Development of an Approach for Estimating Usual Nutrient Intake Distributions at the Population Level[1,2]

Patricia M. Guenther, Phillip S. Kott*[3] and Alicia L. Carriquiry[†]

*U.S. Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Riverdale, MD 20737; *U.S. Department of Agriculture, National Agricultural Statistics Service, Fairfax, VA 22030; and [†]Department of Statistics, Iowa State University, Ames, IA 50011*

ABSTRACT Assessment of the dietary intake of a population must consider the large within-person variation in daily intakes. A 1986 report by the National Academy of Sciences (NAS), commissioned by the U.S. Department of Agriculture (USDA), marked an important milestone in the history of this issue. Since that time, USDA has been working cooperatively with statisticians at Iowa State University (ISU), who have further developed the measurement error model approach proposed by NAS. The method developed by the ISU statisticians can be used to estimate usual dietary intake distributions for a population but not for specific individuals. It is based on the assumption that an individual can more accurately recall and describe the foods eaten yesterday than foods eaten at an earlier time. The method requires as few as two independent days of nutrient intake information or three consecutive days for at least a subsample of the individuals. It removes biases of subsequent reporting days compared with the first day, and temporal effects such as day-of-the-week and seasonal effects can be easily removed. The method developed at ISU is described conceptually and applied to data collected in the 1989–91 USDA Continuing Survey of Food Intakes by Individuals to estimate the proportion of men and women age 20 y and older having ''usual'' (long-run average) intakes below 30% of energy from fat, below the 1989 Recommended Dietary Allowances for vitamin A and folate, and above 1000 μg for folate. These results were compared with the results from the distributions of 1-d intakes and of 3-d mean intakes to demonstrate the effect of within-person variation and asymmetry on usual nutrient intakes in a population. J. Nutr. 127: 1106–1112, 1997.

KEY WORDS: • *dietary assessment* • *dietary fat* • *folate* • *vitamin A* • *humans*

Food consumption data are collected for a wide variety of reasons, using a wide variety of methods and procedures. Dietary data are collected in surveys that monitor the dietary and health status of the population, in epidemiologic studies and in clinical trials. They are used to judge the nutritional adequacy of diets, to evaluate the effectiveness of food assistance programs and in food safety risk assessments. The purpose of this paper is to review the problems associated with estimating distributions of usual dietary intakes of populations and to describe a useful approach for comparing such estimates with standards set for a variety of assessment purposes. It is important to note that this approach does not estimate usual dietary intake distributions for specific individuals.

At present, the most commonly used dietary data collection methods are interviewer-administered 24-h recalls, self-administered food records and food-frequency questionnaires, which may be either interviewer- or self-administered. In a 24-h recall, the interviewer asks the respondent to list all foods and beverages consumed during the previous day. Probing questions are used to obtain the desired level of detail for the descriptions and amounts of foods eaten. Food records require the respondent to provide a written description of the types and amounts of foods eaten. Food-frequency questionnaires provide a list of foods and groups of foods, and respondents are asked how often they eat each item on the list.

One of the most important estimation issues relates to the temporal aspects of dietary intake estimation. If each individual ate the same thing every day, day-to-day, week-to-week or season-to-season changes would not be of concern, but this is not the case. Within-individual variation in daily dietary intakes presents a difficult problem, and its importance has long been recognized (Anderson 1988, Beaton et al. 1979, Garn et al. 1978, Hegsted 1972, Keys 1967, Marr 1971, Sempos et al. 1985). Nutritionists want to measure something called ''usual'' or ''habitual'' or ''customary'' daily intake, but even a definition of this concept has seldom been clearly articulated. Here we define ''usual'' as ''long-run daily average,'' where ''long-run'' is effectively a year.

Many questions of interest about dietary intake can be answered by determining usual intakes of groups or by comparing the usual intakes of different groups. Fortunately, a mean of

[1] Supported in part by Cooperative Agreement No. 58-3198-2-006 between the Agricultural Research Service, U.S. Department of Agriculture and the Center for Agriculture and Rural Development, Iowa State University. A.L.C. is funded in part by Research Grant No. 000149610279 from the Office of Naval Research, U.S. Department of Defense.
[2] The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked ''advertisement'' in accordance with 18 USC section 1734 solely to indicate this fact.
[3] To whom correspondence should be addressed.

Manuscript received 13 September 1996. Initial review completed 25 November 1996. Revision accepted 11 February 1997.

1-d intakes by individuals in a group can be an unbiased estimate of the group's usual mean intake. But this is true only if those single days are a good representation of all days. That is, they must represent an appropriate mix of days of the week, months and seasons if they are to represent the usual intake of the group. For example, if one wishes to estimate the usual mean percentage of fat intake for a group and that group is more likely to eat away from home on Friday than on other days of the week, Fridays should not be over- or underrepresented in the data collected because food eaten away from home is typically higher in fat and alcohol content than food eaten at home (Guenther and Ricart 1990). Similarly, if one wishes to estimate the usual mean intake of milk by children, summer should not be over- or underrepresented in the data because school-age children are less likely to drink milk in the summer than in other seasons of the year. A study of teenagers' beverage consumption in the U.S. showed that milk intake was 20% lower in the summer than in the spring, whereas tea consumption was 90% higher (Guenther 1986).

