.95)	−94.36 (153.16)	−122.34 (158.59)

Variable		1 day	Average of 2 days	Univariate NCI macros	MCMC
		Point estimate (SE)			
	On neither (reference)	—	—	—	—
Mother's age when giving birth		**−14.76 (6.14)**	**−13.09 (6.19)**	**−13.38 (6.05)**	**−13.20 (6.24)**
Number of snacks during the day		**124.64 (29.20)**	**118.74 (31.93)**	**102.53 (32.78)**	**104.50 (30.55)**

Notes: Statistically significant coefficients (p < .05) are marked in bold. The covariates with nonsignificant coefficients are caregiver's race, caregiver's ethnicity, parity, when solid foods were introduced, when salty snacks were introduced, and age of infant (in days) when the mother stopped breastfeeding. The estimates for all covariates are available in the Supplementary Tables.

3.2.2. Linear Model With Sodium Intake as an Outcome, Controlling for Total Energy

In the univariate analysis, energy is added to the model as an extension of the regression model. The regression coefficients of the covariates then reflect the effect after controlling for the total intake (see Table 7). The four methods disagree on the significant effects. For instance, the number of snacks during the day is significant in the two non-NCI methods only. Parity appears to have a significant effect only when using the average of 2 days of recalls; a significant group difference is found between the firstborn and the third or subsequent born. The non-NCI methods and the two NCI methods also disagree on the direction of group differences. For example, the non-NCI methods show that children not introduced to salty snacks have higher energy-adjusted sodium intake levels than children introduced to salty snacks in the first 2 years of life. However, the two NCI methods nominally suggest the opposite, though this difference is not statistically significant in any of the four methods.

Table 7. Coefficients with significant or opposite signs in the regression model, with sodium consumption as the outcome and energy as a control variable.

Variable		1 day	Average of 2 days	Bivariate NCI macros	MCMC
		Point estimate (SE)			
Intercept		**474.26 (211.65)**	**470.27 (210.23)**	631.56 (412.32)	**749.13 (302.59)**
Parity	Firstborn	116.51 (60.36)	**130.83 (57.42)**	92.29 (53.63)	88.78 (49.66)
	Second born	−13.27 (49.19)	0.20 (47.51)	−7.71 (50.75)	−8.30 (48.12)
	Third or subsequent born (reference)	—	—	—	—
When solid foods were	Before 4 months	−12.88 (43.85)	−9.22 (43.87)	−2.59 (54.41)	3.35 (48.47)
	After 4 months (reference)	—	—	—	—

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 November 2025 doi:10.20944/preprints202511.0233.v1

Variable		1 day	Average of 2 days	Bivariate NCI macros	MCMC
		Point estimate (SE)			
Introduced					
When salty snacks were introduced	In child's first year	−116.96 (118.05)	−106.51 (120.29)	12.96 (175.06)	21.26 (117.84)
	In child's second year	−110.16 (146.24)	−92.60 (149.14)	0.42 (181.06)	5.11 (146.32)
	Not in child's first 2 years (reference)	—	—	—	—
Mother's age when giving birth		**−8.30 (4.09)**	**−8.19 (3.66)**	**−11.18 (3.61)**	**−10.93 (3.66)**
Number of snacks during the day		**−78.52 (16.11)**	**−71.02 (16.94)**	−71.01 (37.37)	−56.88 (28.72)
Energy		**1.57 (0.05)**	**1.54 (0.05)**	**1.45 (0.32)**	**1.36 (0.17)**

Notes: Statistically significant coefficients ($p < .05$) are marked in bold. The covariates with nonsignificant coefficients are baby's sex, caregiver's race, caregiver's ethnicity, caregiver's education level, caregiver's marital status, WIC and SNAP participation status, and age of infant (in days) when the mother stopped breastfeeding. The estimates for all covariates are available in the Supplementary Tables.

3.2.3. Logistic Regression with a Binary Indicator of Meeting Reference Intakes for Added Sugars

For this example, the ratio of added sugars to total energy is recoded as a binary variable depending on whether the ratio is below the DGA recommendation (<10% of total energy). The estimates from the four methods generally agree on which covariates are significant (see Table 8). Though not statistically significant, the non-NCI methods and the two NCI methods show nominal effects in opposite directions for group differences based on the child's sex, caregiver's ethnicity, and parity.

Table 8. Coefficients with significant or opposite signs in the logistic regression, with binary indicator of meeting recommended intake for added sugar.

