

Variation in the iodine concentrations of foods: considerations for dietary assessment^{1–4}

Alicia L Carriquiry,^{5*} Judith H Spungen,⁶ Suzanne P Murphy,⁷ Pamela R Pehrsson,⁸ Johanna T Dwyer,⁹ WenYen Juan,⁶ and Mark S Wirtz⁶

⁵Department of Statistics, Iowa State University, Ames, IA; ⁶US Food and Drug Administration, College Park, MD; ⁷University of Hawaii Cancer Center, Honolulu, HI; ⁸Nutrient Data Laboratory, Beltsville Human Nutrition Research Center, Agricultural Research Service, USDA, Beltsville, MD; and ⁹Office of Dietary Supplements, NIH, Bethesda, MD

ABSTRACT

Background: Food-composition tables typically give measured nutrient concentrations in foods as a single summary value, often the mean, without providing information as to the shape of the distribution.

Objective: Our objective was to explore how the statistical approach chosen to describe the iodine concentrations of foods affects the proportion of the population identified as having either insufficient or excessive iodine intakes.

Design: We used food intake data reported by the 2009–2010 NHANES and measured iodine concentrations of Total Diet Study (TDS) foods from 4 US regions sampled in 2004–2011. We created 4 data sets, each by using a different summary statistic (median, mean, and 10th and 90th percentiles), to represent the iodine concentration distribution of each TDS food. We estimated the iodine concentration distribution of each food consumed by NHANES participants as the 4 iodine concentration summary statistics of a similar TDS food and used these, along with NHANES food intake data, to develop 4 estimates of each participant's iodine intake on each survey day. Using the 4 estimates in turn, we calculated 4 usual iodine intakes for each sex- and age-specific subgroup. We then compared these to guideline values and developed 4 estimates of the proportions of each subgroup with deficient and excessive usual iodine intakes.

Results: In general, the distribution of iodine intakes was poorly characterized when food iodine concentrations were expressed as mean values. In addition, mean values predicted lower prevalences of iodine deficiency than did median values. For example, in women aged 19–50 y, the estimated prevalence of iodine deficiency was 25% when based on median food iodine concentrations but only 5.8% when based on mean values.

Conclusion: For nutrients such as iodine with highly variable concentrations in important food sources, we recommend that food-composition tables provide useful variability information, including the mean, SD, and median. *Am J Clin Nutr* 2016;104(Suppl): 877S–878S.

Keywords: food composition, iodine, nutrient content, statistical distribution, variability

INTRODUCTION

The fact that nutrient content often varies across different samples of the same food is well established (1). However, even though variation in nutrient content may be critical to the accuracy of dietary assessment and the results of nutritional epidemiologic studies, the topic has received little attention in the

literature. Indeed, to our knowledge, no systematic study of the variation in nutrients in foods has yet been published.

The development of meaningful summary statistics requires knowledge of the shape of the distribution. Although it is often assumed that the mean of a set of values is a reasonable summary statistic, in actuality this is true only if the underlying distribution is symmetrical around the mean and the variance is low.

As illustrated by Pehrsson et al. (2) in this supplement issue, the iodine content of foods has been found to vary over time; the causes of this variation may include seasonal effects and changes in agricultural and processing technologies that introduce iodine. In addition, there is great variability in the iodine content and bio-availability of soils from different regions and thus in foods from those regions (3). If the variability in the iodine content of a given food is nonnegligible, then how that variability is characterized affects the assessment of nutrient intake. We would like to know what summary statistics describe the most salient aspects of the distribution of iodine concentration in a given food, such that iodine intake assessments based on ≥ 1 of those summary statistics are the most useful for public health purposes.

For more than a decade, the Nutrient Data Laboratory (NDL)¹⁰ of the USDA's Agricultural Research Service has collected and

¹ Presented at the workshop "Assessment of Iodine Intake: Foods and Supplements" held by the NIH Office of Dietary Supplements in Rockville, MD, 22–23 April 2014.

² Supported in part by the Office of Dietary Supplements, NIH, through grant R01-HL091024.

³ Certain commercial equipment, instruments, materials, or software are mentioned in this article. Such mention is not intended to imply recommendation or endorsement by the Food and Drug Administration, the NIH, or the USDA, nor is it intended to imply that the materials, equipment, or software are necessarily the best available for the purpose.

⁴ Supplemental Tables 1–4 are available from the "Online Supporting Material" link in the online posting of the article and from the same link in the online table of contents at <http://ajcn.nutrition.org>.

*To whom correspondence should be addressed. E-mail: alicia@iastate.edu.

¹⁰ Abbreviations used: EAR, Estimated Average Requirement; FDA, US Food and Drug Administration; FNDDS, Food and Nutrient Database for Dietary Studies; NDL, Nutrient Data Laboratory; P10, 10th percentile; P90, 90th percentile; SR, National Nutrient Database for Standard Reference; TDS, Total Diet Study; UL, Tolerable Upper Intake Level.

First published online August 17, 2016; doi: 10.3945/ajcn.115.110353.

analyzed foods as part of the National Food and Nutrient Analysis Program. Iodine is now being added to the list of nutrients analyzed by the NDL and the US Food and Drug Administration (FDA) is working with the NDL to jointly produce an online database that can be used for estimating iodine intake from food in the US population (2). However, until the NDL data on iodine are added to the USDA's National Nutrient Database for Standard Reference (SR) and become available for use, the FDA's Total Diet Study (TDS) will remain the most comprehensive source of data on the iodine content of foods consumed by the US population.

In the present study we investigated the variability in iodine intake in US population subgroups on the basis of food consumption data reported by the NHANES and the iodine content of foods measured as part of the TDS. Because the purpose of our study was to explore whether applying different statistical approaches to describing the iodine content of foods has an impact on the assessment of usual iodine intakes from food, it was not critical to include sources of iodine other than food. We did not consider iodine-containing dietary supplements, iodine in drinking water, or the use of iodized salt at the table. Thus, the estimated usual iodine intake distributions presented herein are underestimates of the true iodine intake for a given sex- and age-specific subgroup given equivalent assumptions with regard to the iodine concentrations of foods.

METHODS

The USDA's Food and Nutrient Database for Dietary Studies (FNDDS) provides the concentrations of certain nutrients, but not iodine, for ~6200 foods. The foods recorded by NHANES participants are drawn from the FNDDS. The USDA's system of food coding, used by NHANES, is distinct from the FDA's system, used by the TDS. For our analysis of usual iodine intakes, we relied on food intake data reported by NHANES participants during the years 2009–2010 (4) and the iodine concentrations of foods collected by the TDS during the years 2004–2011 (5). To account for the iodine concentrations of the foods reported by NHANES participants, we used a file that maps each FNDDS food to a TDS food on the basis of the similarity of their ingredients (JH Spungen, unpublished data, 2014). In the following sections, we describe the NHANES food intake data, the TDS iodine concentration data, the mapping file, and our methods and procedures. A flowchart is provided in **Figure 1**.

NHANES food intake data

NHANES is a continuous, ongoing survey of the non-institutionalized, civilian resident population of the United States (6). The dietary interview component, called "What We Eat in America," is conducted as a partnership between the USDA and the CDC. Under this partnership, the CDC's National Center for Health Statistics is responsible for the survey sample design and all aspects of data collection and the USDA's Food Surveys Research Group is responsible for the dietary data collection methodology, maintenance of the databases used to code and process the data, and data review and processing. The dietary intake data are collected in 2 independent interviews and are based on recall of the day before the interview (midnight to midnight) (7).