The answers to many other important questions, however, require knowledge of the distributions of usual daily intakes. Such distributions are desired for risk analyses related to dietary adequacy and food safety and for measuring progress towards dietary objectives. For example, not only do we want to know the mean usual percentage of energy from fat for a certain population, but we also want to know the proportion of that population with a usual percentage of energy from fat of 30% or less. The second question is more difficult to answer than the first because it requires that the entire distribution of usual intakes be estimated, not just the mean.

A decade ago, the U.S. Department of Agriculture (USDA) commissioned the National Academy of Sciences (NAS) to investigate the question of how to assess the adequacy of nutrient intake (Subcommittee on Criteria for Dietary Evaluation 1986). That report marked an important milestone. The Ten-Year Comprehensive Plan for the National Nutrition Monitoring and Related Research Program calls for implementing the NAS recommendations (Department of Health and Human Services and Department of Agriculture 1993). The method discussed here implements the probability approach outlined in the NAS report, improving it where necessary. The centerpiece of the approach is a measurement error model that treats the intake observed for any individual on any given day as the sum of that individual's true usual intake and a random "disturbance" or "measurement error" for that individual on that day. The NAS approach requires estimating two population distributions, the distribution of nutrient requirements and the distributions of usual nutrient intakes.

We focus here on the challenge of estimating usual intake distributions and recognize the difficulties associated with estimating nutrient requirements. Even without the requirements distributions, it will be possible to use the usual intake distributions to determine what proportion of the population meets various dietary standards and objectives that have been promulgated. Reliable estimates of usual intake distributions should also be helpful to those who formulate such standards and objectives, for example, the Recommended Dietary Allowances (Subcommittee on the Tenth Edition of the RDAs 1989) and the Healthy People 2000 nutrition objectives (U.S. Department of Health and Human Services 1990).

Our approach to estimating usual intake distributions is based on the assumption that an individual can more accurately recall and describe the types and amounts of foods eaten yesterday than the types and amounts of foods eaten over any longer period of time. We also assume that the nutrient database used can adequately reflect the nutrient content of the foods eaten at that time.

METHODS

USDA has worked cooperatively with statisticians at Iowa State University (ISU) who have developed a method for estimating usual intake distributions based on the NAS recommendations. We describe the method only in a conceptual manner here. Nusser et al. (1996) present the technical details and discuss a simulation study conducted to validate the method by assessing its performance and comparing it with other procedures.

The method developed at ISU controls the within-person variation, or the day-to-day variability, of nutrient intake. Dietary data contain both within- and between-person variation. When analyzing such data, only the between-person variation is of interest. That is, the within-person variation should be removed, i.e., the variation of "usual" intakes is of interest.

Nutritionists have sought to remove or minimize this within-person variation by lengthening the observation period from 1 to 3, 6, 7, 14, 30, 90 or 365 d. After having done so, questions remain about the accuracy of the dietary intake information collected (see, for example, Smith et al. 1991) and the accuracy of the nutrient intake estimates. Is it at all possible to find subjects who provide this information accurately over what may be an extended period of time yet still represent populations of interest in a scientifically defensible manner? If not, can the errors be measured and dealt with successfully in the estimation process?

The method developed at ISU takes much of the burden of estimating usual intake away from the subjects. Because the procedure's main goal is to obtain usual intake distributions at the population level, only a few days of intake data for sampled individuals are required. If dietary intake data are collected on independent days, 2 d suffice for at least some individuals in the sample; if data are collected on consecutive days, 3 d are needed for at least some individuals in the sample. It should be noted, however, that the method does *not* produce estimates of usual intakes for particular individuals in the sample.

A second fundamental problem addressed by the method developed at ISU is that distributions of usual nutrient intakes are often right-skewed. Furthermore, USDA food intake data are often collected in surveys having complex designs. Many statistical procedures are based on the assumptions that data are normally distributed and are from simple random samples. These procedures typically are not robust to violations of those assumptions. The method developed at ISU deals with both problems directly. Furthermore, because the method does not require that intake distributions be normal, intake values that are extreme, but perhaps valid, need not be discarded.

The method developed at ISU assumes that a reported 1-d nutrient intake for an individual found in a food intake survey dataset can be represented conceptually as

$$y_{its} = x_i + c_t + b_s + e_{it} \quad (1)$$

where y_{its} is the reported nutrient intake by individual i for date t, on day s, which is the sequence number of the day for which the individual has provided intake information for the survey. The value we are interested in estimating is x_i, the usual intake of individual i. The second term on the right side of Equation (1), c_t, represents the temporal effect on nutrient intake caused by the particular day of the week and time of the year. The third term, b_s, denotes the bias associated with intakes on a particular reporting day of the survey. The last term, e_{it}, is simply the difference between the reported intake, y_{its}, and the other three terms.

