Review

Statistical Methods for Estimating Usual Intake of Nutrients and Foods: A Review of the Theory

KEVIN W. DODD, PhD; PATRICIA M. GUENTHER, PhD, RD; LAURENCE S. FREEDMAN, PhD; AMY F. SUBAR, PhD, MPH, RD; VICTOR KIPNIS, PhD; DOUGLAS MIDTHUNE, MS; JANET A. TOOZE, PhD; SUSAN M. KREBS-SMITH, PhD, MPH, RD

ABSTRACT

Although 24-hour recalls are frequently used in dietary assessment, intake on a single day is a poor estimator of long-term usual intake. Statistical modeling mitigates this limitation more effectively than averaging multiple 24-hour recalls per respondent. In this article, we describe the statistical theory that underlies the four major modeling methods developed to date, then review the strengths and limitations of each method. We focus on the problem of estimating the distribution of usual intake for a population from 24-hour recall data, giving special attention to the problems inherent in modeling usual intake for foods or food groups that a proportion of the population does not consume every day (ie, episodically consumed foods). All four statistical methods share a common framework. Differences between the methods arise from different assumptions about the measurement characteristics of 24-hour recalls and from the fact that more recently developed methods build upon their predecessor(s). These differences can result in estimated usual intake distributions that differ from one another. We also demonstrate the need for an improved method for estimating usual intake distributions for episodically consumed foods.

J Am Diet Assoc. 2006;106:1640-1650.

K. W. Dodd, V. Kipnis, and D. Midthune are mathematical statisticians and A. F. Subar and S. M. Krebs-Smith are nutritionists, National Cancer Institute, Bethesda, MD. P. M. Guenther is a nutritionist, US Department of Agriculture Center for Nutrition Policy and Promotion, Alexandria, VA. L. S. Freedman is director, Biostatistics Unit, Gertner Institute for Epidemiology and Health Policy Research, Sheba Medical Center, Tel Hashomer, Israel. J. A. Tooze is an assistant professor, Section on Biostatistics, Department of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC.

Address correspondence to: Kevin W. Dodd, PhD, Mathematical Statistician, National Cancer Institute, 6116 Executive Blvd, MSC 8317, Bethesda, MD 20892. E-mail: doddk@mail.nih.gov

Copyright © 2006 by the American Dietetic Association.

0002-8223/06/10610-0012$32.00/0
doi: 10.1016/j.jada.2006.07.011

A common purpose of dietary assessment is to evaluate the dietary intake of a group or population in relation to some standard, with respect to both nutrient adequacy and the prevention of chronic disease. Standards for nutrient adequacy have changed in recent years, moving beyond the Recommended Dietary Allowances to include several types of Dietary Reference Intakes (1). Recommendations for food intake include those found in the *Dietary Guidelines for Americans 2005* (2) and in *Tracking Healthy People 2010*, a statement of national health goals and objectives (3).

For simplicity in consumer communication, food intake recommendations intended to achieve nutrient adequacy and promote health are often expressed in terms of daily targets (4). However, because nutrients can be stored in the body and because dietary intake varies from day to day, it is both unnecessary and impractical to achieve those targets every day (2,3). Therefore, a key concept in assessing adherence to such recommendations is usual intake, which is defined as long-run average intake (1).

Food frequency questionnaires (FFQs) and 24-hour recalls are two of the major dietary data collection instruments. The 24-hour recall has been the primary instrument used in surveillance, and the FFQ has been the primary instrument used in epidemiology. FFQs are designed to measure long-term behavior and are relatively inexpensive to field compared to 24-hour recalls. However, FFQs are limited to a finite list of foods and are hampered by the inability of individuals to accurately report their food intake retrospectively over a long period of time. Both of these shortcomings introduce substantial error into usual intake estimates based on FFQs (5-8).

In contrast, 24-hour recalls provide rich detail about the types and amounts of foods consumed. By focusing on a single day, the magnitude of systematic errors on 24-hour recalls is reduced. However, individual diets can vary greatly from day to day. In addition, measurement errors plague 24-hour recalls and are compounded by error resulting from the use of standardized recipe files and food composition databases. All of these factors contribute to considerable within-person variability, ensuring that measured intake on a single day is a poor estimator of long-term intake (9,10). Early attempts to compensate for this limitation by averaging multiple (two to seven) 24-hour recalls per respondent were deemed unsatisfactory due to high respondent burden and low quality of reported information. Moreover, averages over a small number of days do not adequately represent individual usual intakes. Thus, more sophisticated methods based on statistical modeling evolved (11).