TDS iodine concentration data

The TDS analyzes a "market basket" of ~286 foods collected 4 times/y, once from each of 4 geographic regions: North Central, West, South, and Northeast (8, 9). Since 2003–2004, iodine has been among the analytes measured (10). The analytic methods used by the TDS are described by Pehrsson et al. (2).

Mapping file linking NHANES foods to TDS foods

In the mapping file, each FNDDS food is linked to a single TDS food. Because there are only 286 TDS foods, multiple FNDDS foods are linked to the same TDS food. For example, 16 natural cheeses in the FNDDS are linked to the TDS designation for natural cheddar cheese. Similarly, 23 processed and imitation cheeses in the FNDDS are linked to the TDS designation for processed American cheese.

Analysis of temporal and regional variability in the iodine content of TDS foods

We analyzed the variability in the measured iodine concentrations of the 286 TDS foods collected quarterly between 2004 and 2011 in 4 US regions. Over this period, nominally 32 measurements ($4/y \times 8\ y$) were available for each food. We computed the CV as the SD divided by the mean. To explore the sources of variability in the iodine content of foods, we fitted a mixed linear model to the 32 measurements of iodine concentration for each food using maximum likelihood estimation. The methodology was implemented by the function "lmer" in the lme4 package of the R statistical programming language (11).

To illustrate the variability in the iodine concentrations of foods over time and over regions, we selected 8 TDS foods collected in the years 2004–2011 that exhibited both high iodine content and a high apparent variability in iodine content. These 8 foods are as follows:

- milk: skim, fluid (TDS code 4)
- cheese: cheddar, natural (TDS code 12)
- fish sticks or patties: frozen, oven-cooked (TDS code 34)
- eggs, scrambled with oil (TDS code 35)
- bread, whole-wheat (TDS code 62)
- meatloaf: beef, homemade (TDS code 148)
- cheese: Swiss, natural (TDS code 236)
- pizza: cheese and pepperoni, from pizza carry-out (TDS code 281).

Estimation of iodine intakes for NHANES 2009–2010 participants

We computed 4 summary statistics to describe the distribution of measured iodine concentrations across samples of TDS foods: 10th percentile (P10), median, mean, and 90th percentile (P90) values. Given that these summary statistics describe 32 measurements, P10 is the third smallest and P90 is the 29th highest.

We estimated the iodine concentration distribution of each FNDDS food as the 4 iodine concentration summary statistics of the TDS food to which it is linked by the mapping file. Using the 4 summary statistics thus computed and the food consumption data from NHANES 2009–2010, we developed 4 estimates of the iodine intake of each NHANES participant on each survey day. This process yielded 4 iodine intake data sets, each providing

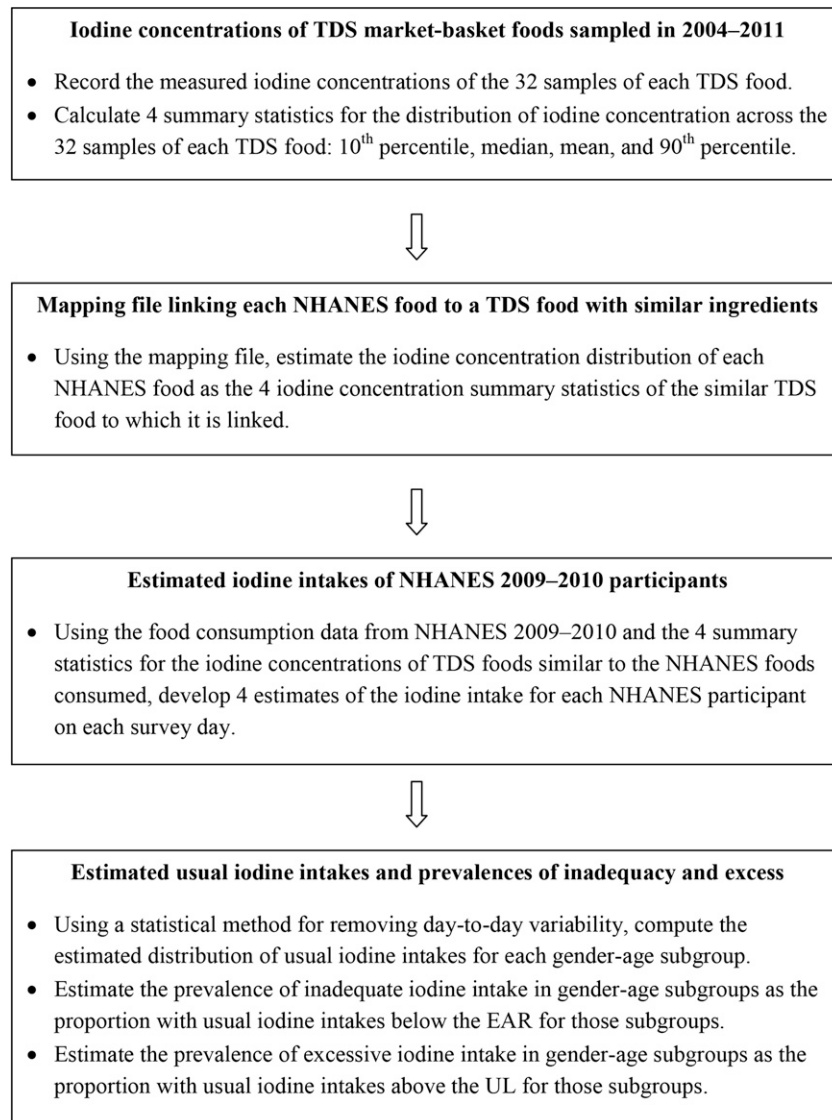


FIGURE 1 Flowchart of methods and procedures leading to the computation of usual iodine intakes and estimated prevalences of iodine deficiency and excess. EAR, Estimated Average Requirement; TDS, Total Diet Study; UL, Tolerable Upper Intake Level.

estimates of the daily iodine consumption under 1 of 4 scenarios in which the iodine concentration of each food is assumed to be its P10, median, mean, or P90 value.

Computation of estimated usual iodine intakes

We computed estimates of the distribution of usual iodine intakes based only on food for each sex- and age-specific subgroup using Personal Computer Software for Intake Distribution Estimation (PC-SIDE version 1.1; Department of Statistics and Center for Agricultural and Rural Development, Iowa State University, 2003) and the Iowa State University (ISU) method (12). Usual intakes of a nutrient, which are defined as the long-run average intakes by individuals, can be estimated as long as 2 or more daily intakes are available for at least a subsample of participants. Daily intakes are subject to both day-to-day and person-to-person variability. By removing the day-to-day variability from the daily intakes, we obtained an estimate of the usual iodine intake distribution for each subgroup in which the

tails resemble the “true” tails of usual intake. That is, the percentiles of the estimated distribution are approximately unbiased estimates of the true percentiles of usual iodine intake from food sources, and thus the proportion of individuals in the subgroup with excessive or deficient iodine intakes can be estimated in an unbiased fashion as well. Because NHANES 2009–2010 is not a self-weighting sample, we used the appropriate survey weights in all calculations. Before computing the distributions of usual intake, we removed the effect of day of the week on iodine intake. Rather than a linear adjustment (i.e., subtraction of the day effect from the observations), we used a ratio adjustment based on a regression model to ensure that adjusted observations are positive (12).