It is assumed that the bias effects, b_s, for the first day's reported nutrient intake values are negligible. This means not only that the individual reports food intakes for that day without systematic bias, but also that the nutrient values assigned to those intakes are unbiased.

Whether or not there is bias caused by underreporting of food intake on the first day of reported intake should not obscure one of the important attributes of the ISU method, namely, it removes biases

TABLE 1

Outline of steps in the method developed at Iowa State University for estimating usual nutrient intake distributions

1. Initial data adjustments
   - A. Shift of data away from zero
   - B. Initial power (or log) transformation
   - C. Regression adjustment for nuisance effects
   - D. Homogenization of subsequent intake days to match distribution of first day
   - E. Creation of equal-weight sample
   - F. Back transformation to original scale
2. Measurement error model
   - A. Full transformation to normal scale (power transformation followed by grafted polynomial)
   - B. Estimation of within- and between-individual variances
   - C. Test for heterogeneity of within-individual variances and consequent adjustment of variance component estimates
   - D. Estimation of usual intake distribution in normal scale
   - E. Bias-adjusted back transformation to original scale of a representative sample from the intake distribution
3. Estimation of usual intake distribution in original scale

of subsequent reporting days compared with the first, such as training or conditioning effects (Pao et al. 1985, U.S. Department of Agriculture 1987). In addition, temporal effects, such as day-of-the-week and time-of-year effects, can also be removed from datasets in which such temporal factors are recorded.

The output of the method developed at ISU is an estimation of the distribution of usual intakes in the population. The procedure involves a number of steps, which are summarized in **Table 1**. First, the temporal effects are removed using a power transformation followed by a regression adjustment that has been modified to avoid negative adjusted intake values.

This regression adjustment corrects only for the bias in the means of reported intake for days following the first survey day. Biases in higher-order moments, such as variances, also have to be removed from the later-day reported intakes. The technique for doing this in the method developed at ISU can be called homogenization because it results in each day having virtually the same distribution of intake values.

After the regression adjustment and homogenization, the basic question still remains: what is the distribution of the x_i, the usual intakes? Intakes of many nutrients, even after they have been regression-adjusted and homogenized, still do not have normal distributions. The ISU method produces a continuous transformation that maps homogenized intakes for a nutrient into a standard normal distribution. This step goes beyond the simple transformations suggested in the NAS report (Subcommittee on Criteria for Dietary Evaluation 1986). By rigorously transforming to normality, we can properly take advantage of standard statistical techniques developed for the normal distribution when estimating parameters such as between-person variances in intakes.

An assumption underlying the very existence of such a transformation is that the distribution of the original 1-d intakes is continuous. This means, for example, that individual 1-d intakes cannot cluster at specific values. A transformation to normality can be developed for nutrients because they have continuous 1-d distributions; but for many foods, 1-d distributions cluster at zero. As a result, the method developed at ISU cannot at present be directly applied to foods. Research on estimating usual food intake distributions is underway (Nusser et al. 1997).

Another important feature of the method developed at ISU is that survey sampling weights can be incorporated into the homogenized intake distribution so that it truly estimates the intake distribution of the target population and not just the sample. In the example below, the weights for the 1989–91 USDA Continuing Survey of Food Intakes by Individuals are calibrated so that the weighted sample matches what is known about the U.S. population with respect to 13 variables believed to be associated with eating behavior. These variables range from the presence of children in the household to income level (Fuller et al. 1994, Kott and Guenther 1993).

After the homogenized data are transformed into a normally distributed data set, our equation is the same as before

$$h^*_{it} = x^*_i + e^*_{it} \qquad (2)$$

but now the transformed intake values, the h^*_{it}, are normally distributed. Standard statistical techniques are then applied to this measurement error model to estimate the distribution of usual intakes from the transformed variables, the x^*_i in Equation (2). Then these estimated, normally distributed x^*_i values are mapped into the original scale through a bias-adjusted back transformation, and the distribution of original-scale usual intakes is estimated. A more technical discussion of these steps is found in Nusser et al. (1996).