Variable		1 day	Average of 2 days	Bivariate NCI macros	MCMC
		Point estimate (SE)			
Intercept		**2.93 (0.74)**	**2.98 (0.75)**	**2.87 (0.96)**	**4.14 (1.36)**
Baby's sex	Female	0.07 (0.15)	0.05 (0.16)	−0.12 (0.18)	−0.06 (0.24)
	Male (reference)	—	—	—	—
Caregiver's race	African American	−0.17 (0.15)	−0.22 (0.17)	−0.17 (0.27)	−0.29 (0.31)
	Other	**0.85 (0.25)**	**0.79 (0.25)**	**0.75 (0.36)**	**1.07 (0.44)**
	White (reference)	—	—	—	—

doi:10.20944/preprints202511.0233.v1

Variable		**1 day**	**Average of 2 days**	**Bivariate NCI macros**	**MCMC**
		Point estimate (SE)			
Caregiver's ethnicity	Hispanic or Latino	−0.22 (0.18)	−0.14 (0.18)	0.15 (0.24)	0.14 (0.27)
	Non-Hispanic or non-Latino (reference)	—	—	—	—
Parity	Firstborn	−0.15 (0.20)	−0.15 (0.21)	0.12 (0.23)	0.20 (0.31)
	Second born	−0.05 (0.19)	−0.08 (0.18)	0.24 (0.21)	0.42 (0.29)
	Third or subsequent born (reference)	—	—	—	—
When solid foods were introduced	Before 4 months	−0.39 (0.21)	**−0.41 (0.20)**	−0.38 (0.22)	−0.47 (0.26)
	After 4 months (reference)	—	—	—	—
When salty snacks were introduced	In child's first year	**−1.45 (0.45)**	**−1.52 (0.48)**	**−1.74 (0.63)**	**−2.80 (0.99)**
	In child's second year	−1.32 (0.66)	**−1.39 (0.68)**	**−1.71 (0.82)**	**−2.70 (1.13)**
	Not in child's first 2 years (reference)	—	—	—	—
Age of the infant (in days) when the mother stopped breastfeeding		0.0003 (0.0006)	0.0006 (0.0005)	**0.001 (0.0007)**	0.002 (0.001)
Number of snacks during the day		**−0.32 (0.07)**	**−0.31 (0.07)**	**−0.37 (0.08)**	**−0.48 (0.15)**

Notes: Statistically significant coefficients ($p < .05$) are marked in bold. The covariates with nonsignificant coefficients are caregiver's education level, caregiver's marital status, WIC and SNAP participation status, and mother's age when giving birth. The estimates for all covariates are available in the Supplementary Tables.

3.2.4. Linear Model with HEI Total Score as Outcome

Multivariate regression analysis explored the HEI total score as a function of demographic characteristics and feeding practices. Table 9 presents the regression coefficients for the three methods. The two non-NCI methods show no significant association between HEI total scores and any of the covariates selected, whereas the MCMC method identifies four: baby's sex, TV on while eating, pattern of WIC participation, and age of the infant (in days) when the mother stopped breastfeeding. For comparisons among children with varying WIC participation patterns, all three methods agree that the longer children are enrolled in WIC, the nominally higher their HEI total score. The MCMC method produces the largest difference among all three methods. Moreover, MCMC results suggest that the group difference (coefficient = −3.37 [SE = 1.49]) between children who are on WIC for 2–3 years and those who are consistently on WIC is significant ($p = .024$), whereas such a difference is not significant in the two non-NCI methods.

 doi:10.20944/preprints202511.0233.v1

Table 9. Coefficients with significant or opposite signs in the regression model, with HEI total score as the outcome.

Variable		1 day	Average of 2 days	MCMC
		Point estimate (SE)		
Intercept		**58.01 (11.10)**	**57.23 (10.93)**	**62.12 (2.90)**
Baby's sex	Female	2.06 (3.63)	2.47 (3.67)	**3.25 (1.02)**
	Male (reference)	—	—	—
TV on while eating	Most or sometimes	−2.28 (3.96)	−2.35 (3.81)	**−2.32 (0.99)**
	Never or rarely (reference)	—	—	—
Pattern of WIC participation	1 year or less	−2.49 (6.41)	−2.29 (6.13)	−2.62 (1.63)
	2–3 years	−2.62 (5.28)	−2.35 (5.12)	**−3.37 (1.49)**
	4–5 years	−0.03 (5.94)	0.10 (6.08)	−0.65 (1.62)
	Intermittently	−1.00 (5.14)	−0.76 (5.18)	−2.00 (1.45)
	Consistently (reference)	—	—	—
Age of the infant (in days) when the mother stopped breastfeeding		0.01 (0.02)	0.01 (0.02)	**0.01 (0.00)**

