UC Santa Barbara
UC Santa Barbara Previously Published Works

Title

Estimating national and subnational nutrient intake distributions of global diets.

Permalink

https://escholarship.org/uc/item/7bj2w0kh

Journal

American Journal of Clinical Nutrition, 116(2)

Authors

Passarelli, Simone

Free, Christopher

Allen, Lindsay

et al.

Publication Date

DOI

Peer reviewed

Estimating national and subnational nutrient intake distributions of global diets

Simone Passarelli,[1] Christopher M Free,[2,3] Lindsay H Allen,[4] Carolina Batis,[5] Ty Beal,[6,7] Anja Pia Biltoft-Jensen,[8] Sabri Bromage,[1] Ling Cao,[9] Analí Castellanos-Gutiérrez,[5] Tue Christensen,[8] Sandra P Crispim,[10] Arnold Dekkers,[11] Karin De Ridder,[12] Selma Kronsteiner-Gicevic,[1,13] Christopher Lee,[14] Yanping Li,[1] Mourad Moursi,[15] Isabelle Moyersoen,[12] Josef Schmidhuber,[16] Alon Shepon,[1,17] Daniel F Viana,[1,18] and Christopher D Golden[1,19,20]

[1]Department of Nutrition, Harvard T. H. Chan School of Public Health, Boston, MA, USA; [2]Bren School of Environmental Science and Management, University of California, Santa Barbara, Santa Barbara, CA, USA; [3]Marine Science Institute, University of California, Santa Barbara, Santa Barbara, CA, USA; [4]ARS Western Human Nutrition Research Center, USDA, Davis, CA, USA; [5]Nutrition and Health Research Center, National Institute of Public Health, Cuernavaca, Morelos, Mexico; [6]Department of Environmental Science and Policy, University of California, Davis, Davis, CA, USA; [7]Global Alliance for Improved Nutrition, Washington, DC, USA; [8]Division of Food Technology, National Food Institute, Technical University of Denmark, Lyngby, Denmark; [9]School of Oceanography, Shanghai Jiao Tong University, Shanghai, China; [10]Department of Nutrition, Federal University of Paraná, Curitiba, Brazil; [11]Centre for Nutrition, Prevention and Health Services, National Institute for Public Health and the Environment, Bilthoven, The Netherlands; [12]Department of Epidemiology and Public Health, Sciensano, Brussels, Belgium; [13]Institute for Statistics of the Federation of Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina; [14]Harvard College, Cambridge, MA, USA; [15]Intake, Center for Dietary Assessment, FHI Solutions, Washington, DC, USA; [16]Trade and Markets Division, UN's Food and Agricultural Organization, Rome, Italy; [17]Department of Environmental Studies, The Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv, Israel; [18]Betty and Gordon Moore Center for Science, Conservation International, Arlington, VA, USA; [19]Department of Environmental Health, Harvard T. H. Chan School of Public Health, Boston, MA, USA; and [20]Department of Global Health and Population, Harvard T. H. Chan School of Public Health, Boston, MA, USA

ABSTRACT

Background: Access to high-quality dietary intake data is central to many nutrition, epidemiology, economic, environmental, and policy applications. When data on individual nutrient intakes are available, they have not been consistently disaggregated by sex and age groups, and their parameters and full distributions are often not publicly available.

Objectives: We sought to derive usual intake distributions for as many nutrients and population subgroups as possible, use these distributions to estimate nutrient intake inadequacy, compare these distributions and evaluate the implications of their shapes on the estimation of inadequacy, and make these distributions publicly available.

Methods: We compiled dietary data sets from 31 geographically diverse countries, modeled usual intake distributions for 32 micronutrients and 21 macronutrients, and disaggregated these distributions by sex and age groups. We compared the variability and skewness of the distributions and evaluated their similarity across countries, sex, and age groups. We estimated intake inadequacy for 16 nutrients based on a harmonized set of nutrient requirements and bioavailability estimates. Last, we created an R package—nutriR—to make these distributions freely available for users to apply in their own analyses.