© 2006 by the American Dietetic Association

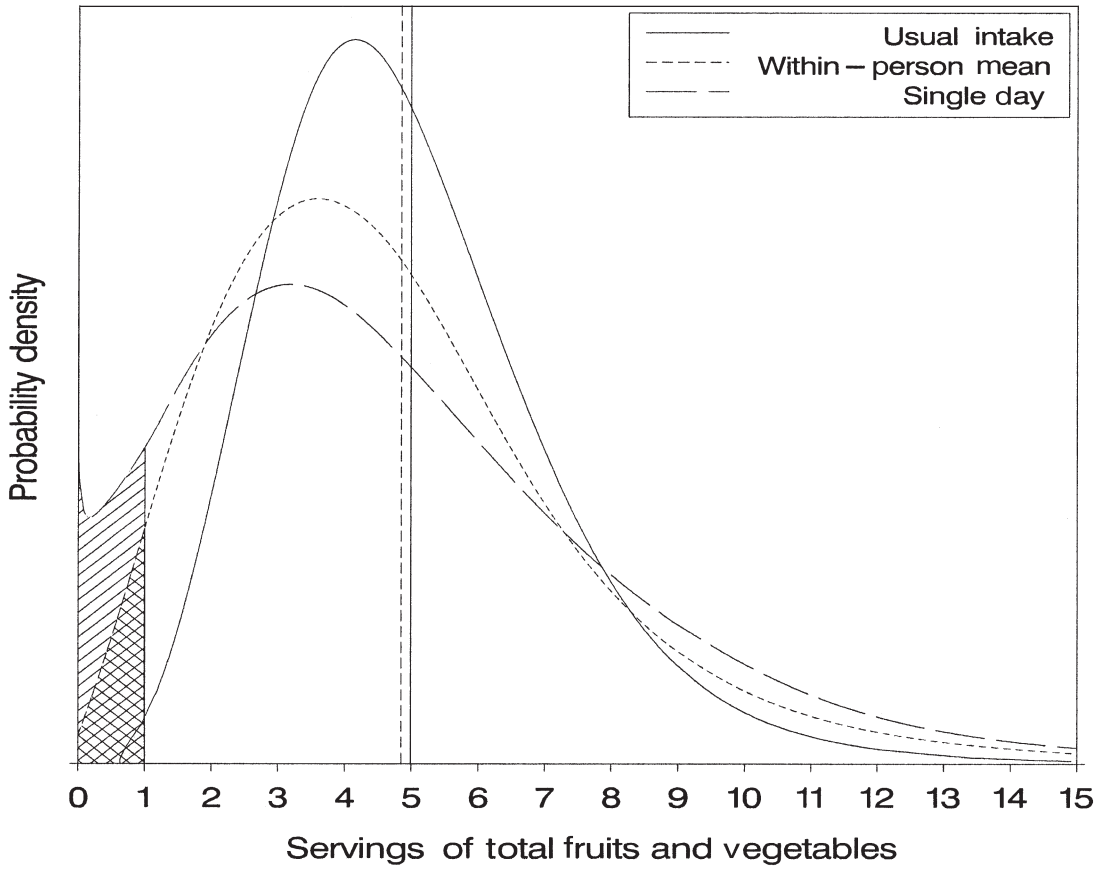

Figure 1. Comparison of estimated distributions for intake of total fruits and vegetables in the US population for the period 1994-1996, based on one 24-hour recall per respondent (broken line), on within-person means of two 24-hour recalls (dashed line), and on two 24-hour recalls per respondent, using a statistical model. The dashed vertical line marks the mean of the distribution of within-person means. The solid vertical line marks the mean of the other two distributions. The percent of the population with intake <1 serving is estimated by the crosshatched/shaded areas under each curve to the left of 1. Adapted with permission from reference 16.

Our research focused on the problem of using 24-hour recall data to estimate the distribution of usual intake in a population. We reviewed the four major statistical modeling methods developed to date. All four methods are based on a common framework, but differ in assumptions made about the measurement characteristics of 24-hour recalls and in statistical complexity.

We begin by explaining why simple approaches to estimating usual intake are unsatisfactory. Next, we describe the common framework and provide a rudimentary method that illustrates it. We then describe how each of the four major methods builds on the common framework and highlight each method's strengths and weaknesses. We give special attention to the fourth method, which addresses problems inherent in modeling usual intake of foods or food groups that are not typically consumed every day. In the Discussion, we demonstrate the need for an improved method that will allow estimation of usual intake distributions for these episodically consumed foods. This article serves as background for two companion articles (12,13).

REVIEW OF STATISTICAL METHODOLOGY

Why Simple Approaches to Estimating Usual Intake Are Unsatisfactory

When researchers in the nutrition community recognized that a single day's reported intake poorly reflected usual intake (14,15), their first solution was to measure several single-day intakes for each respondent with 24-hour recalls and average the observations. The empirical distribution of these within-person means was used to estimate the distribution of usual intake for a population. However, for many dietary components of interest, the mean of any financially and operationally feasible number of 24-hour recalls for an individual still contains considerable within-person variation. Thus, the distribution of within-person means has a larger variance than the true usual intake distribution, leading to a biased estimate of the fraction of the population with usual intake above or below some standard. The degree of bias decreases for standards that are closer to the population mean intake. These potential biases are illustrated in Figure 1, taken from Guenther and colleagues (16). Figure 1 compares

estimated distributions of intake of total fruits and vegetables in the US population for the period 1994-1996, based on one 24-hour recall per respondent (broken line), on the within-person means of two 24-hour recalls (dashed line), and on two 24-hour recalls using a statistical model. The dashed vertical line marks the mean of the distribution of within-person means. The solid vertical line marks the mean of the other two distributions.

As expected, the area under each curve to the left of 5—estimating the percent of the population with a usual intake of fewer than five servings per day—is approximately 40% for all three distributions. However evaluation of dietary adequacy often involves standards that fall in the tails of the intake distribution (17), where the biases can be substantial. For example, the percent of the population with a usual intake of less than one serving per day is estimated to be 9.3%, 3.6%, and 0.4%, for the 1-day distribution, the within-person means distribution, and the usual intake distribution, respectively (16).

The Common Framework of Statistical Models Based on 24-Hour Recalls

Statistical modeling mitigates some of the limitations of 24-hour recalls by analytically estimating and removing the effects of within-person variation in dietary intake. Each method described in the following sections performs the same sequence of steps: Step 1: Describe the assumed relationship between individual 24-hour recall measurements and individual usual intake; Step 2: Partition the total variation in 24-hour recall measurements into within- and between-person components; and Step 3: Estimate the usual intake distribution accounting for within-person variation.

Different assumptions about the measurement characteristics of a 24-hour recall lead to differences in the methods, primarily in Steps 1 and 3. The more complex methods include Step 0, in which initial adjustments are made to observed 24-hour recalls. In what follows, we illustrate the statistical motivation for each step as we develop a rudimentary method for estimating the usual intake distribution.

Step 1: Describe the Assumed Relationship between Individual 24-Hour Recall Measurements and Individual Usual Intake

The usual assumption is that 24-hour recall intake is an unbiased estimator of usual intake. Lack of bias does not imply lack of error—it means that a particular 24-hour recall may over- or underestimate an individual's true usual intake, but over repeat applications to the same individual, the estimation errors cancel out. This assumption is equivalent to the assumption that a 24-hour recall is unbiased for true single-day intake.

Step 2: Partition the Total Variation in 24-Hour Recall Measurements into Within- and Between-Person Components

Individual usual intakes may be expressed as the sum of the group's mean usual intake and person-specific deviations from the group mean; these deviations (in parentheses below) represent between-person variation:

$$\text{[I]}\quad \text{individual usual intake} = \text{group mean usual intake} + (\text{individual usual intake} - \text{group mean usual intake}).$$

Each 24-hour recall may be expressed as:

$$\text{[II]}\quad \text{24-hour recall intake} = \text{group mean usual intake} + (\text{individual usual intake} - \text{group mean usual intake}) + (\text{24-hour recall intake} - \text{individual usual intake}),$$

where the third term represents within-person variation. Parameters of the model [II] are estimated using standard methods. Estimation of the within-person variance requires replicated measurements, so at least some individuals must provide two or more 24-hour recalls.