Notes: Statistically significant coefficients (p < .05) are marked in bold. The covariates with nonsignificant coefficients are caregiver's race, caregiver's ethnicity, caregiver's education level, caregiver's marital status, currently using regular childcare, usual number of hours the child sleeps, frequency of family meals together in a week, timing of the introduction of sugar-sweetened beverages, early introduction of complementary foods, timing of the introduction of sugar-sweetened foods, and SNAP participation status. The estimates for all covariates are available in the supplementary tables.

4. Discussion

The application of the MCMC method for UI analyses was motivated by the complexity of dietary patterns comprising various foods and beverages. These should be assessed collectively (Zhang et al., 2011b). Although the estimation of HEI scores was a motivating factor for implementing the MCMC method, the MCMC method can be used for any study in which joint estimation of multiple dietary constituents is of interest (Tooze, 2020). When using data that represent a population, researchers can implement the MCMC method to construct a pseudo population that represents the intake distribution. This pseudo population provides UI estimates from multiple dietary constituents while considering covariates of interest.

This study explored the use of the pseudo population generated from the MCMC method for UI analyses beyond those involving HEI scores. The analyses considered were population and subpopulation intake distribution for a single dietary constituent and regression models with intake estimates as either the outcome or a covariate. To evaluate the performance of the MCMC method, three other methods were also considered: two non-NCI methods that use 1 day of dietary intake data or the average of 2 days of dietary intake data (when available), and the NCI method developed for one or two dietary constituents. Previous studies focusing on dietary intake overwhelmingly rely on data from NHANES. The current study investigated diet quality in children aged 5 using WIC ITFPS-2 dietary recall data.

Comparing the results of the two NCI methods and the two non-NCI methods, the distributions of single dietary constituents or ratios of intake to energy show similar patterns, whereby the two NCI methods usually produce higher estimates for the first quartile and the median, but lower estimates for the third quartile. This finding demonstrates the effectiveness of NCI methods for normalizing the intake distribution. Without UI estimation, the distribution using 1 day of intake data or the average of 2 days of intake data is usually right-skewed, with a larger median than the mean and a long tail at the third quantile and beyond. Such a pattern directly impacts deficiency detection, which relies on the ratio of the percentage of energy from a certain dietary constituent. In the current example, compared with the two NCI methods, the two non-NCI methods consistently report a higher percentage of the population as meeting the DGA recommendation for added sugars, which may lead to an overly optimistic conclusion about the diet quality of young children.

The four estimation methods examined generally agree on the association of intakes with covariates for most of the models assessed in this study. However, in certain situations, one method may report an effect to be significant when the others do not. Nonetheless, the difference in the magnitude is seldom over two SEs. As indicated by the two regression models in the bivariate analyses, the non-NCI and NCI methods report group differences in opposite directions for some effects. However, this occurred only for effects that were not statistically significant. These findings are consistent with those of other studies (Fulgoni et al., 2020; Herrick et al., 2018).

Of particular interest in the current study is the comparison between the two NCI methods. In most cases, the mean and median of intake estimates from the MCMC method are similar to those from the univariate or bivariate NCI methods. However, the study found an exception when examining episodically consumed food, for example, whole grains, because the distribution in population and subpopulation differs by the method used. UI estimation using the bivariate NCI method assumes the probability of consumption is correlated with consumption quantity, whereas the MCMC method does not explicitly incorporate this assumption. Further investigation is needed to identify the cause of the difference. With the incorporation of seafood in the WIC food package and the launch of the HEI-Toddler-2020, understanding the estimation issues around episodically consumed foods is crucial, as very young children in the United States do not regularly consume seafood.

However, the findings from the bivariate analyses suggest that the two NCI methods can reach comparable estimations on ratios between dietary constituents and energy. One possible explanation is that the magnitude of energy is much larger than the intake estimates of the dietary constituents. Given the similar energy distribution, dividing a dietary constituent by energy to develop a ratio reduces the difference between the NCI methods. In other words, the distribution of energy drives the distribution of ratios. Though there are other ways to compute a ratio, such as the population ratio method, they are not considered here because other studies have found that the estimation from these methods is similar to the NCI methods (Herrick et al., 2018; Wambogo et al., 2020).