Estimation of prevalences of iodine deficiency and excess

The Institute of Medicine has derived Estimated Average Requirements (EARs) and Tolerable Upper Intake Levels (ULs) for iodine intake in age- and life stage-specific subgroups,

including pregnant and lactating women (13). However, the EARs and ULs derived for pregnant and lactating women were not used in our analysis. We included all female NHANES participants in age-categorized analytic data sets without regard to pregnancy or lactation status.

We estimated the prevalence of deficient iodine intake in a given sex- and age-specific subgroup as the proportion with usual iodine intakes below the EAR for that subgroup (14–16). Likewise, we estimated the prevalence of excessive iodine intake in a given sex- and age-specific subgroup as the proportion with usual iodine intakes above the UL for that subgroup.

RESULTS

Temporal and regional variability in the iodine content of 8 foods

Figures 2 and 3 provide summary statistics in the form of box plots for the 8 high-iodine, high-variability TDS foods selected for illustrative purposes. Figure 2 shows the iodine concentration variability in each year from 2004 to 2011; Figure 3 shows the iodine concentration variability in each of the 4 geographic regions over the same time period. In both Figures 2 and 3, the

heavy horizontal line in each box marks the median iodine concentration, the bottom and top of each box mark the 25th [quartile 1 (Q1)] and 75th (Q3) percentiles, and the whiskers extend from the lowest value that is no less than $Q1 - (1.5 \times IQR)$ to the highest value that is no greater than $Q3 + (1.5 \times IQR)$, where IQR is the difference between the 75th and the 25th percentiles.

Overall, the iodine concentrations of the 8 foods appear to vary over time and across regions. Visual inspection of Figure 2 shows that for some foods in some years (including whole-wheat bread in 2004–2007; fish sticks in 2005, 2007, and 2010; and pizza with cheese and pepperoni in 2008–2011), the high end of the distribution extends much farther than the low end (i.e., the distribution is right-skewed), implying that the mean of the distribution greatly exceeds the median and thus the mean is not very useful as a summary statistic. Likewise, visual inspection of Figure 3 shows substantial variability within and across geographic regions; in a few cases, the distribution is right-skewed (including fish sticks in region 3, Swiss cheese in regions 1 and 4, and pizza with cheese and pepperoni in region 2).

The TDS study design is balanced; that is, for each food there are the same number of measurements of iodine concentration for each year and each region. Therefore, it is straightforward to

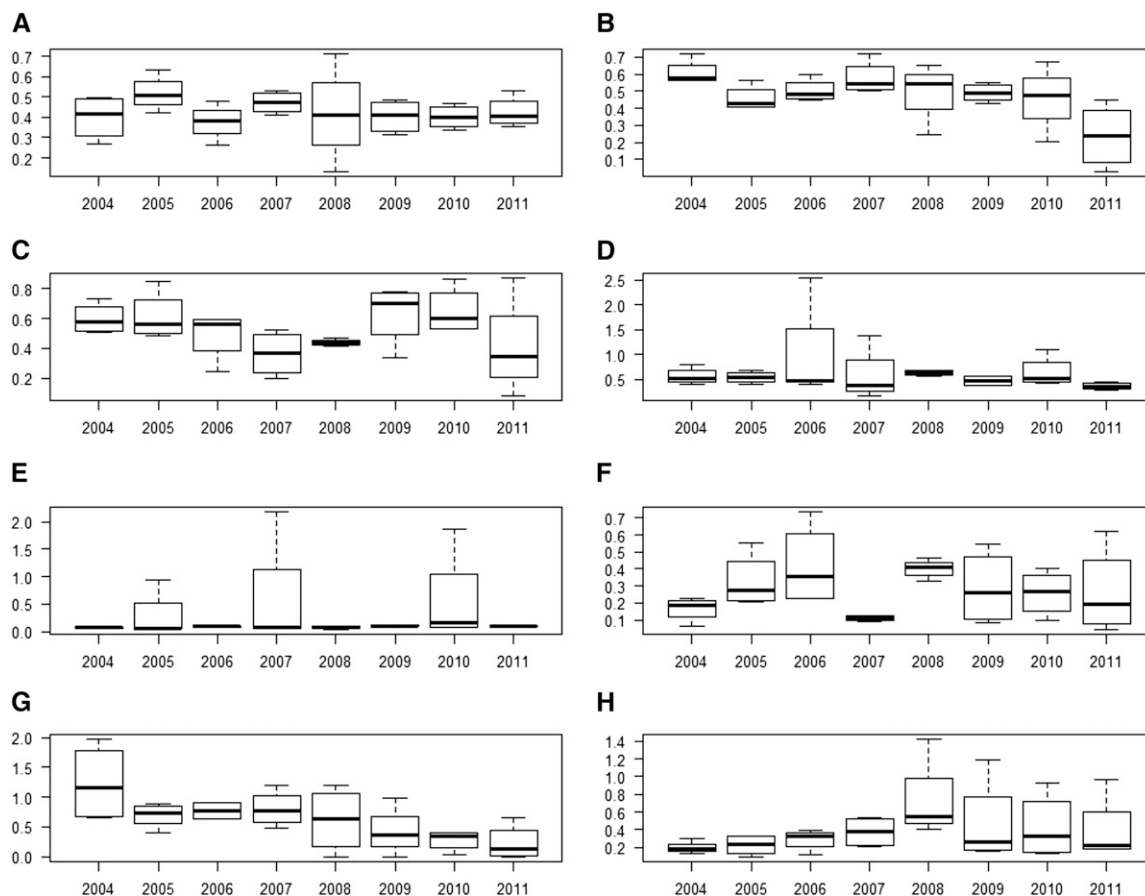


FIGURE 2 Empirical distributions of iodine concentrations, by year, for skim milk, fluid (A); whole-wheat bread (B); cheddar cheese, natural (C); meatloaf, prepared in the home (D); fish sticks or patties, frozen, oven-cooked (E); Swiss cheese, natural (F); eggs, scrambled with oil (G); and pizza with cheese and pepperoni, ordered out (H). The y axis of each panel represents the iodine concentration in units of milligrams per kilogram. The heavy horizontal line in each box is the median. The bottom of each box is drawn at the 25th percentile of the observations (Q1), and the top of each box is drawn at the 75th percentile (Q3). The whiskers extend from the lowest value that is no lower than $Q1 - (1.5 \times IQR)$ to the highest value that is no higher than $Q3 + (1.5 \times IQR)$, where the IQR is the difference between Q3 and Q1. Values are based on measured values of iodine concentrations from the Total Diet Study for foods sampled in 2004–2011 (5). Q, quartile.