Step 3: Estimate the Usual Intake Distribution Accounting for Within-Person Variation

If the within-person variance is $\sigma_w{}^2$, then the variance of the average of n independent 24-hour recall intakes for an individual is $\sigma_w{}^2/n$. If the between-individual variance (ie, the variance of the usual intake distribution) is $\sigma_b{}^2$, it follows that the empirical distribution of within-person means has variance $\sigma_b{}^2 + \sigma_w{}^2/n$. A set of intermediary values with the desired variance $\sigma_b{}^2$ is constructed by shrinking each individual mean toward the overall mean:

$$\text{[III]}\quad \text{intermediary value} = (1 - w) \times (\text{overall mean}) + w \times (\text{individual mean}),$$

where w, the shrinkage factor, is the square root of the ratio of the between-person variance to the variance of the within-person means distribution:

$$\text{[IV]}\qquad w = \sqrt{\frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2/n}}.$$

By construction, the empirical distribution of the intermediary values has a mean equal to the overall 24-hour recall mean. The percentage of the group with intake less than some threshold value, such as a dietary standard, is estimated by the percentage of the intermediary values that are less than the threshold value. Although the intermediary values [III] are based on individual means, they are not suitable estimates of individual usual intake; their purpose is solely to describe the distribution of usual intake. Alternative methodologies have been developed to produce appropriate estimates of individual usual intake.

Equations [III] and [IV] show that when the between-person variation is small or when n is small, w is close to zero, and each intermediary value is close to the overall mean. When the within-person variation is small or when n is large, w is close to one, and the intermediary values are close to the individual means.

The Consequences of Routine Use of Transformations

The rudimentary method described above requires the entire set of intermediary values to describe the esti-

mated usual intake distribution. In some cases, the description can be simplified. For example, if the within- and between-person deviations are normally distributed, then so is the usual intake distribution, which is then described solely by its mean and variance.

Such cases are rare, in part because intake (usual or on a single day) is by definition non-negative. Some individuals can have usual intakes more than twice the group usual intake, but no individuals can have usual intake equally far below the group usual intake. Similarly, the magnitude of deviations from intakes that exceed an individual's usual intake may be larger than those deviations from intakes that fall short. Thus, observed intake distributions tend to be right-skewed (having a small number of very large values) instead of exhibiting a normal distribution's symmetry about its mean.

To reconcile the desire to use the statistical properties of the normal distribution with the need to model inherently non-normal data, statisticians often assume that a normal distribution approximates the distribution of a (nonlinear) transformation of the observed data, rather than the observed data themselves. For example, if the data have a highly skewed distribution, then the distribution obtained by taking the logarithm of each observation may be symmetric, and therefore be better approximated by a normal distribution. In this example, we say that the data have been transformed to the log scale. For less-skewed data, weaker transformations, such as the square root and cube root, are often sufficient to achieve approximate normality. If a particular transformation produces normally distributed data, the distribution of untransformed data can be described in terms of the normal distribution and the transformation. This fact is crucial when estimating usual intake distributions because standards by which intake are to be assessed are expressed in units on the scale of untransformed intake; for example, in grams rather than in the square root of grams.

The expression that relates values in the transformed scale to usual intake in the original scale is called the back-transformation. The form of the back-transformation depends on the assumptions about the 24-hour recall as an assessment instrument; one must assume that a 24-hour recall measurement is unbiased for usual intake in a particular scale. The methods we discuss assume either that transformed 24-hour recall intake is an unbiased estimator of transformed usual intake (Assumption A), or that 24-hour recall intake is an unbiased estimator of usual intake in the untransformed scale (Assumption B).

Assumption A leads to a simple back-transformation that is just the inverse of the original transformation. Assumption B follows directly from the definition of usual intake if within-person variation in 24-hour recalls is solely due to day-to-day variability in diet, but in general requires that all components of within-person variation tend to average out on the original scale. As described later, the back-transformation under Assumption B requires an additional adjustment. Reasons to favor one assumption over the other are given in the Discussion.

A 1986 report by the National Research Council (18) was groundbreaking because it was the first to suggest applying model [II] to 24-hour recalls to estimate distributions of usual intake. The report also described how transformations could be incorporated (under Assumption A) into usual intake estimation. Although it did not provide an in-depth methodological description, the report served as the launching point for the four methods described in the following sections and summarized in Figure 2.

The Institute of Medicine Method: Building on the Common Framework

Requiring the full set of intermediary values to describe the usual intake distribution, rather than relying on a simplified parameterization, allows the rudimentary approach suggested by equations [II] through [IV] to withstand mild departures from normality. The desire to retain this data-driven robustness, even when the empirical distribution of single-day intake is highly skewed, motivated the Institute of Medicine (19) to develop a detailed method that includes a power or log transformation as an initial adjustment to the 24-hour recall data. The estimation of between- and within-person components of variance is carried out on the transformed scale, and intermediary values are constructed by shrinking individual means of transformed 24-hour recall intakes to the overall mean of transformed 24-hour recall intake, then applying the inverse of the original transformation to each shrunken mean. As in the National Research Council report, the Institute of Medicine back-transformation is consistent with Assumption A.

The Iowa State University Method: An Extension with a Different Assumption

Nusser and colleagues (20) and Guenther and colleagues (21) described a method developed at Iowa State University for modeling usual intake. In contrast to the National Research Council/Institute of Medicine methods, the Iowa State University method is based on Assumption B; that is, a 24-hour recall is unbiased for usual intake on the original scale rather than on the transformed scale. Whereas the Institute of Medicine method is limited to situations where 24-hour recalls are obtained from a simple random sample of individuals, the Iowa State University method can also be applied to 24-hour recall data from complex surveys.

The Iowa State University method is based on a complex model that uses a two-stage transformation to obtain 24-hour recalls that are almost exactly normally distributed. Choosing the transformation requires a fairly large sample (in our experience, several hundred individuals, of which at least 50 must have at least two 24-hour recalls). The Iowa State University method also allows the within-person variance to vary across individuals, reflecting the fact that some individuals may have a more varied diet than others. The method sacrifices some robustness for this added flexibility—intermediary values are based on theoretical quantiles from a normal distribution instead of on individual means.