One advantage of using the MCMC method is that, when working with episodically consumed dietary constituents, it rarely faces convergence failures commonly encountered in the univariate or bivariate NCI macros (Fulgoni et al., 2020; Herrick et al., 2018; Kipnis et al., 2009; Zhang et al., 2011b). For the intake estimation for whole grains, the univariate and bivariate NCI macros failed to converge for five and one replicate weights, respectively. When this occurs, the degrees of freedom are adjusted to the number of successful runs among the replicate weights for SE estimation using the univariate or bivariate NCI methods. Additionally, the computation time is shorter and the file sizes are smaller for the MCMC method relative to the other NCI methods when working on the same number of dietary constituents—another practical reason to choose the MCMC method for intake estimation.

4.1. Limitations

This study used a nationally representative sample of children aged 5 years. Importantly, their dietary intake patterns are substantially different from those of an adolescent or adult population. The findings presented could differ from those obtained by applying the MCMC method to other population groups. The sample used in this study challenges the MCMC method because dietary constituents are consumed daily or episodically. Many of the dietary constituents used to calculate the HEI score (e.g., seafood, legumes, dark vegetables, whole grains) are considered episodically consumed by young children. Large numbers of episodically consumed constituents could cause convergence failure in the application of the MCMC method. All constituents need to be carefully disaggregated to achieve a relatively balanced number of episodic and daily components. It is unclear whether disaggregation strategies could impact the final estimation results.

Although the current study shows that the MCMC method can produce UI estimation comparable to the other NCI methods using data from the WIC ITFPS-2 study, it does not assess how close these estimates are to the true UI. A simulation study is necessary to evaluate the bias and precision of UI estimates obtained from the MCMC method.

5. Conclusions

This study illustrates the use of MCMC for UI analyses that involve one, two, or more dietary constituents. In general, estimation based on the MCMC method is largely consistent with its NCI counterparts for univariate and bivariate analyses. More investigation is required to understand the cause of large differences in the distributions of episodically consumed dietary constituents. The examples used in this study demonstrate the applicability of the MCMC method for not only multivariate analysis but also a range of diet quality and intake evaluation. For practical reasons, including reduced computation time, smaller file sizes, and increased probability of model convergence, the MCMC method may be preferred.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

References

1. Ahuja, J. K. C., Montville, J. B., Omolewa-Tomobi, G., Heendeniya, K. Y., Martin, C. L., Steinfeldt, L. C., Anand, J., Adler, M. E., LaComb, R. P., & Moshfegh, A. J. (2012). *USDA food and nutrient database for dietary studies, 5.0 – documentation and user guide*. U.S. Department of Agriculture, Agricultural Research Service, Food Surveys Research Group.
2. Au, L., Thompson, H., Ritchie, L., Sun, B., Zimmerman, T. P., Whaley, S. E., Reat, A., Sankavaram, K., & Borger, C. (2025). Longer WIC participation is associated with higher diet quality and consumption of WIC-eligible foods among children 2-5 years old. *Journal of the Academy of Nutrition and Dietetics, 125*(10), 1443–1457.e1. https://doi.org/10.1016/j.jand.2024.11.012
3. Borger, C., Zimmerman, T., Vericker, T., DeMatteis, J., Gollapudi, B., Whaley, S., Ritchie, L., Au, L., Sallack, L., & May, L. (2020). *WIC Infant and Toddler Feeding Practices Study-2: Fourth year report*. U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support. https://www.fns.usda.gov/research/wic/itfps-2-fourth-year-report
4. Borger, C., Zimmerman, T., DeMatteis, J., Gollapudi, B., Whaley, S., Ritchie, L., Au, L., & May, L. (2022). *WIC Infant and Toddler Feeding Practices Study-2: Fifth year report*. U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support. https://www.fns.usda.gov/research/wic-itfps-2-fifth-year-report