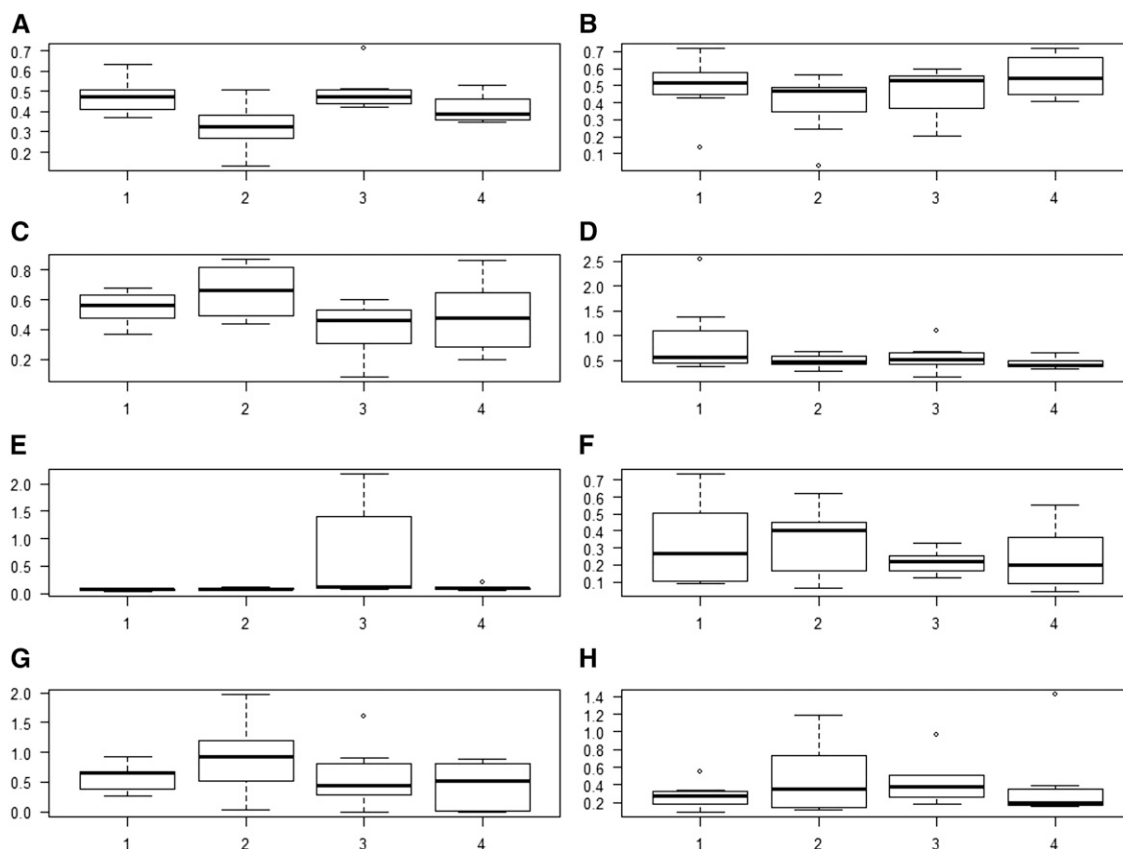


FIGURE 3 Empirical distributions of iodine content, by region, for skim milk, fluid (A); whole-wheat bread (B); cheddar cheese, natural (C); meatloaf, prepared in the home (D); fish sticks and patties, oven baked (E); Swiss cheese, natural (F); scrambled eggs with oil (G); and pizza with cheese and pepperoni, ordered out (H). The y axis of each panel represents the iodine concentration in units of milligrams per kilogram. The heavy horizontal line in each box is the median. The bottom of each box is drawn at the 25th percentile of the observations (Q1), and the top of each box is drawn at the 75th percentile (Q3). The whiskers extend from the lowest value that is no lower than $Q1 - (1.5 \times IQR)$ to the highest value that is no higher than $Q3 + (1.5 \times IQR)$, where the IQR is the difference between Q3 and Q1. Values are based on measured values of iodine concentration from the Total Diet Study for foods sampled in 2004–2011 (5). Q, quartile.

estimate the proportion of the observed variance in iodine concentration attributable to year and the proportion attributable to region. **Table 1** provides mean (SD) iodine concentrations in each of the 8 selected TDS foods. Table 1 also provides the CVs and the proportions of the total variance attributable to variation across years (Var_y) and across regions (Var_r). For each food, the proportion of the variance in iodine concentration that is due to other sources is the difference between 100% and the sum of the proportions attributable to year and region. Inspection of Table 1 shows that for the 8 foods selected, the unexplained variability in iodine concentration is larger than the variability explained by either year-to-year or region-to-region variability individually. However, the variation explained by year-to-year or region-to-region differences is substantial for some foods. For meatloaf and whole-wheat bread, ~35% of the variability is attributable to year-to-year differences; for Swiss cheese and fish sticks, the respective proportions of the variability attributable to regional differences are 34% and 25%.

It is often helpful to think of variability in terms of the CV. For some of the foods in this sample (skim milk, meatloaf, pizza with cheese and pepperoni, fish sticks, and whole-wheat bread), the CV exceeds 50%, indicating that the overall variability in iodine concentration is quite high. Indeed, for whole-wheat bread, which has a highly skewed distribution of iodine concentration and a

CV >200%, we must question whether using a single summary statistic to represent iodine concentration can be meaningful.

Although we only discuss the results obtained for the 8 foods selected, we note that several other TDS foods had higher variability in iodine concentration. Examples include canned pork and beans; salted margarine; lamb chops, pan-cooked with oil; and canned,

TABLE 1
Variability in the iodine concentrations of 8 high-iodine foods¹

| | Mean, mg/kg | SD, mg/kg | CV, % | Var _y , % | Var _r , % |
|-----------------------------|----------------|--------------|-------|----------------------|----------------------|
| Skim milk | 0.41 | 0.43 | 1.03 | 0.0 | 11.4 |
| Cheddar cheese | 0.49 | 0.18 | 0.37 | 13.1 | 0.0 |
| Fish sticks or patties | 0.53 | 0.51 | 0.97 | 0.8 | 25.4 |
| Eggs, scrambled with oil | 0.59 | 0.20 | 0.34 | 15.2 | 19.5 |
| Whole-wheat bread | 0.23 | 0.47 | 2.09 | 34.1 | 10.6 |
| Meatloaf | 0.27 | 0.16 | 0.58 | 35.4 | 9.8 |
| Swiss cheese | 0.69 | 0.11 | 0.16 | 1.8 | 38.4 |
| Pizza, cheese and pepperoni | 0.38 | 0.32 | 0.84 | 3.0 | 0.0 |

¹Values are based on measured values of iodine concentration from the Total Diet Study for foods sampled in 2004–2011 (5). Var_r, proportion of the total variance attributable to the variability across regions; Var_y, proportion of the total variance attributable to the variability across years.

refried beans. However, most foods that have high variability in iodine concentration have low mean and median concentrations, and thus their impact on dietary intake is also likely to be low.

The sole TDS seafood that we selected for analysis of region-to-region and year-to-year variability (fish sticks or patties, oven-cooked) is processed with other ingredients that could affect the iodine content. The TDS market basket includes 2 types of natural seafood:

- shrimp, boiled (TDS code 244)
- salmon: steaks/fillets, baked (TDS code 318)

On the basis of published values of the mean, SD, and median of the iodine concentrations measured in the nominally 24 samples of each food collected during the years 2006–2011 (17), these 3 types of seafood had very similar variability. For the TDS samples of fish sticks or patties, shrimp, and salmon, respectively, the CVs were 0.41, 0.53, and 0.40 and the mean-to-median ratios were 1.0, 1.2, and 1.1.