Under Assumption B, simply applying the inverse of the initial transformation to intermediary values as in the Institute of Medicine method produces a biased estimate of the usual intake distribution. The back-transfor-

Method			
NRC[a]/IOM[b]	ISU[c]	BP[d]	ISUF[e]
Step 0: Apply initial data adjustments			
i. Apply a power or log transformation to observed 24-hour recalls, so that the distribution of transformed data is approximately normal	i. Adjust observed 24-hour recalls for non–person-specific biases such as season, day-of-week, and time-in-sample effects ii. Construct a two-stage transformation that makes the distribution of the transformed adjusted 24-hour recalls look as close to normal as possible	i. Adjust observed 24-hour recalls for non–person-specific biases such as season, day-of-week, and time-in-sample effects ii. Apply a power or log transformation to observed 24-hour recalls, so that the distribution of transformed data is approximately normal	i. Estimate the distribution of the probability to consume on a given day, based on the relative frequency of nonzero 24-hour recalls ii. Set aside the zero 24-hour recalls and adjust the consumption-day 24-hour recalls for non–person-specific biases such as season, day-of-week, and time-in-sample effects iii. Construct a two-stage transformation that makes the distribution of the transformed adjusted consumption-day 24-hour recalls look as close to normal as possible
Step 1: Describe the assumed relationship between individual 24-hour recall measurements and individual usual intake			
i. A transformed 24-hour recall is unbiased for transformed usual intake (Assumption A)	i. A 24-hour recall is unbiased for usual intake in the untransformed scale (Assumption B)	i. A 24-hour recall is unbiased for usual intake in the untransformed scale (Assumption B)	i. Usual intake is the probability to consume on a given day multiplied by the usual intake amount for a day the food is consumed ii. A 24-hour recall measures zero consumption exactly iii. A nonzero 24-hour recall is unbiased for usual consumption-day intake in the untransformed scale (Assumption B)
Step 2: Partition the total variation in 24-hour recall measurements into within- and between-person components			
i. Within-person variance is the same for all individuals	i. Within-person variance can vary among individuals	i. Within-person variance is the same for all individuals	i. Within-person variance can vary among individuals
Step 3: Estimate the usual intake distribution accounting for within-person variation			
i. Construct a set of intermediary values that retain mean and between-person variance of transformed 24-hour recalls ii. **BACK-TRANSFORMATION:** Apply inverse of initial power or log transformation to each intermediary value iii. The empirical distribution of back-transformed intermediary values is the estimated usual intake distribution	i. Construct a set of intermediary values that retain mean and average between-person variance of transformed 24-hour recalls ii. **BACK-TRANSFORMATION:** Using the inverse of the initial two-stage normality transformation in conjunction with a bias-adjustment, adjust each normal-scale intermediary value to obtain its original-scale counterpart iii. The empirical distribution of original-scale intermediary values is the estimated usual intake distribution	i. Construct a set of intermediary values that retain mean and average between-person variance of transformed 24-hour recalls ii. **BACK-TRANSFORMATION:** Using the inverse of the initial power or log transformation in conjunction with a bias-adjustment, adjust each normal-scale intermediary value to obtain its original-scale counterpart iii. The empirical distribution of original-scale intermediary values is the estimated usual intake distribution	i. **BACK-TRANSFORMATION:** Using the inverse of the initial two-stage normality transformation in conjunction with a bias-adjustment, mathematically describe the distribution of original-scale consumption-day usual intake ii. Mathematically combine the estimated consumption-day usual intake distribution with the estimated distribution of consumption probability to obtain a set of intermediary values that represent usual intake, assuming that consumption probability and consumption-day usual intake are statistically independent iii. The empirical distribution of original-scale intermediary values is the estimated usual intake distribution

[a]NRC=National Research Council.
[b]IOM=Institute of Medicine.
[c]ISU=Iowa State University.
[d]BP=Best-Power.
[e]ISUF=Iowa State University foods.

Figure 2. Steps required for estimating distributions of usual intake using different statistical modeling methods. (Data from this figure are available online at www.adajournal.org as part of a PowerPoint presentation featuring additional online-only content.)

mation of the Iowa State University method incorporates an adjustment for this bias. The Iowa State University method can also account for biases due to seasonality, day-of-week (22), and time-in-sample (23) effects as part of its initial adjustments. Here, time-in-sample effects include the often-seen phenomenon in which the average of the measurements for a first 24-hour recall is higher than the average of measurements for subsequent applications of the instrument in the same group of people. The adjustment is illustrated in Figure 1, where the vertical line depicting the mean of the usual intake distribution (estimated by the Iowa State University method) overlays the line for the mean of the distribution of the first 24-hour recall and is to the right of the line depicting the mean of the distribution of within-person means.

The Best-Power Method: A Simplification of the Iowa State University Method

Extrapolating from the original National Research Council report, Nusser and colleagues (20) also proposed a simplified alternative to the Iowa State University method. This so-called Best-Power method shares the Iowa State University method's applicability to complex surveys, but uses only a one-stage power or log transformation. Moreover, it does not allow the within-person variance to vary across individuals. The simple form of the Best-Power method's initial transformation leads to a straightforward adjustment for transformation-induced bias. This adjustment is described in Figure 3. A simulation study comparing the Iowa State University and Best-Power methods indicated that although the Iowa State University method is (statistically) uniformly superior to the Best-Power method, the differences are very small in practical terms (20).

The Iowa State University Food Method: Application to Episodically Consumed Dietary Components

The National Research Council/Institute of Medicine, Iowa State University, and Best-Power methods were developed to model usual intake where the distributions of single-day intakes can be transformed (at least approximately) to normality. This is the case for intake of most nutrients and for some commonly consumed food groups such as total fruit and vegetables (16). However, for episodically consumed foods, food groups, and nutrients (eg, broccoli, whole grains, and lycopene), it is possible to observe zero consumption on a particular day, even for individuals who sometimes consume these foods. This leads to distributions of observed intake with a clump of zero observations in the left tail. Because the normal distribution is continuous, normality is not attainable by transformation (which would preserve the clump).

Nusser and colleagues (24) proposed a method for modeling usual food intake that explicitly accounts for the clumping at zero. Zero observations are treated separately from positive observations. This separation is motivated by writing the simple n-day within-person mean as the product of two parts:

$$[\mathrm{V}] \quad \frac{\text{total intake of the food}}{n} = \frac{k}{n} \times \frac{\text{total intake of the food}}{k},$$

where k is the number of days on which the food is observed to be consumed (consumption days). For large values of n, equation [V] expresses long-term average intake as the product of two components: the estimated probability of consuming on a given day (k/n) multiplied by the long-term average amount consumed on consumption days (total intake of the food/k). For example, a person who consumes, on average, 4 oz whole grains every other day has a usual intake of 2 oz/day.