5. Borger, C., Zimmerman, T. P., DeMatteis, J., Gollapudi, B., Thorn, B., Whaley, S. E., & Ritchie, L. (2025). *WIC Infant and Toddler Feeding Practices Study-2: Year 9 report*. U.S. Department of Agriculture, Food and Nutrition Service. https://www.fns.usda.gov/research/wic/itfps2/ninth-year
6. Brassard, D. (2022). *National Cancer Institute (NCI) methods workshop*. GitHub. https://github.com/didierbrassard/nci_workshop
7. Chi, D. L., & Scott, J. M. (2019). Added sugar and dental caries in children: A scientific update and future steps. *Dental Clinics of North America*, *63*(1), 17–33. https://doi.org/10.1016/j.cden.2018.08.003
8. Dodd, K. W., Guenther, P. M., Freedman, L. S., Subar, A. F., Kipnis, V., Midthune, D. N., Tooze, J. A., & Krebs-Smith, S. M. (2006). Statistical methods for estimating usual intake of nutrients and foods: A review of the theory. *Journal of the American Dietetic Association*, *106*(10), 1640–1650. https://doi.org/doi:%2010.1016/j.jada.2006.07.011
9. Fulgoni, V. L., III, Gaine, P. C., & Scott, M. O. (2020). Comparison of various methods to determine added sugars intake to assess the association of added sugars intake and micronutrient adequacy. *Nutrients*, *12*(9), 2816. https://doi.org/10.3390/nu12092816
10. Gelfand, A. R., & Tangney, C. C. (2021). Dietary quality differs among cannabis use groups: Data from the National Health and Nutrition Examination Survey 2005–16. *Public Health Nutrition*, *24*(11), 3419–3427. https://doi.org/10.1017/S1368980020001846
11. Gillespie, K. M., White, M. J., Kemps, E., Moore, H., Dymond, A., & Bartlett, S. E. (2024). The impact of free and added sugars on cognitive function: A systematic review and meta-analysis. *Nutrients*, *16*(1), 75. https://doi.org/10.3390/nu16010075
12. Harrison, G. G., Hirschman, J. D., Owens, T. A., McNutt, S. W., & Sallack, L. E. (2014). WIC Infant and Toddler Feeding Practices Study: Protocol design and implementation. *The American Journal of Clinical Nutrition*, *99*(3), 742S–746S. https://doi.org/10.3945/ajcn.113.073585
13. Herrick, K. A., Rossen, L. M., Parsons, R., & Dodd, K. W. (2018). Estimating usual dietary intake from National Health and Nutrition Examination Survey data using the National Cancer Institute method. *Vital and Health Statistics*, *2*(178), 1–63.
14. Hutchinson, J. M., Dodd, K. W., Guenther, P. M., Lamarche, B., Haines, J., Wallace, A., Perreault, M., Williams, T. E., Louzada, M. L. D. C., Jessri, M., Lemieux, S., Olstad, D. L., Prowse, R., Simpson, J. R., Vena, J. E., Szajbely, K., & Kirkpatrick, S. I. (2023). The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: Scoring system and construct validity. *Applied Physiology, Nutrition, and Metabolism*, *48*(8), 620–633. https://doi.org/10.1139/apnm-2023-0018
15. Kline, N., Zvavitch, P., Wroblewska K., Worden, M., Mwombela, B., & Thorn, B. (2022). *WIC participant and program characteristics 2020*. U.S. Department of Agriculture, Food and Nutrition Service. https://www.fns.usda.gov/research/wic-participant-program-characteristics-2020
16. Kipnis, V., Midthune, D., Buckman, D. W., Dodd, K. W., Guenther, P. M., Krebs-Smith, S. M., Subar, A. F., Tooze, J. A., Carroll, R. J., & Freedman, L. S. (2009). Modeling data with excess zeros and measurement error: Application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics*, *65*(4), 1003–1110. https://doi.org/10.1111/j.1541-0420.2009.01223.x
17. Kirkpatrick, S. I., & Collins, C. E. (2016). Assessment of nutrient intakes: Introduction to the special issue. *Nutrients*, *8*(4):184. https://doi.org/10.3390/nu8040184
18. Leyvraz, M., Chatelan, A., da Costa, B. R., Taffé, P., Paradis, G., Bovet, P., Bochud, M., & Chiolero, A. (2018). Sodium intake and blood pressure in children and adolescents: A systematic review and meta-analysis of experimental and observational studies. *International Journal of Epidemiology*, *47*(6), 1796–1810. https://doi.org/10.1093/ije/dyy121
19. National Cancer Institute. (n.d.). *User's guide for analysis of usual intakes: For use with versions 2.1 of the MIXTRAN, DISTRIB, and INDIVINT SAS macros*. https://epi.grants.cancer.gov/diet/usualintakes/Users_Guide_v2.1.pdf
20. National Cancer Institute. (2013). *SAS macros for fitting multivariate measurement error models and estimating multivariate usual intake distributions*.