Effect of variability in food iodine concentrations on subgroup iodine intakes

Four distributions of estimated usual iodine intakes in each sex- and age-specific subgroup of NHANES 2009–2010 participants (each distribution computed by using 1 of the 4 iodine intake databases described in the Methods) are shown in **Figures 4–7**. Figure 4 shows the distributions for children aged 1–3 and 4–8 y; Figure 5 shows the distributions for boys aged 9–13 y and girls aged 9–13 y; Figure 6 shows the distributions for males aged 14–50 y and females aged 14–50 y; and Figure 7 shows the distributions for men aged ≥ 51 y and women aged

≥ 51 y. In each plot, the 4 curves correspond to the distributions computed by using the P10, median, mean, and P90 iodine concentrations of each food as the point estimates. The 2 vertical lines on each plot are drawn at the iodine EAR and UL for the corresponding sex- and age-specific subgroup.

Additional summary statistics for distributions of estimated usual iodine intakes in sex- and age-specific subgroups of NHANES 2009–2010 participants are presented in **Supplemental Tables 1–4**. Supplemental Table 1 summarizes the results when iodine intake is computed by using the P10 iodine concentration of each food as the point estimate. Similarly, Supplemental Tables 2–4 summarize the respective results when iodine intake is computed by using the median, the mean, and the P90 iodine concentrations as the point estimates. Supplemental Tables 1–4 provide, for each sex- and age-specific subgroup, the number of persons in the subgroup, 6 summary statistics describing the estimated distribution of usual iodine intakes, the EAR, and the proportion of usual intakes below the EAR.

When the third lowest measured iodine concentration of each food is taken as the point estimate (Supplemental Table 1), the prevalence of deficient iodine intake is large for individuals aged ≥ 14 y. Recall, however, that the distributions of usual iodine intakes shown in Figures 4–7 (and in Supplemental Tables 1–4) do not include iodine contributed by supplements, drinking water, or salt used at the table; therefore, we expect that these values underestimate the total iodine intake of population subgroups.

It is clear from inspection of Figures 4–7 (and Supplemental Tables 1–4) that the estimated usual iodine intake distributions vary tremendously depending on the iodine concentration summary statistic selected to represent the food consumed by survey participants. Comparison of the iodine intakes corresponding to

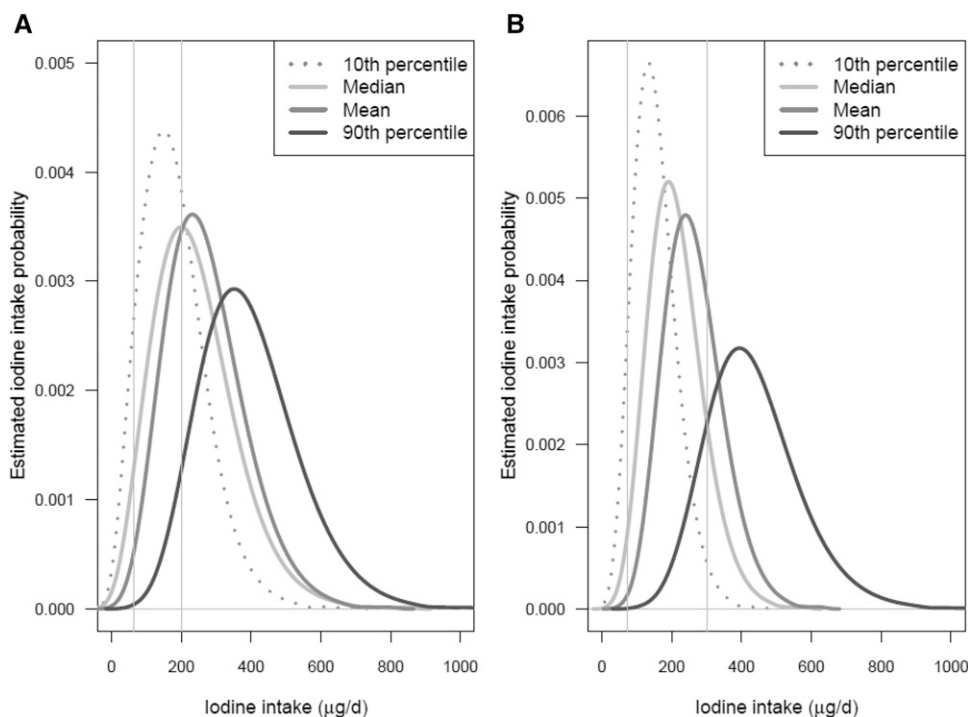


FIGURE 4 Distributions of estimated usual iodine intakes for children aged 1–3 y (A) and children aged 4–8 y (B). The 4 estimated distributions are based on the 4 different point estimates of iodine concentration shown in the key. In each graph, the left and right vertical lines are drawn at the subgroup-specific Estimated Average Requirement and the Tolerable Upper Intake Level, respectively. Values are based on food intake data reported by NHANES 2009–2010 (4) and iodine concentration measurements from the Total Diet Study for foods sampled in 2004–2011 (5).

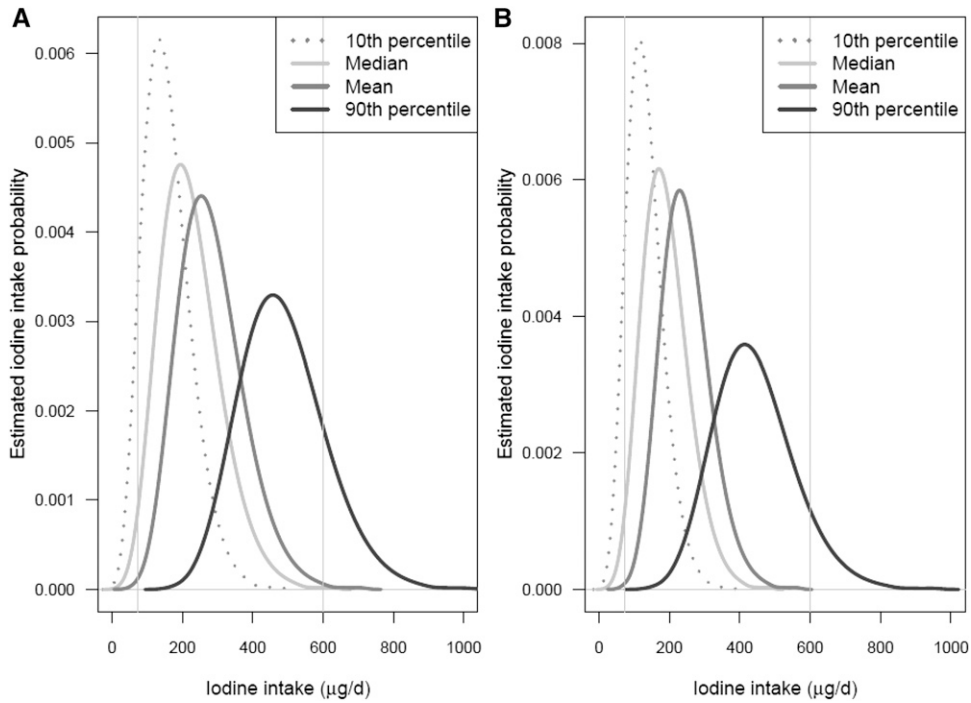


FIGURE 5 Distributions of estimated usual iodine intakes for boys (A) and girls (B) aged 9–13 y. The 4 estimated distributions are based on the 4 different point estimates of iodine concentration shown in the key. In each graph, the left and right vertical lines are drawn at the subgroup-specific Estimated Average Requirement and the Tolerable Upper Intake Level, respectively. Values are based on food intake data reported by NHANES 2009–2010 (4) and iodine concentration measurements from the Total Diet Study for foods sampled in 2004–2011 (5).

the third lowest and third highest iodine concentrations offers a dramatic illustration of what might happen if extreme nutrient concentrations were used to construct a food-composition table.