The first step in the Iowa State University Foods method is to estimate the distribution of (single-day) consumption probabilities in the population. For simplicity, this distribution is modeled discretely, with possible values: 0, 0.02, . . . , 0.98, 1.00. The estimated proportion of individuals in the population having each value is calibrated to observed counts of individuals who consumed on 0, 1, . . . , n out of n possible 24-hour recalls. When n is small, additional constraints are enforced to ensure a unique estimate of the consumption probability distribution for a given set of observed counts.

Next, the distribution of usual consumption-day intake is estimated by applying the Iowa State University method (20) to the nonzero 24-hour recalls, including any adjustments for time-related biases, such as seasonality, day-of-week, or time-in-sample effects. Thus, the Iowa State University Foods method assumes that Assumption B holds for each nonzero 24-hour recall; that is, that each 24-hour recall is unbiased for usual consumption-day intake in the original scale. In addition, by setting aside all of the zero 24-hour recalls, the Iowa State University Foods method makes the implicit assumption that a 24-hour recall measures zero intake exactly—if a food is not reported on a 24-hour recall, the food was not consumed on that day.

Finally, the Iowa State University Foods method combines the distributions of consumption probability and usual consumption-day intake to obtain the estimated distribution of usual intake. The combining process relies on the assumption that usual consumption-day intake is unrelated to consumption probability. However, Dodd and colleagues (25) demonstrated that the amount consumed on consumption days can be positively correlated with the probability to consume. In such cases, the Iowa State University Foods method may introduce bias by overestimating the amount consumed by those with a low probability of consumption, and underestimating the amount consumed by those with a high probability of consumption.

DISCUSSION

The mean usual intake for a group may be monitored by tracking the average 24-hour recall intake from appropriate surveys over time (26,27). However, evaluating dietary adequacy in relation to recommended standards involves the entire distribution of usual intake. Several methods (within-person means, National Research Coun-

Let T_{ij} denote the true intake for individual i on day j. By definition, the individual's usual intake, denoted by τ_i, is the mean of infinitely many such single-day intakes, ie:

$$\tau_i = \mathrm{E}[T_{ij}|i].$$

Denote the 24-hour recall–reported intake for individual i on day j as R_{ij}, where in general R_{ij} may measure T_{ij} with error ε_{ij}:

$$R_{ij} = T_{ij} + \varepsilon_{ij}, \text{ where } \mathrm{E}[\varepsilon_{ij}|i] = \mu_\varepsilon.$$

Let $g(\cdot)$ be a one-to-one transformation—with inverse transformation $h(\cdot) = g^{-1}(\cdot)$. Denote the transformed 24-hour recall–reported intake for individual i on day j as

$$r_{ij} = g(R_{ij}).$$

The transformation $g(\cdot)$ is such that the following normal-theory model applies:

$$r_{ij} = \mathrm{E}[r_{ij}|i] + \{r_{ij} - \mathrm{E}[r_{ij}|i]\} = b_i + w_{ij},$$

where $\{w_{ij}\} \sim N(0, \sigma_w^2)$ and $\{b_i\} \sim N(\mu, \sigma_b^2)$. Under Assumption A, $b_i = g(\tau_i)$, and therefore $\tau_i = h(b_i)$. Estimates of the unknown parameters μ, σ_b^2, and σ_w^2 are obtained from standard components-of-variance formulas applied to the transformed 24-hour recall–reported intakes.

Now suppose that a 24-hour recall is unbiased for true single-day intake, ie, $\mu_\varepsilon = 0$. Then

$$\tau_i = \mathrm{E}[T_{ij}|i] = \mathrm{E}[R_{ij}|i],$$

and now Assumption B holds: a 24-hour recall is unbiased for true usual intake on the original scale. Then

$$\tau_i = \mathrm{E}[R_{ij}|i] = \mathrm{E}[h(r_{ij})|i] = \mathrm{E}[h(b+w)|b = b_i],$$

where w symbolizes within-person variation in transformed 24-hour recall–reported intake. If h is nonlinear, τ_i does not, in general, equal $h(b_i)$. Using an approximation to the expectation of a function of a random variable (29),

$$\mathrm{E}[h(b+w)|b = b_i] \approx h(\mathrm{E}[b+w|b = b_i]) + \tfrac{1}{2}h''(\mathrm{E}[b+w|b = b_i])\mathrm{Var}[b+w|b = b_i] = h(b_i) + \tfrac{1}{2}h''(b_i)\mathrm{Var}[w] = h(b_i) + \tfrac{1}{2}h''(b_i)\sigma_w^2,$$

where $h''(b_i)$ denotes the second derivative of the function h evaluated at the value b_i.

Nusser and colleagues' Best-Power method (20) constructs a set of intermediary values $\{b_i^*\}$ that has the same mean and variance (and therefore the same normal distribution) as the $\{b_i\}$. Therefore, the back-transformed intermediary values

$$\tau_i^* = h(b_i^*) + \tfrac{1}{2}h''(b_i^*)\sigma_w^2, \qquad \text{[Ia]}$$

have the property that the distribution of $\{\tau_i^*\}$ is approximately the same as that of $\{\tau_i\}$. The second term on the right-hand side of Equation [Ia] is the bias-adjustment term. The chart at the end of this text presents the bias-adjustment terms corresponding to some transformations that are routinely used to achieve normally distributed 24-hour recall data. For example, if the square root of observed 24-hour recall data is approximately normal (ie, $g(R) = R^{1/2}$), the corresponding bias adjustment is obtained by setting $\lambda = 2$ in the "Power" row of the chart. Power or Box-Cox transformations using values of λ greater than 1 are generally required to correct for the right-skewness of observed 24-hour recall distributions; in such cases, the bias correction term is always positive, although it is different for each intermediary value.

Transformation	$g(R)$	$h(r)$	**Bias-adjustment term**
Logarithm	$r = \log(R)$	$R = \exp(r)$	$\tfrac{1}{2}\exp(b_i^*)\sigma_w^2$
Power	$r = R^{1/\lambda}$	$R = r^\lambda$	$\tfrac{1}{2}\lambda(\lambda-1)(b_i^*)^{\lambda-2}\sigma_w^2$
Box-Cox	$r = \lambda(R^{1/\lambda} - 1)$	$R = (r/\lambda + 1)^\lambda$	$\tfrac{1}{2}[(\lambda-1)/\lambda][b_i^*/\lambda + 1]^{\lambda-2}\sigma_w^2$

Bias-adjustment terms used in Nusser and colleagues' Best-Power method (20) for selected choices of normality transformation $g(R)$, with inverse function $h(r)$.