RESULTS

The method described above has been implemented in a software package called SIDE (Software for Intake Distribution Estimation) at ISU (Department of Statistics and Center for

TABLE 2

Estimated distributions of a single day's intake of fat, means of 3 d of intake and usual daily intake in populations of men and women 20 y and older, 1989–91

	Fat					
	Men			Women[1]		
Percentile	One day	Three-day mean % of energy	Usual[2]	One day	Three-day mean % of energy	Usual[2]
10	23.2	26.4	29.0	21.5	24.5	27.3
25	29.2	30.7	32.0	27.6	29.3	30.5
50	35.3	35.3	35.1	34.2	34.2	34.0
75	41.1	39.7	38.3	40.4	38.7	37.4
90	46.6	43.4	41.1	45.9	42.8	40.4

[1] Pregnant and lactating women excluded.
[2] Adjusted for day of week, month of year, and sequence of surveyed days.
Source: U.S. Department of Agriculture, Continuing Survey of Food Intakes by Individuals (1996) (3,371 men; 4,606 women; these numbers exclude individuals who fasted on one of the three reported intake days).

TABLE 3

Folate: Estimated distributions of a single day's intake of folate, means of 3 d of intake, and usual daily intake in populations of men and women 20 y and older, 1989–91

	Folate					
	Men			Women[1]		
Percentile	One day	Three-day mean	Usual[2]	One day	Three-day mean	Usual[2]
		μg/d			*μg/d*	
10	101	128	158	72	96	118
25	148	173	202	115	133	153
50	233	242	269	182	188	205
75	355	342	353	281	264	270
90	533	489	454	398	357	344

[1] Pregnant and lactating women excluded.
[2] Adjusted for day of week, month of year, and sequence of surveyed days.
Source: U.S. Department of Agriculture, Continuing Survey of Food Intakes by Individuals (1996) (3,381 men; 4,621 women).

Agricultural and Rural Development, ISU 1996). This software (Version 1.0) was applied to data from the 1989–91 Continuing Survey of Food Intakes by Individuals (CSFII) to estimate the distributions of usual dietary intakes for men and nonpregnant, nonlactating women age 20 y and older (U.S. Department of Agriculture 1996). The data were collected using an interviewer-administered 24-h recall of food intake, followed by a self-administered 2-d food record (U.S. Department. of Agriculture 1995). Intake estimates do not include nutrients consumed in the form of dietary supplements. The method developed at ISU does not require that all individuals in the sample have the same number of days of intake data; however, here we used only those individuals who provided all 3 d of intake information.

The tables and figures display estimates of the population's intake on any day (based on each respondent's first day of intake data), estimates of the population's mean daily intake during any 3-d period (based on each respondent's 3-d mean intake), and estimates of the population's usual daily intake (based on each respondent's 3 d of intake data). The percentiles (**Tables 2–4**) show the general patterns of the distributions shrinking towards the center as we move from the 1-d to the 3-d mean to the usual intake distributions, that is, the 10th and 25th percentiles increase and the 75th and 90th decrease. For the more asymmetric distributions, folate and vitamin A, the median values (50th percentile) increase slightly from the 1-d to the 3-d mean to the usual. The means of the distributions (**Table 5**) are quite similar to each other as expected.

In **Table 6**, we show the proportion of the population meeting recommended levels of intake on any single day, the proportion meeting the recommendation during any 3-d period, and the proportion whose usual intake met the recommendation. For other nutrients not shown here, the extent of over- or underestimation of the proportion of the population having usual intakes above or below any cut point caused by using only 1 d or the mean of a few days will be a function of the ratio of the within-person variation to the between-person variation in intake, of how skewed the true distribution is, and of where the cut point lies on the distribution.

TABLE 4

Estimated distributions of a single day's intake of vitamin A, means of 3 d of intake, and usual daily intake in populations of men and women 20 y and older, 1989–91

	Vitamin A					
	Men			Women[1]		
Percentile	One day	Three-day mean	Usual[2]	One day	Three-day mean	Usual[2]
		μg RE/d			*μg RE/d*	
10	212	345	508	161	262	388
25	404	510	603	313	415	551
50	746	834	987	608	680	789
75	1345	1341	1455	1061	1139	1171
90	2281	2085	2076	1942	1728	1683

[1] Pregnant and lactating women excluded.
[2] Adjusted for day of week, month of year, and sequence of surveyed days.
Source: U.S. Department of Agriculture, Continuing Survey of Food Intakes by Individuals (1996) (3,381 men; 4,621 women).

TABLE 5

Estimated mean intakes of selected nutrients using 1 d of intake per individual, individual 3-d means, and usual intake program described in text in populations of men and women 20 y and older, 1989–91

Population group	Nutrient	*n*	Estimated mean nutrient intake		
			One day	Three days	Usual[1]
Men	Energy from fat, %	3371	35.0	35.1	35.1
	Folate, μg/d	3381	286	290	292
	Vitamin A, μg RE/d	3381	1117	1149	1182
Women[2]	Energy from fat, %	4606	33.9	33.9	33.9
	Folate, μg/d	4621	219	220	221
	Vitamin A, μg RE/d	4621	912	928	947

[1] Adjusted for day of week, month of year, and sequence of surveyed days.
[2] Pregnant and lactating women excluded.
Source: U.S. Department of Agriculture, Continuing Survey of Food Intakes by Individuals (1996).