To demonstrate the impact of using different iodine concentration values for computing iodine intakes, we estimated the prevalence of potentially deficient intake (**Table 2**) as well as the

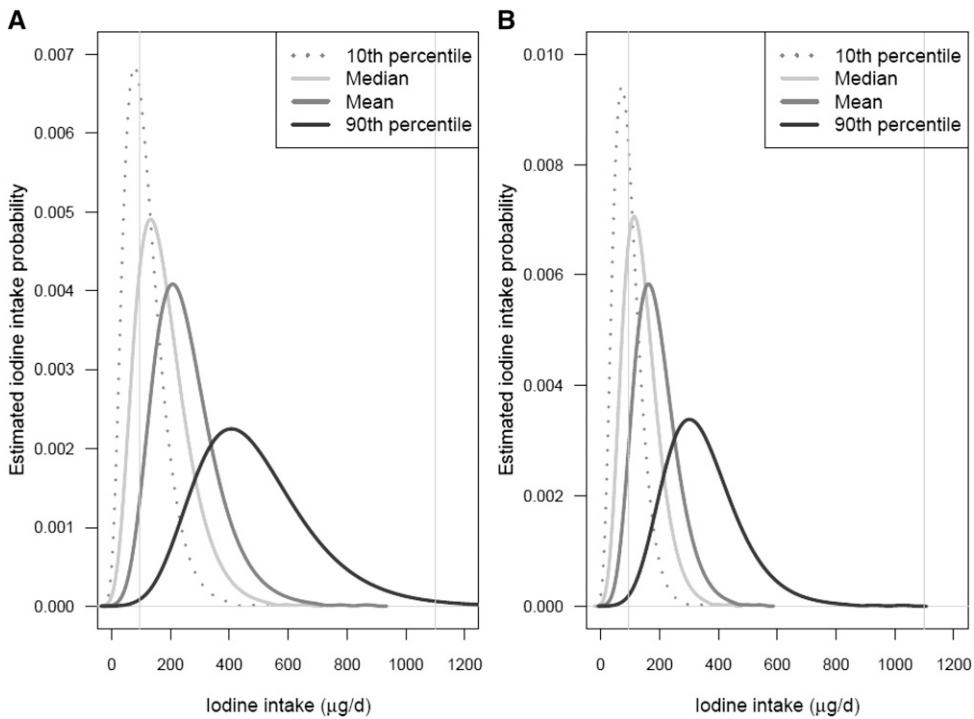


FIGURE 6 Distributions of estimated usual iodine intake distributions for males (A) and females (B) aged 14–50 y. The 4 estimated distributions are based on the 4 different point estimates of iodine concentration shown in the key. In each graph, the left and right vertical lines are drawn at the subgroup-specific Estimated Average Requirement and the Tolerable Upper Intake Level, respectively. Values are based on food intake data reported by NHANES 2009–2010 (4) and iodine concentration measurements from the Total Diet Study for foods sampled in 2004–2011 (5).

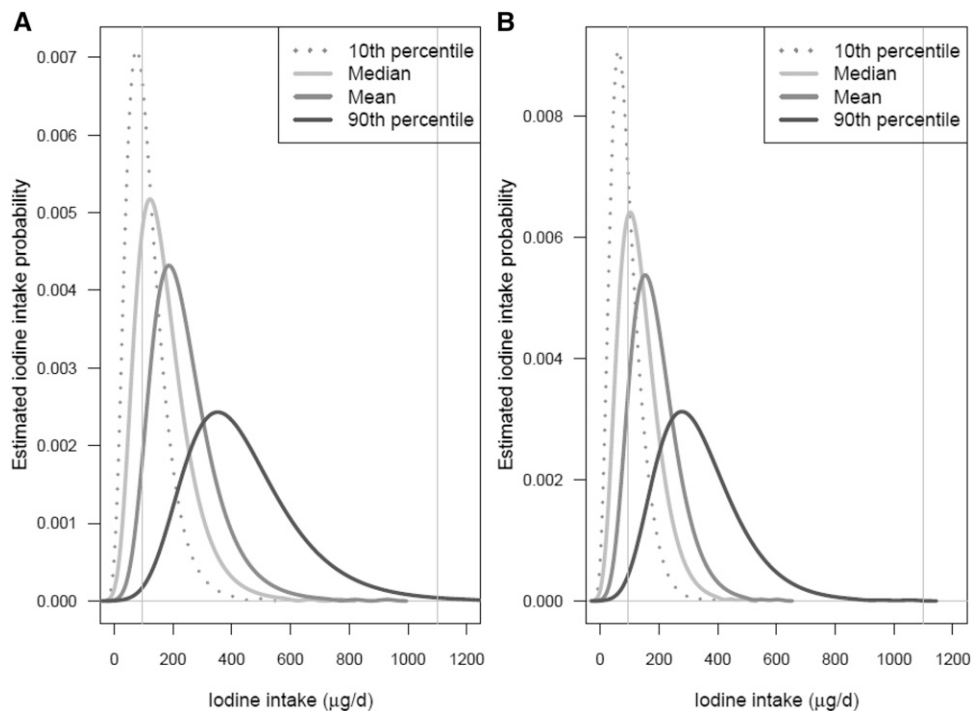


FIGURE 7 Distributions of estimated usual iodine intake distributions for men (A) and women (B) aged ≥ 51 y. The 4 estimated distributions are based on the 4 different point estimates of iodine concentration shown in the key. In each graph, the left and right vertical lines are drawn at the subgroup-specific Estimated Average Requirement and the Tolerable Upper Intake Level, respectively. Values are based on food intake data reported by NHANES 2009–2010 (4) and iodine concentration measurements from the Total Diet Study for foods sampled in 2004–2011 (5).

prevalence of potentially excessive intake (**Table 3**). The estimated prevalence of iodine deficiency is dramatically different across the columns of Table 2. As an example, for women aged 31–50 y, the prevalence of iodine intakes below the EAR is 58% when estimates are based on the P10 iodine concentration in foods and <1% when based on the P90. Similarly, large differences in the proportion of persons with intakes above the UL are observed across the columns of Table 3 for children aged 1–13 y. For example, the proportion of children aged 4–8 y with usual intakes above the UL is 2% when intake estimates are based on the P10 iodine concentration and 85% when based on the P90.

What is more significant from a practical point of view is the comparison of usual iodine intakes based on median iodine concentrations with those based on mean iodine concentrations. Because these 2 summary statistics are the most plausible contenders for use in dietary assessments, it is important to understand the consequences of adopting one over the other. On average, the usual iodine intakes based on mean iodine concentrations are shifted to the right by 30–92 $\mu\text{g}/\text{d}$ (depending on sex and age) relative to those based on median iodine concentrations, yielding markedly lower percentages of individuals with estimated iodine intakes below the relevant subgroup-specific EAR. For example, the percentage of women aged 51–70 y with iodine intakes below the EAR is 8.7% when based on the mean iodine concentration and 31.3% when based on the median. In the women overall, prevalence estimates for potential iodine deficiency that are very small when iodine intake is based on mean iodine concentrations increase to values that might be considered important from a public health standpoint when median iodine concentrations are used instead.