Figure 3. Adjusting for transformation-induced bias in the Best-Power method for estimating usual intake distributions.

cil/Institute of Medicine, Iowa State University, and Best-Power) may be used to estimate this distribution for dietary components consumed nearly every day by almost everyone, but only two methods (within-person means and Iowa State University Foods method) may be used for episodically consumed foods. The distribution of within-person means is generally a poor estimator of the true usual intake distribution because it produces inflated estimates of the prevalence of inadequate or excessive usual intake. The modeling methods (National Research Council/Institute of Medicine, Iowa State University, Best-Power, and Iowa State University Foods method) can—if assumptions are met—produce better estimates, although it may be difficult to decide which method to use. Figure 4 summarizes the strengths and weaknesses of each of the modeling methods. Some methods are im-

	Method			
	NRC[a]/IOM[b]	ISU[c]	BP[d]	ISUF[e]
Strengths	• Permits the use of simple transformations to obtain approximate normality • Robust to mild departures from normality • Easy to apply to real data—SAS[f] code to implement the IOM method has been published in nutrition literature	• Can be applied to datasets from complex surveys • Software (SIDE, C-SIDE, PC-SIDE[g]) is available to implement the ISU method • Software produces standard errors for estimated parameters of usual intake distributions • Can apply preliminary data adjustments for covariates such as interview sequence or day-of-week	• Permits the use of simple transformations to obtain approximate normality • Robust to mild departures from normality • Can be applied to datasets from complex surveys • Software (SIDE) is available to implement the BP method • Software produces standard errors for estimated parameters of usual intake distributions • Can apply preliminary data adjustments for covariates such as interview sequence or day-of-week	• Uses a two-part model that can be applied to datasets with many observed zero intakes • Can be applied to datasets from complex surveys • Software (C-SIDE) is available to implement the ISUF method • Can apply preliminary data adjustments for covariates such as interview sequence or day-of-week in one of the two parts of the model
Weaknesses	• Implementation for datasets from complex surveys (as in original NRC report) is less straightforward, not well documented • Published SAS code does not produce standard errors for estimated parameters of usual intake distributions • Cannot adjust for the effects of covariates such as interview sequence or day-of-week • Cannot be applied to datasets with many observed zero intakes	• Adjustment for covariates is not done in the context of a unified model • Complex two-stage transformations may fail to obtain exact normality, but the subsequent steps of the method rely heavily on the (unsatisfied) normality assumption • Cannot be applied to datasets with many observed zero intakes	• Adjustment for covariates is not done in the context of a unified model • Cannot be applied to datasets with many observed zero intakes	• The two parts of the model are estimated independently, ignoring the possibility of correlation between probability to consume and usual amount consumed • Adjustment for covariates is only done for one part of the model, not in the context of a unified model • Complex two-stage transformations may fail to obtain exact normality, but the subsequent steps of the method rely heavily on the (unsatisfied) normality assumption • Software does not produce standard errors for estimated parameters of usual intake distributions

[a]NRC=National Research Council.
[b]IOM=Institute of Medicine.
[c]ISU=Iowa State University.
[d]BP=Best-Power.
[e]ISUF=Iowa State University Foods.
[f]SAS Institute Inc, Cary, NC.
[g]SIDE=Software for Intake Distribution Estimation (Iowa State University, Ames, IA).

Figure 4. Summary of the strengths and weaknesses of the different statistical modeling methods for estimating usual intake distributions. (Data from this figure are available online at www.adajournal.org as part of a PowerPoint presentation featuring additional online-only content.)

plemented only in specialized software, whereas others may be implemented in standard statistical software packages. Furthermore, obtaining standard errors of estimated parameters of usual intake is difficult, if not impossible, for certain methods. Although these practical differences between the methods may play an important role, the remainder of this report focuses on how theoretical differences between the methods may influence one's choice of method.

The National Research Council, Iowa State University, and Best-Power methods have been applied to data from food consumption surveys to estimate usual nutrient intake for the US population (16-18,28). The Institute of Medicine (19) recommends the use of the Iowa State University method or alternatively (for small, simple random samples) their own method. However, this recommendation ignores the fact that the two methods make different assumptions concerning the 24-hour recall as an assessment instrument—whether or not a 24-hour recall is unbiased for usual intake in the original scale (Assumption B) vs the transformed scale (Assumption A). We believe a better recommendation is to consistently use modeling methods that share Assumption B. That is, when estimating distributions of usual nutrient intake, to use the Iowa State University method whenever possible (per the Institute of Medicine's recommendation), and to use the Best-Power method with samples that are too small for the full Iowa State University method. This recommendation also ensures that the estimate of the group mean usual intake coincides with the overall average of the 24-hour recalls and is, therefore, consistent with the practice of tracking average intake over time.

Some evidence exists that neither Assumption A nor Assumption B truly holds. An analysis (8) of the Observing Protein and Energy Nutrition Study investigated how systematic biases or random error associated with the assessment instrument affects the estimation of usual intake distributions from two FFQs or two 24-hour recalls, through comparison to distributions based on unbiased (recovery) biomarkers for protein and energy intake. Both instruments displayed systematic bias as well as within-person variation. The 24-hour recall instrument showed substantially smaller bias than the FFQ and greater within-person variation. Accordingly, the distributions estimated from 24-hour recalls—after adjustment for within-person variation using a method similar to that described in Figure 3—agreed with the biomarker-derived distributions much more closely than did the distributions estimated from FFQs. This is demonstrated in Figure 5. Figure 5 shows distributions of usual energy intake for female participants in the Observing Protein and Energy Nutrition study, as estimated by different methods and assessment instruments. The distribution of doubly labeled water measurements, labeled Biomarker, represents the gold standard. The distribution labeled FFQ is based on the first FFQ measurement per respondent, whereas the distribution labeled 24-hour recall-A is estimated by the Institute of Medicine method (under Assumption A) and the distribution labeled 24-hour recall-B is estimated by the Iowa State University method (under Assumption B). The adjusted back-transformations required by Assumption B shift usual intake distributions estimated with either of Nusser and colleagues' methods (Iowa State University or Best-Power method) to the right of distributions estimated with the Institute of Medicine/National Research Council methods. The amount of shift depends on both the magnitude of within-person variation and the normality transformation used.