High fat intake is a current public health concern in the United States (Federation of American Societies for Experimental Biology, Life Sciences Research Office 1995). A recommended level of intake is 30% of energy or less (U.S. Department of Agriculture 1995). If this level is considered to be the target level for the population's usual intake, as it is in the Year 2000 Objectives (Department of Health and Human Services 1990), then using any of the estimated mean or median intakes as an indicator of usual intake would yield similar conclusions: the population's usual intake of fat was 34–35% of energy in 1989–91 (Tables 2 and 5). However, if 30% is considered to be the desired level of usual intake for all individuals in the population, then using either the 1-d or the 3-d mean distribution or the usual intake distribution yields conclusions that are quite different. As illustrated in **Figure 1**, for example, on any given day in 1989–91, 28% of men had a fat intake below 30% of energy and during any 3-d period, 22% of men had a mean daily intake below that level. However, what is of interest is their usual daily intake, and only 15% of men had usual intakes below the recommended level. Differences of this magnitude have important implications for nutrition and public health policy.

We note that the percentage of energy from fat should not be treated as a single nutrient because it is a ratio of two variables, fat and energy. It is known that, in general, ratios have different statistical properties than single variables. For illustrative purposes here, however, the example is appropriate. For these examples, we computed the ratio of fat to energy for each individual day prior to any further analysis.

Folate is a nutrient of current public health interest in part because of its relationship to the development of neural-tube defects. About two fifths of nonpregnant, nonlactating women and one fourth of men had usual intakes of folate from food below the 1989 Recommended Dietary Allowances (Subcommittee on the Tenth Edition of the RDAs 1989) (Table 6). Dietary guidance to increase fruit and vegetable consumption and fortification of the food supply have been proposed to improve this situation. Because high folate intake can mask the hematologic signs of pernicious anemia, an upper limit of 1000 μg/d of folate (from all sources) has been proposed (U.S. Department of Health and Human Services 1992). We estimate that in 1989–91, 1.4% of men and 0.6% of women had intakes (from food only) as high as the upper limit on any single day, and 0.05% of men and <0.01% of women had usual intakes exceeding that level. **Figures 2** and **3** show the differences in the distributions of 1-d intakes, distributions of 3-d mean intakes, and the usual distributions for folate and vitamin A intakes by men. The vertical lines indicate the 1989 Recommended Dietary Allowances.

For vitamin A, a nutrient having a notoriously skewed intake distribution, it is interesting to note that although the means of the three distributions are similar (Table 5), the medians differ greatly (Table 4). Among both men and women, the median value for the estimated distribution of

TABLE 6

Estimated proportion of the population meeting recommended intake levels for selected nutrients on 1 d, on 3 d, and the estimated proportion whose usual intake meets the recommendation in populations of men and women 20 y and older, 1989–91

Population group	Nutrient	Recommended level	Proportion meeting recommended level		
			One day	Three days	Usual[1]
Men	Fat[2]	30% of energy	0.28	0.21	0.14
	Folate	200 μg/d	0.59	0.65	0.76
	Vitamin A	1000 μg RE/d	0.37	0.40	0.49
Women[3]	Fat[2]	30% energy	0.33	0.28	0.22
	Folate	180 μg/d	0.49	0.53	0.61
	Vitamin A	800 μg RE/d	0.36	0.42	0.49

[1] Adjusted for day of week, month of year, and sequence of surveyed days.
[2] Individuals who fasted on one of the three reported intake days excluded.
[3] Pregnant and lactating women excluded.
Source: U.S. Department of Agriculture, Continuing Survey of Food Intakes by Individuals (1996) (3,381 men; 4,621 women).

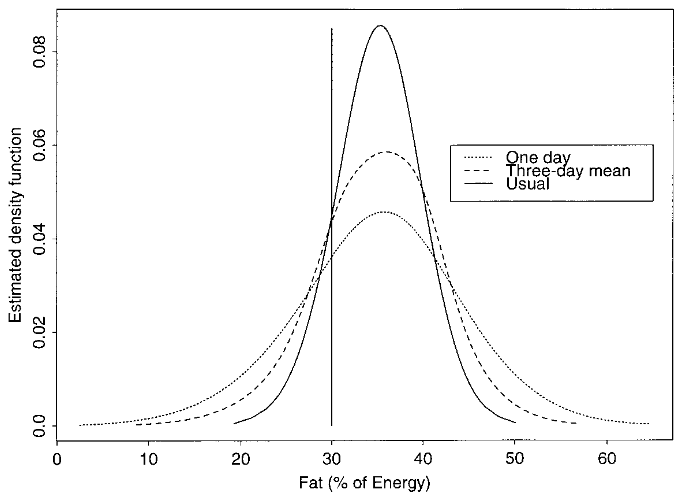

FIGURE 1 Estimated density functions of fat intake (expressed as a percentage of energy intake) on a single day, during a 3-d period, and the distribution of usual daily intakes, for the population of men age 20 y and older based on data from the Continuing Survey of Intakes by Individuals, 1989–91 (U.S. Department of Agriculture 1996). Total area under each curve is equal to 1 (100%). Area under solid curve to the left of the vertical line estimates the proportion of men with usual intakes meeting the 30% recommendation. The analogous area under the 1-d (or 3-d) curve estimates the proportion of men with 1-d (or 3-d mean) intakes meeting the recommendation.