As can be seen from inspection of Table 3, the estimated proportion of children ≤ 8 y of age with usual iodine intakes above the UL is greater when iodine intakes are based on the mean iodine concentration than on the median. In all other sex- and age-specific subgroups, the proportion of individuals with usual iodine intakes above the UL is negligible irrespective of whether the iodine concentration summary statistic is the mean or the median.

DISCUSSION

We have shown, using the iodine concentrations of foods sampled by the TDS in 2004–2011 and food intake data collected by NHANES in 2009–2010, that the use of different iodine concentration summary statistics to determine iodine intake can have a dramatic effect on dietary assessment. The results suggest that even a change from the use of the mean iodine concentration to the median can lead to large differences in estimates of the prevalence of deficiency for most population subgroups. The reason for the large difference is that the distribution of iodine concentrations in foods appears to be nonsymmetric, at least on the basis of the 32 iodine measurements obtained by the TDS for each of 286 foods.

In right-skewed distributions, the mean is always larger than the median; the more pronounced the skewness, the larger the difference. The difference between mean and median is also a function of outliers; even a single outlier can have a large effect on the mean of a distribution but will typically have a negligible effect on the median. In the case of the iodine concentration of foods, we find that for at least some high-iodine foods the distributions are right-skewed. This suggests that the median is

TABLE 2
Estimates of the prevalence of deficient iodine intake, by subgroup¹

| | <i>n</i> | EAR, μg/d | Prevalence of iodine intake <EAR, % (SE) | | | |
|-----------------|----------|--------------|---|------------|------------|------------|
| | | | P10 | Median | Mean | P90 |
| Children | | | | | | |
| Age 1–3 y | 406 | 65 | 9.0 (2.3) | 3.5 (1.4) | 0.8 (0.6) | 0.0 (NA) |
| Age 4–8 y | 782 | 65 | 5.1 (1.8) | 0.9 (0.6) | 0.1 (0.1) | 0.0 (NA) |
| Males | | | | | | |
| Age 9–13 y | 375 | 73 | 7.7 (3.6) | 1.2 (1.3) | 0.1 (0.1) | 0.0 (NA) |
| Age 14–18 y | 356 | 95 | 23.1 (4.7) | 6.1 (3.4) | 0.3 (0.5) | 0.0 (NA) |
| Age 19–30 y | 482 | 95 | 49.0 (2.2) | 18.6 (4.2) | 3.7 (2.2) | 0.1 (0.2) |
| Age 31–50 y | 803 | 95 | 47.3 (2.2) | 18.9 (2.9) | 3.1 (1.5) | 0.1 (0.1) |
| Age 51–70 y | 795 | 95 | 52.1 (3.2) | 21.1 (2.6) | 3.8 (1.4) | 0.2 (0.2) |
| Age ≥71 y | 383 | 95 | 47.9 (3.3) | 24.7 (4.0) | 5.5 (2.9) | 0.1 (–0.2) |
| Females | | | | | | |
| Age 9–13 y | 387 | 73 | 11.3 (4.7) | 1.5 (1.6) | 0.1 (0.1) | 0.0 (NA) |
| Age 14–18 y | 327 | 95 | 52.9 (4.7) | 18.5 (6.6) | 4.0 (3.3) | 0.1 (0.2) |
| Age 19–30 y | 495 | 95 | 64.2 (4) | 25.0 (4.8) | 5.1 (3.4) | 0.0 (NA) |
| Age 31–50 y | 940 | 95 | 57.8 (2.4) | 25.1 (3.1) | 6.1 (2.2) | 0.2 (0.2) |
| Age 51–70 y | 773 | 95 | 63.4 (2.5) | 31.3 (2.7) | 8.7 (2.5) | 0.4 (0.4) |
| Age ≥71 y | 442 | 95 | 60.4 (2.9) | 32.0 (3.2) | 10.1 (3.0) | 1.0 (0.9) |

¹Estimates are based on iodine intake from food sources only and are calculated as the proportion of usual intakes below the EAR. Values are based on food intake data reported by NHANES 2009–2010 (4) and measured values of iodine concentration from the Total Diet Study for foods sampled in 2004–2011 (5). EAR, Estimated Average Requirement; NA, not available; P10, 10th percentile; P90, 90th percentile.

a more robust point estimate of the available measurements than the mean. Furthermore, there is no “statistical cost” in using the median in place of the mean because in symmetric distributions they are approximately equal.

Although the median is robust to skewness and to outliers, and is therefore a more appropriate summary of the distribution of

iodine concentrations in many of the foods in the TDS, it is not clear how best to summarize the dietary contribution of a nutrient when the variability in nutrient concentration is very large. Consider, for example, the case of whole-wheat bread. Across the 32 iodine measurements that were obtained by the TDS, the iodine content varied between 0.030 and 2.169 mg/kg, which resulted in

TABLE 3
Estimates of the prevalence of excessive iodine intake, by subgroup¹

| | <i>n</i> | UL, μg/d | Prevalence of iodine intake >UL, % (SE) | | | |
|-----------------|----------|----------|---|------------|------------|------------|
| | | | P10 | Median | Mean | P90 |
| Children | | | | | | |
| Age 1–3 y | 406 | 200.0 | 36.3 (2.7) | 58.0 (2.6) | 71.1 (3.0) | 96.0 (2.5) |
| Age 4–8 y | 782 | 300.0 | 2.0 (1.1) | 13.2 (2.8) | 29.2 (3.0) | 85.0 (3.5) |
| Males | | | | | | |
| Age 9–13 y | 375 | 600.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 17.9 (7.9) |
| Age 14–18 y | 356 | 900.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 1.3 (1.9) |
| Age 19–30 y | 482 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 1.4 (1.3) |
| Age 31–50 y | 803 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 0.9 (0.7) |
| Age 51–70 y | 795 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 1.0 (0.7) |
| Age ≥71 y | 383 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) |
| Females | | | | | | |
| Age 9–13 y | 387 | 600.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 8.0 (6.5) |
| Age 14–18 y | 327 | 900.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) |
| Age 19–30 y | 495 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) |
| Age 31–50 y | 940 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) |
| Age 51–70 y | 773 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) |
| Age ≥71 y | 442 | 1100.0 | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) | 0.0 (NA) |

¹Estimates are based on iodine intake from food sources only and are calculated as the proportion of usual intakes above the UL. Values are based on food intake data reported by NHANES 2009–2010 (4) and measured values of iodine concentration from the Total Diet Study for foods sampled in 2004–2011 (5). NA, not available; P10, 10th percentile; P90, 90th percentile; UL, Tolerable Upper Intake Level.

a CV of >200%. In such a case, the range of possible nutrient contributions is not well summarized by any single value.

The USDA's SR is the largest US food-composition database, although, as noted in the Introduction, it does not yet provide iodine concentrations. The most recent version, SR28, reports on the composition of 8789 foods. The SR28 provides the number of analyses, the SEM, the 95% CI calculated on the basis of assumptions about the distribution, and the mean, minimum, and maximum values of nutrient concentrations (18). The median value is not provided in the SR28. Thus, a user seeking a point estimate has little choice but to use the mean value. The FDA's TDS food-composition database, which reports on only 286 foods, provides the number of analyses, the SD, and the mean, median, minimum, and maximum values of nutrient concentrations (17).