Except for the small number of dietary components for which unbiased biomarkers have been discovered, it is impossible to know which of Assumptions A or B is more appropriate. The 24-hour recall may overestimate true usual intake of some specific foods or food groups, but based on the Observing Protein and Energy Nutrition study results for energy (arguably a measure of total food intake) we conclude that the 24-hour recall has a general tendency toward underestimation. The bias-adjustment of the Iowa State University, Best-Power, and Iowa State University Foods methods partially offsets this tendency, making Assumption B more attractive from a practical standpoint.

One limitation common to all of the methods is that looking for differences in usual intake between sexes or among age groups requires separate estimation for each subgroup. A more efficient approach is to model the group means themselves as a function of covariates so that the person-specific deviation is from the group mean of similar individuals. In addition to allowing statistical tests for differences between subgroups defined by covariates such as age and sex, such an approach also allows the adjustment for time-related biases as part of a single, unified model (13), rather than being a preliminary data adjustment as in the Iowa State University methods.

One of the main tasks in modeling usual intake from 24-hour recalls is to estimate within- and between-person variance, which requires that multiple 24-hour recalls be available for at least some individuals in the sample. In the case of the Iowa State University Foods method, the same number of 24-hour recalls must be available for all individuals. The small number of 24-hour recalls typically available per individual poses a special problem when estimating usual food intake. When 2 days of data are available, observed consumption probabilities take on only three values: zero, one-half, and one. Estimating a smooth distribution of consumption probabilities from such discrete data can be difficult for the Iowa State University Foods method. Also, respondents who truly consume a food regularly could still have zero consumption on one or both 24-hour recalls, making it more difficult to separate within- from between-person variability in the remaining nonzero 24-hour recalls.

Even in cases where the Iowa State University Foods method can successfully estimate the distributions of consumption probability and usual consumption amount, the method combines the two distributions assuming independence of probability and amount. This can lead to unsatisfactory estimates of the usual intake distribution when the assumption does not hold. Furthermore, the Iowa State University Foods method permits adjustment for potential time-related biases only in the estimation of the consumption-day usual intake distribution, and not in the estimation of the consumption probability distribution. This can also lead to less than satisfactory estimates.

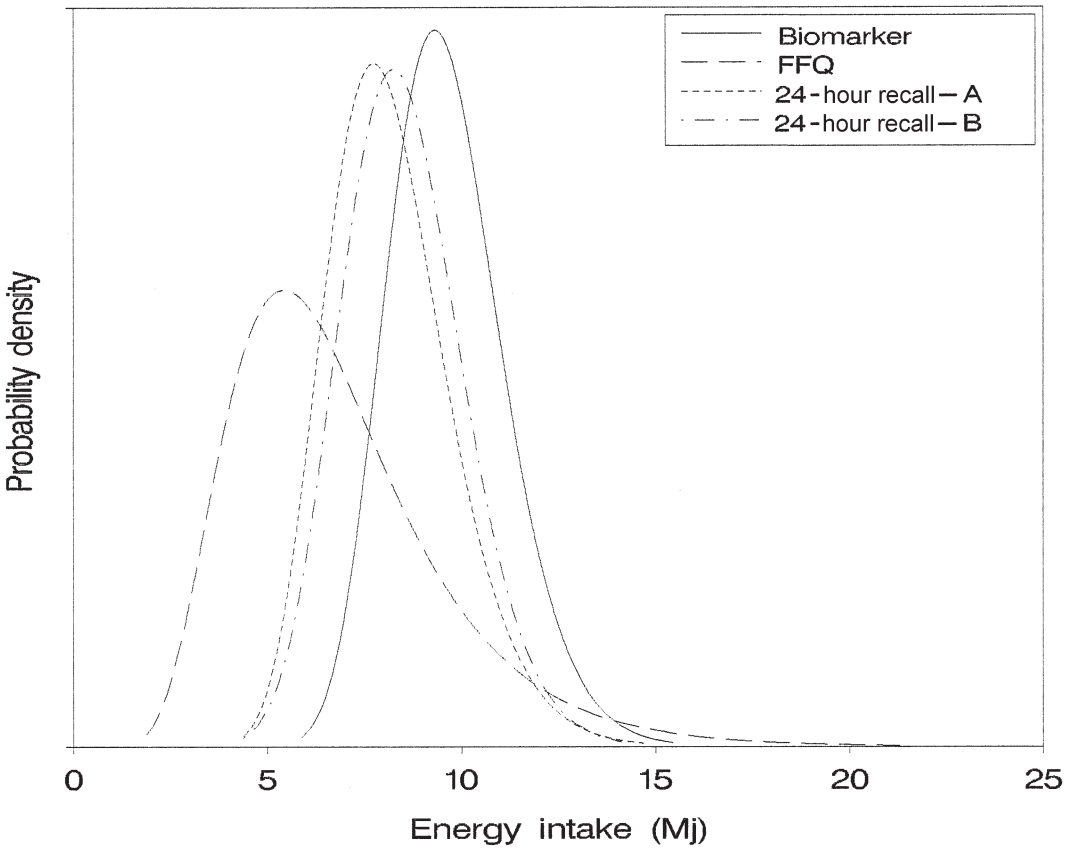

Figure 5. Distributions of usual energy intake for female participants in the Observing Protein and Energy Nutrition study, as estimated by different modeling methods and assessment instruments. Biomarker=the distribution of doubly labeled water measurements; this represents the gold standard. FFQ=food frequency questionnaire distribution. 24-hour recall–A=this distribution is based on two 24-hour recalls per respondent and modeled under the assumption that the respective instrument is unbiased for usual intake on some transformed scale. 24-hour recall–B=this distribution is based on two 24-hour recalls per respondent and modeled under the assumption that the respective instrument is unbiased for usual intake on the untransformed scale.

CONCLUSIONS

Modeling usual intake of episodically consumed foods from surveys in which a limited number of 24-hour recalls per respondent are the primary instrument presents special challenges. The Iowa State University Foods method addresses these challenges because it accounts for days without consumption; accounts for consumption-day amounts that are positively skewed, with extreme values in the upper tail; and distinguishes within-person variability, consisting of reporting errors and day-to-day variation in intake, from between-individual variability due to differences in usual intake.

We have identified some areas where the Iowa State University Foods method could be extended and improved. In addition to meeting the challenges mentioned above, an improved method should allow correlation between the probability of consuming a food and the amount consumed on a consumption day and be able to incorporate covariate information relating to 24-hour recall intake, to allow efficient estimation and comparison for subgroups.