https://epi.grants.cancer.gov/diet/usualintakes/users_guide_multivariate_mcmc_distrib_brr_macros_v1.0.pdf
21. National Cancer Institute. (2024). *Usual dietary intakes: SAS macros for the NCI method*. https://epi.grants.cancer.gov/diet/usualintakes/macros.html
22. National Institutes of Health, National Cancer Institute, Division of Cancer Control & Population Sciences. (2025, 14 March). *Overview & background of the Healthy Eating Index*. https://epi.grants.cancer.gov/hei
23. Nusser, S. M., Fuller, W. A., & Guenther, P. M. (1995). Estimating usual dietary intake distributions: Adjusting for measurement error and nonnormality in 24-hour food intake data. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin D (Eds.), *Survey measurement and process quality* (pp. 689–709). John Wiley & Sons.
24. O'Neil, C. E., Nicklas, T. A., Zanovec, M., Cho, S. S., & Kleinman, R. (2011). Consumption of whole grains is associated with improved diet quality and nutrient intake in children and adolescents: The National Health and Nutrition Examination Survey 1999–2004. *Public Health Nutrition*, *14*(2), 347–355. https://doi.org/doi:10.1017/S1368980010002466
25. Raper, N., Perloff, B., Ingwersen, L., Steinfeldt, L., & Anand, J. (2004). An overview of USDA's Dietary Intake Data System. *Journal of Food Composition and Analysis*, *17*(3-4), 545–555. https://doi.org/10.1016/j.jfca.2004.02.013
26. Siegfried, Y., DeMatteis, J. M., & Gollapudi, B. (2023). Implementation of a sample design for a survey of program participants using time-location sampling. *PLOS One*, *18*(5), e0285877. https://doi.org/10.1371/journal.pone.0285877
27. Tooze, J. A. (2020). Estimating usual intakes from dietary surveys: Methodologic challenges, analysis approaches, and recommendations for low- and middle-income countries. Intake – Center for Dietary Assessment/FHI Solutions.
28. Tooze, J. A., Kipnis, V., Buckman, D. W., Carroll, R. J., Freedman, L. S., Guenther, P. M., Krebs-Smith, S. M., Subar, A. F., & Dodd, K. W. (2010). A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method. *Statistics in Medicine*, *29*(27), 2857–2868. https://doi.org/10.1002/sim.4063
29. Tooze, J. A., Midthune, D., Dodd, K. W., Freedman, L. S., Krebs-Smith, S. M., Subar, A. F., Guenther, P. M., Carroll, R. J., & Kipnis, V. (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of the American Dietetic Association*, *106*(10), 1575–1587. https://doi.org/10.1016/j.jada.2006.07.003
30. U.S. Department of Agriculture, U.S. Department of Health & Human Services. (2020). *Dietary guidelines for Americans, 2020-2025* (9th ed.). www.dietaryguidelines.gov
31. U.S. Department of Agriculture, Food and Nutrition Service. (2022, August 2). *WIC benefits and services*. https://www.fns.usda.gov/wic/benefits-services
32. Wambogo, E., Reedy, J., Herrick, K., Shams-White, M., & Lerman, J. (2020). Comparing Healthy Eating Index scoring methods and different dietary assessment instruments: The interactive diet and activity tracking in AARP (IDATA) Study. *Current Developments in Nutrition*, *4*(Suppl 2), 573. https://doi.org/10.1093/cdn/nzaa046_073
33. Zhang, S., Krebs-Smith, S. M., Midthune, D., Perez, A., Buckman, D. W., Kipnis, V., Freedman, L. S., Dodd, K. W., & Carroll, R. J. (2011a). Fitting a bivariate measurement error model for episodically consumed dietary components. *International Journal of Biostatistics*, *7*(1). https://doi.org/10.2202/1557-4679.1267
34. Zhang, S., Zhang, S., Carroll, R. J., Midthune, D., Guenther, P. M., Krebs-Smith, S. M., Kipnis, V., Dodd, K. W., Buckman, D. W., Tooze, J. A., & Freedman, L. (2011b). A new multivariate measurement error model with zero-inflated dietary data, and its application to dietary assessment. *Annals of Applied Statistics*, *5*(2B):1456–1487. https://doi.org/10.1214/10-AOAS446