At a minimum, we suggest that nutrient databases provide the number of analyses, the SD, and the mean and median values of nutrient concentrations. That would allow a considered decision to be made as to whether to select the median or the mean as a point estimate while providing some information about the underlying variability. If a single point estimate of these concentrations is to be used for public health assessment and the choice is between the mean and the median, then the results of the present investigation suggest that the median would be a more robust choice than the mean.

In cases in which having a good estimate of usual nutrient intake is needed for public health purposes, it might be reasonable to rethink how to construct a food-composition database. Ideally, food-composition databases would include, for each food, the distribution of concentrations of each nutrient. If the distribution of nutrient concentrations is available for each food, a question of interest is how we might make use of that information when calculating nutrient intakes for a sample of individuals. For example, suppose that instead of presenting a single point estimate of the nutrient concentration of each food, database developers were to present all validated, available measurements. The number of nutrient concentration measurements might vary across foods. To estimate the daily consumption of the nutrient by a sample of individuals, it would be possible to implement an approach that is reminiscent of bootstrapping, as follows:

- 1) For each person, each day, and each food, draw a random nutrient concentration from among those listed in the food-composition database. Follow each draw by replacement such that the same concentration can be drawn more than once.
- 2) Add up the contributions of the nutrient for each person and each day, thereby developing an intake database for the nutrient.
- 3) Repeat the above 2 steps multiple times.

Suppose that we repeat the first 2 steps 5 times. After implementing the approach, the analyst would have constructed 5 food intake databases for the sample. By using each database, the analyst can estimate distributions of usual nutrient intake and the prevalence of deficiency and excess. By averaging the estimated distributions of usual intake over the 5 replicated databases, the analyst accounts for the variability in nutrient concentration. Note that our mention of 5 replicates is arbitrary and used only for illustration. There is an extensive literature on methods for selecting

the number of bootstrap replicates needed to achieve a desired degree of accuracy in estimation (19, 20).

The range of estimated distributions reflects the variability in the nutrient's concentration in the foods consumed by the sample individuals. By proceeding in this manner, the issue of selecting a single summary nutrient content for each food is eliminated and instead it is possible to incorporate the inherent variability in nutrient content of foods into dietary assessments. Our results suggest that this approach would significantly improve the reliability of iodine intake assessment. It is likely that this method would likewise improve the assessment of intake for other nutrients as well. Quantifying the temporal and regional variability in nutrient concentrations in major food sources of those nutrients is a necessary first step to determining whether rethinking the structure of food-composition databases is important from a public health perspective.

We thank Christine Swanson for her important contribution to the study concept and Gay Goodman, Iodine Initiative Consultant to the NIH Office of Dietary Supplements, for expert technical editing and related contributions.

The authors' responsibilities were as follows—ALC and SPM: designed the study; JHS, PRP, and MSW: provided databases used in the study; ALC and WJ: performed the statistical analyses; ALC, JHS, SPM, PRP, JTD, and WJ: contributed written portions of the manuscript; and ALC: had responsibility for the final content. The authors reported no conflicts of interest related to the study.

REFERENCES

1. Greenfield H, Southgate DAT. Food composition data: production, management and use. New York: Springer; 2003.
2. Pehrsson PR, Patterson KY, Spungen JH, Wirtz MS, Andrews KW, Dwyer JT, Swanson CA. Iodine in food- and dietary supplement-composition databases. *Am J Clin Nutr* 2016;104(Suppl):868S–76S.
3. Mehra A, Saikat SQ, Carter JE. Bioavailability of iodine in the UK-Peak District environment and its human bioaccessibility: an assessment of the causes of historical goitre in this area. *Environ Monit Assess* 2014;186:987–99.
4. Centers for Disease Control and Prevention, National Center for Health Statistics. National Health and Nutrition Examination Survey: NHANES 2009–2010 dietary data. May 2013 [cited 2014 Sep 1]. Available from: <http://wwwn.cdc.gov/nchs/nhanes/search/DataPage.aspx?Component=Dietary&CycleBeginYear=2009>.
5. US Food and Drug Administration. Total Diet Study: analytical results [cited 2014 Oct 15]. Available from: <http://www.fda.gov/Food/FoodScienceResearch/TotalDietStudy/ucm184293.htm>.
6. Centers for Disease Control and Prevention, National Center for Health Statistics. National Health and Nutrition Examination Survey: plan and operations, 1999–2010. *Vital Health Stat* 2013;1(56). [cited 2015 Apr 6]. Available from: http://www.cdc.gov/nchs/data/series/sr_01/sr01_056.pdf.
7. Centers for Disease Control and Prevention, National Center for Health Statistics. National Health and Nutrition Examination Survey: NHANES 2009–2010, questionnaire instruments, dietary recall. Dietary interview component. [cited 2015 Apr 6]. Available from: http://www.cdc.gov/nchs/data/nhanes/nhanes_09_10/dietaryrecall_f.pdf.
8. Egan SK, Bolger PM, Carrington CD. Update of US FDA's Total Diet Study food list and diets. *J Expo Sci Environ Epidemiol* 2007;17:573–82.
9. US Food and Drug Administration. Total Diet Study: study design. 2009. Version current 10 April 2013 [cited 2014 Oct 15]. Available from: <http://www.fda.gov/Food/FoodScienceResearch/TotalDietStudy/ucm184232.htm>.
10. Murray CW, Egan SK, Kim H, Beru N, Bolger PM. US Food and Drug Administration's Total Diet Study: dietary intake of perchlorate and iodine. *J Expo Sci Environ Epidemiol* 2008;18:571–80.
11. R Development Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2011. Available from: <http://www.R-project.org/>.
12. Nusser SM, Carriquiry AL, Dodd KW, Fuller WA. A semiparametric transformation approach to estimating usual daily intake distributions. *J Am Stat Assoc* 1996;91:1440–9.

13. Food and Nutrition Board, Institute of Medicine. Dietary Reference Intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc. Washington (DC): National Academies Press; 2001.
14. Beaton GH. Criteria for an adequate diet. In: Shils ME, Olson JA, Shike M, editors. Modern nutrition in health and disease. 8th ed. Philadelphia: Lea & Febiger; 1994. p. 1491–505.
15. Carriquiry AL. Assessing the prevalence of nutrient inadequacy. *Public Health Nutr* 1999;2:23–33.
16. Food and Nutrition Board, Institute of Medicine. Dietary Reference Intakes: applications in dietary assessment. Washington (DC): National Academies Press; 2000.
17. US Food and Drug Administration. Total Diet Study. Element results summary statistics: market baskets 2006 through 2011. April 15, 2014. [cited 2015 Apr 6]. Available from: <http://www.fda.gov/downloads/Food/FoodScienceResearch/TotalDietStudy/UCM184301.pdf>.
18. USDA, Agricultural Research Service. Composition of foods: raw, processed, prepared. USDA National Nutrient Database for Standard Reference, release 28 (2015). Documentation and user guide. September 2015. [cited 2015 Oct 30]. Available from: http://www.ars.usda.gov/sp2UserFiles/Place/80400525/Data/SR/SR28/sr28_doc.pdf.
19. Hall P. The bootstrap and Edgeworth expansion. New York: Springer; 1992.
20. Chernick MR. Bootstrap methods: a guide for practitioners and researchers. 2nd ed. Hoboken (NJ): Wiley; 2008.