1-d intakes is about 23% lower than the estimated median of the usual intakes. The first quartile is more than 40% lower, whereas the third quartile of the 1-d intakes is within 6% of the third quartile of usual intakes.

These results demonstrate the effects of not removing within-person variation. In addition, the ISU method accommodates any skewed distributions directly.

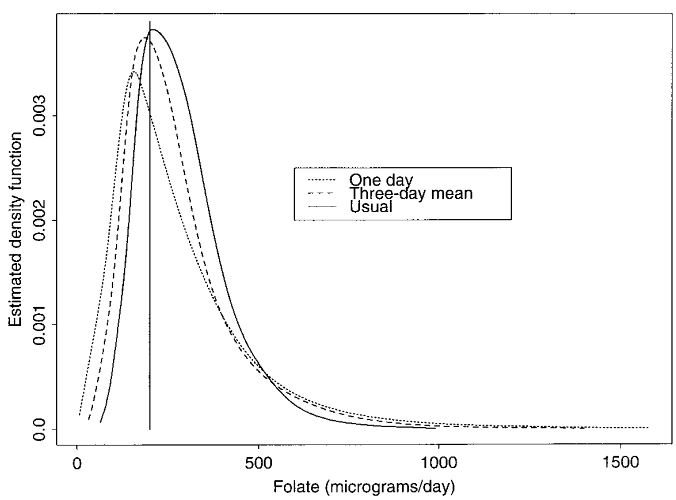

FIGURE 2 Estimated density functions of folate intake on a single day, during a 3-d period, and the distribution of usual daily intakes, for the population of men age 20 y and older based on data from the Continuing Survey of Intakes by Individuals, 1989–91 (U.S. Department of Agriculture 1996). Total area under each curve is equal to 1 (100%). Area under the solid curve to the right of the vertical line estimates the proportion of men with usual intake meeting the Recommended Dietary Allowance. The analogous area under the 1-d (or 3-d) curve estimates the proportion of men with 1-d (or 3-d mean) intakes meeting the recommendation.

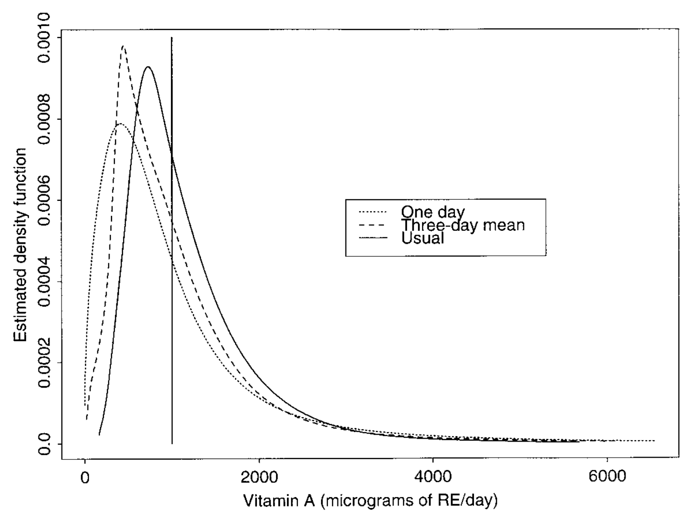

FIGURE 3 Estimated density functions of vitamin A intake on a single day, during a 3-d period, and the distribution of usual daily intakes, for the population of men age 20 y and older based on data from the Continuing Survey of Intakes by Individuals, 1989–91 (U.S. Department of Agriculture 1996). Total area under each curve is equal to 1 (100%). Area under the solid curve to the right of the vertical line estimates the proportion of men with usual intake meeting the Recommended Dietary Allowance. The analogous area under the 1-d (or 3-d) curve estimates the proportion of men with 1-d (or 3-d mean) intakes meeting the recommendation.

DISCUSSION

As stated above, our approach to estimating the usual intake distribution of a population is based on two assumptions about the survey data: *1*) individuals report their food intakes on the first day of the survey without systematic bias, and *2*) these intakes are linked correctly to the food composition database, which contains accurate quantities of particular nutrients in 100 g of each food. Although sources of error may exist to varying degrees in the estimates compiled in the food composition database, we believe that the overall error in the average nutrient content of a food is small compared with the variation in nutrient composition across foods of different types—small enough to be ignored in most cases. For example, even potential errors for folate, such as those described by Martin et al. (1990), are less important than the differences in the folate content of oranges vs. apples, beef liver vs. roast beef, or a folate-fortified cereal vs. a nonfortified cereal.