The first of two companion articles to this one, by Subar and colleagues (12), describes the concept and development of an FFQ designed to provide covariate information to help explain between-person differences in consumption probability and its correlate, consumption-day usual intake. The second, by Tooze and colleagues (13), outlines an integrated modeling framework that addresses all of the challenges listed here.

The authors thank Phillip S. Kott, Joseph D. Goldman, and Richard P. Troiano for thoughtful reviews and Anne Brown Rodgers for her expert editing assistance.

References

1. Institute of Medicine, Food and Nutrition Board. Dietary Reference Intakes: Applications in dietary assessment. Available at: http://www.nap.edu/books/0309071836/html. Accessed August 21, 2006.
2. *Nutrition and Your Health: Dietary Guidelines for Americans 2005*. Washington, DC: US Department of Agriculture; 2005. Home and Garden Bulletin No. 232.
3. US Department of Health and Human Services. Tracking Healthy People 2010. Available at: http://

www.cdc.gov/nchs/hphome.htm. Accessed July 7, 2005.
4. US Department of Agriculture. MyPyramid. Available at: http://www.mypyramid.gov. Accessed July 22, 2005.
5. Flegal KM, Larkin FA. Partitioning macronutrient intake estimates from a food frequency questionnaire. *Am J Epidemiol.* 1990:131:1046-1058.
6. Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, Sharbaugh CO, Trabulsi J, Runswick S, Ballard-Barbash R, Sunshine J, Schatzkin A. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The OPEN study. *Am J Epidemiol.* 2003;158:1-13.
7. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, Bingham S, Schoeller DA, Schatzkin A, Carroll RJ. Structure of dietary measurement error: Results of the OPEN biomarker study. *Am J Epidemiol.* 2003;158:14-21.
8. Freedman LS, Midthune D, Carroll RJ, Krebs-Smith S, Subar AF, Troiano RP, Dodd K, Schatzkin A, Ferrari P, Kipnis V. Adjustments to improve the estimation of usual dietary intake distributions in the population. *J Nutr.* 2004;134:1836-1843.
9. Beaton GH, Milner J, Corey P, McGuire V, Cousins M, Stewart E, de Ramos M, Hewitt D, Grambsch PV, Kassim N, Little JA. Sources of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. *Am J Clin Nutr.* 19797;32: 2546-2559.
10. Beaton GH, Milner J, McGuire V, Feather TE, Little JA. Source of variance in 24-hour dietary recall data: Implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. *Am J Clin Nutr.* 1983;37:986-995.
11. Carriquiry AL. Estimation of usual intake distributions of nutrients and foods. Proceedings from the workshop Future Directions for What We Eat in America—NHANES: The integrated CSFII-NHANES. *J Nutr.* 2003;133(suppl 2):609S-623S.
12. Subar AF, Dodd KW, Guenther PM, Kipnis V, Midthune D, McDowell M, Tooze JA, Freedman LS, Krebs-Smith SM. 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*. 2006;106:1556-1563.
13. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, Guenther PM, Carroll RJ, Kipnis V. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc*. 2006;106:1575-1587.
14. Garn SM, Larkin FA, Cole PE. The real problem with 1-day diet records. *Am J Clin Nutr.* 1978;31:1114-1116.
15. Hegsted DM. Problems in the use and interpretation of the Recommended Dietary Allowances. *Ecol Food Nutr.* 1972;1:255-265.
16. Guenther PM, Dodd KW, Reedy J, Krebs-Smith SM. Most Americans eat much less than recommended amounts of fruits and vegetables. *J Am Diet Assoc.* 2006;106:1371-1379.
17. US Department of Agriculture Agricultural Research Service, Food Surveys Research Group, Beltsville Human Nutrition Research Center. What we eat in America, NHANES 2001-2002: Usual nutrient intakes from food compared to Dietary Reference Intakes. Available at: http://www.ars.usda.gov/ba/bhnrc/fsrg. Accessed August 24, 2005.
18. National Research Council, Subcommittee on Criteria for Dietary Evaluation. *Nutrient Adequacy: Assessment Using Food Consumption Surveys.* Washington, DC: National Academy Press; 1986.
19. Institute of Medicine, Food and Nutrition Board. Dietary Reference Intakes: Applications in dietary planning. Available at: http://www.nap.edu/books/0309088534/html. Accessed June 24, 2006.
20. Nusser SM, Carriquiry AL, Dodd KW, Fuller WA. A semiparametric transformation approach to estimating usual daily intake distributions. *J Am Stat Assoc.* 1996;91:1440-1449.
21. Guenther PM, Kott PS, Carriquiry AL. Development of an approach for estimating usual nutrient intake distributions at the population level. *J Nutr.* 1997; 127:1106-1112.
22. Haines PS, Hama MY, Guilkey D, Popkin BM. Weekend eating in the United States is linked with greater energy, fat, and alcohol intake. *Obes Res.* 2003;11: 945-949.
23. Bailar BA. The effects of rotation group bias on estimates from panel surveys. *J Am Stat Assoc.* 1975;70: 23-29.
24. Nusser SM, Fuller WA, Guenther PM. Estimating usual dietary intake distributions: Adjusting for measurement error and non-normality in 24-hour food intake data. In: Lyberg L, Biemer P, Collins M, Deleeuw E, Dippo C, Schwartz N, Trewin D, eds. *Survey Measurement and Process Quality.* New York, NY: Wiley; 1997:689-709.
25. Dodd KW, Carriquiry AL, Fuller WA, Chen C-F. On the estimation of usual intake distributions for foods. *Proceedings of the Section of Survey Research Methods*. Alexandria, VA: American Statistical Association; 1996:304-310.
26. Wright JD, Kennedy-Stephenson J, Wang CY, McDowell MA, Johnson CL. Trends in intake of energy and micronutrients—United States, 1971-2000. *MMWR Morbid Mortal Wkly Rep.* 2004;53:80-82.
27. Enns CW, Goldman JD, Cook A. Trends in food and nutrient intakes by adults: NFCS 1977-78, CSFII 1989-91, and CSFII 1994-95. *Fam Econ Nutr Rev.* 1997;10:1-15.
28. Carriquiry AL. Assessing the prevalence of nutrient inadequacy. *Public Health Nutr.* 1999;2:23-33.
29. Bain LJ, Engelhardt M. *Introduction to Probability and Mathematical Statistics.* Boston, MA: PWS-Kent Publishing; 1989:197.