The method developed at ISU assumes that the first day of intake data is free from systematic reporting bias but that the subsequent days are not because of potential training or conditioning effects. The assumption that the first reported day has the least bias is reasonable with the CSFII data used here. Data from other sources may behave differently, however, and the optional adjustments of subsequent-day data to the first day's mean and variance may not be necessary.

Another technical assumption implicit in the method developed at ISU is that the measurement errors in the adjusted and transformed data are not correlated with the individual means. This assumption does not appear to be violated in any of the data we have investigated.

The method developed at ISU removes the within-person, or day-to-day variability of nutrient intake and addresses any skewness in the data. This within-person variation can be a combination of true variation in an individual's daily intake and random error in the reporting of intake. The method developed at Iowa State does not handle any systematic bias due to underreporting, however. To reduce the potential for such a bias, USDA has worked with researchers at the Census Bureau's Center for Survey Methods Research to investigate the cognitive aspects of the 24-h dietary recall task (DeMaio et al. 1993, Guenther et al. 1996). We have developed a multiple-pass approach to the 24-h recall that gives the respondents more opportunities to recall foods initially forgotten. Continuing research is required to improve the completeness of the reported list of foods eaten and other types of reporting error such as error in estimating the amounts of foods eaten. Such research could reduce the degree of underreporting and improve the quality of the estimates of nutrient intakes, including estimates of usual intake distributions.

Research in statistical methodology is required to develop estimates of the reliability of the estimated usual intake distributions, for example, standard errors for the percentiles. Some ratio variables are important for population assessment, for example, vitamin B-6/protein and thiamin/energy. Research must address such ratios. Researchers at ISU have proposed a method for estimating the usual intake distributions for ratios of dietary components (Carriquiry et al. 1995), but the results are preliminary. Further work is required in this area.

The problem of estimating the proportion of the population at risk for dietary inadequacy has yet to be resolved. Beaton (1993) has suggested that the proportion of the population having usual intakes below the mean requirement could be used as an estimate of the proportion at risk, at least for some nutrients. It is clear that an assessment of nutrient adequacy of a population requires reliable estimates of usual intake distributions. In addition, estimates of mean requirements are needed. They should be expressed in terms of nutrients as consumed in foods in order to correspond to available food composition databases. At present, statistical methods are available that are fit for use for estimating the proportion of the population above or below a given standard.

In this paper, we have applied the method developed at ISU to the problem of estimating usual nutrient intake distributions. The method can also be applied to estimate many other distributions of interest to nutritionists, ranging from the number of hours of television usually watched to biochemical indices of nutritional status.

ACKNOWLEDGMENTS

We gratefully acknowledge the advice and assistance of our coworkers Lori G. Borrud and Alanna J. Moshfegh, USDA Agricultural Research Service, and Kevin W. Dodd and Helen H. Jensen, Iowa State University. In addition, George H. Beaton, Christopher T. Sempos and Johanna T. Dwyer reviewed an earlier version of the manuscript and provided invaluable comments. Any remaining errors in the text are our own.

LITERATURE CITED

Anderson, S. A. (1988) Guidelines for use of dietary intake data. J. Am. Diet. Assoc. 88: 1258–1260.

Beaton, G. H. (1993) Criteria of an adequate diet. In: Modern Nutrition in Health and Disease, 8th ed. (Shils, M. E., Olson, J. A. & Shike, M., eds.), pp. 1491–1505. Lea and Febiger, Philadelphia, PA.

Beaton, G. H., Milner, J., Corey, P., McGuire, V., Cousins, M., Stewart, E., de Ramos, M., Hewitt, D., Grambsch, V., Kassim, N. & Little, J. A. (1979) Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Am. J. Clin. Nutr. 32: 2546–2559.

Carriquiry, A. L., Fuller, W. A., Goyeneche, J. J. & Dodd, K. W. (1995) Estimation of usual intake distributions of ratios of dietary components. Dietary Assessment Research Series Report 5, Staff Report 95-SR 79. Center for Agricultural and Rural Development, Iowa State University, Ames, Iowa 50011–1070.

DeMaio, T. J., Ciochetto, S. & Davis, W. L. (1993) Research on the Continuing Survey of Food Intakes by Individuals. 1993 Proceedings of the Section on Survey Research Methods, vol. II, pp. 1021–1025. American Statistical Association, Alexandria, VA.

Department. of Health and Human Services and Department of Agriculture (1993) Ten-Year Comprehensive Plan for the National Nutrition Monitoring and Related Research Program. Fed. Register 58: 32769–32770.

Department of Statistics and Center for Agricultural and Rural Development, Iowa State University (1996) SIDE [software for Intake Distribution Estimation], Version 1.0., Iowa State University, Ames, IA.

Federation of American Societies for Experimental Biology, Life Sciences Research Office. Prepared for the Interagency Board on Nutrition Monitoring and Related Research (1995) Third Report on Nutrition Monitoring in the United States: Executive Summary. U.S. Government Printing Office, Washington, DC.

Fuller, W. A., Loughin, M. M. & Baker, H. D. (1994) Regression weighting in the presence of nonresponse with application to the 1987–88 Nationwide Food Consumption Survey. Survey Methodol. 20: 75–85.

Garn, S. M., Larkin, F. A. & Cole, P. (1978) The real problem with one-day records. Am. J. Clin. Nutr. 31: 1114–1116.

Guenther, P. M. (1986) Beverages in the diets of American teenagers. J. Am. Diet. Assoc. 86: 493–499.

Guenther, P. M., Cleveland, L. C. & Ingwersen, L. A. (1997) Questionnaire development and data collection methods. In: Design and Operation: The Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, 1994–96 (Cypel, Y. C. & Tippett, K. S., eds.), NFS Report 96-1. U.S. Department of Agriculture, Agricultural Research Service, Washington, DC (in press).

Guenther, P. M. & Ricart, G. (1990) Effects of eating at food service establishments on the nutritional quality of women's diets. Topics Clin. Nutr. 4: 41–45.

Hegsted, D. M. (1972) Problems in the use and interpretation of the Recommended Dietary Allowances. Ecol. Food Nutr. 1: 255–265.

Keys, A. (1967) Dietary epidemiology. Am. J. Clin. Nutr. 20: 1151–1157.

Kott, P. S. & Guenther, P. M. (1993) Demographic characteristics of the sample, In: Evaluation of Nonresponse in the Nationwide Food Consumption Survey, 1987–88 (Guenther, P. M. & Tippett, K. S., eds.), Nationwide Food Consumption Survey Report No. M-2, pp. 6–12. U.S. Department of Agriculture, Human Nutrition Information Service, Washington, DC.

Marr, J. W. (1971) Individual dietary surveys: purposes and methods. World Rev. Nutr. Diet. 13: 105–164.

Martin, J. I., Landen, W. O., Soliman, A. M. & Eitenmiller, R. R. (1990) Application of a tri-enzyme extraction for total folate determination in foods. J. Assoc. Off. Anal.Chem. 73: 805–808.

Nusser, S. M., Carriquiry, A. L., Dodd, K. W. & Fuller, W. A. (1996) A semiparametric transformation approach to estimating usual daily intake distributions. J. Am. Stat. Assoc. 91: 1440–1449.

Nusser, S. M., Fuller, W. A. & Guenther, P. M. (1997) Estimating usual dietary intake distributions: adjusting for measurement error and nonnormality in 24-hour food intake data. In: Survey Measurement and Process Quality (Lyberg, L., Biemer, P., Collins, M., De Leeuw, E., Dippo, C., Schwarz, N. & Trewin, D., eds.). Wiley and Sons, New York, NY (in press).

Pao, E. M., Mickle, S. J. & Burk, M. C. (1985) One-day and 3-day nutrient intakes by individuals. J. Am. Diet. Assoc. 85: 313–324.

Sempos, C. T., Johnson, N. E., Smith, E. L. & Gilligan, C. (1985) Effects of intraindividual and interindividual variation in repeated dietary records. Am. J. Epidemiol. 121: 120–130.

Smith, A. F., Jobe, J. B. & Mingay, D. J. (1991) Retrieval from memory of dietary information. Appl. Cognitive Psych. 5: 269–296.

Subcommittee on the Tenth Edition of the RDAs (1989) Recommended Dietary Allowances, 10th ed. National Academy Press, Washington, DC.

Subcommittee on Criteria for Dietary Evaluation (1986) Nutrient Adequacy: Assessment Using Food Consumption Surveys. National Academy Press, Washington, DC.

U.S. Department of Agriculture (1987) Women 19–50 Years and Their Children 1–5 Years, 4 Days, Continuing Survey of Food Intakes by Individuals Report No. 85-4, pp. 164–178. U.S. Department of Agriculture, Human Nutrition Information Service, Washington, DC.

U.S. Department of Agriculture (1996) Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey 1989–91, CD-ROM, accession no. PB96-501747. National Technical Information Service, Springfield, VA.

U.S. Department of Agriculture & U.S. Department of Health and Human Services (1995) Nutrition and Your Health: Dietary Guidelines for Americans, 4th ed. Home and Garden Bull. No. 232. U.S. Government Printing Office, Washington, DC.

U.S. Department of Health and Human Services (1990), Healthy People 2000. National Health Promotion and Disease Prevention Objectives. DHHS publication no. (PHS) 91-50212, Washington, DC.

U.S. Department of Health and Human Services (1992) Recommendations for the use of folic acid to reduce the number of cases of spina bifida and other neural tube defects. Morb. Mortal. Wkly. Rep. 41/No. RR-14, September 11, pp. 1–7.