Results: Usual intake distributions were rarely symmetric and differed widely in variability and skewness across nutrients and countries. Vitamin intake distributions were more variable and skewed and exhibited less similarity among countries than other nutrients. Inadequate intakes were high and geographically concentrated, as well as generally higher for females than males. We found that the shape of usual intake distributions strongly affects estimates of the prevalence of inadequate intakes.

Conclusions: The shape of nutrient intake distributions differs based on nutrient and subgroup and strongly influences estimates of nutrient intake inadequacy. This research represents an important contribution to the availability and application of dietary intake data for diverse subpopulations around the world. *Am J Clin Nutr* 2022;116:551–560.

Keywords: nutrient, dietary data, intake, global health, methods, subgroup, distribution, epidemiology, nutrient intake, nutrition

Introduction

Access to robust dietary intake data has been a persistent limitation in the field of nutrition research ([1]). Individual-level dietary data sets are scarce because of the large effort and expense

involved in data collection. It is especially difficult to find data sets describing an individual's usual intake, which requires >1 d of dietary recall to calculate. Even when these data are collected, the resulting data sets are often proprietary. Data availability may also be limited to a mean or median level of intake of a given nutrient in a country rather than a full population distribution. Altogether, these barriers inhibit our understanding of nutritional vulnerabilities of countries and their subpopulations (2).

Consequently, researchers must often rely on coarser food supply or household expenditure data (1). Although these sources provide invaluable information about trends in the average supply of nutrients, they mask important differences in food allocation patterns among sex and age subgroups (3, 4) and are insufficient for estimating the burden of disease risk and dietary inadequacies or excesses. For this, researchers need information on the distribution of usual intakes within sex and age subgroups to assess the proportion of subpopulations at risk (5).

In recent years, there have been significant advances in access to individual-level dietary intake data (6–9). For example, the Global Dietary Database (GDD) project has compiled nutrition data from >1600 surveys (10), representing one of the greatest contributions to democratizing access to nutrition data to date. Still, ~14% of these data sets are not representative at the national or subnational level, and about a quarter are not publicly available.

Dietary data from 24-h recalls or food diaries are the best available data for population-level dietary surveillance because they estimate absolute intake with greater precision (11). However, 1 d of recall does not represent usual intake (long-term intake free from day-to-day variation) (12, 13). Collecting a second day of intake in at least a subsample makes it possible to differentiate within- and between-person variation and use statistical procedures to estimate usual intake distributions that exhibit only the between-person variation. Although other methods have been developed for estimating usual intake without multiple days (14, 15), these approaches have limitations for accurate global estimation, especially in low-income populations in which many foods are not consumed daily (16). Furthermore, these data are not consistently disaggregated by sex and age groups, although there is ample evidence showing that consumption patterns differ across dietary factors, subgroups, and countries (17–19).

The objective of this study was to fill these gaps by harnessing data from geographically and demographically diverse surveys.

Using individual dietary intake data from food consumption surveys incorporating ≥2 recall days, we calculated usual nutrient intake distributions for age and sex subpopulations in 31 countries for 32 micronutrients and 21 macronutrients. We then found the best-fitting distribution shape to describe the intake patterns for each sex–age group within a country. We compared these distributions to highlight differences in intake patterns across country, sex, and age groups based on their mean, variance, and skewness and to show how shape is associated with inadequate intakes. To make these data and distributions accessible and applicable for future research, we developed an R package called "nutriR" as an accompaniment to this article.

Methods

Diet recall data

Most data were obtained from publicly available sources, including the GDD (10) and the FAO/Global Individual Food consumption data Tool website (20) (**Figure 1**; **Supplemental Table 1**; **Supplemental Figure 1**). For some countries, including the Netherlands, Brazil, Belgium, Denmark, China, and Mexico, data were made available upon request. We reviewed data sets to determine their eligibility for inclusion in our analysis, based on whether they had *1*) individual-level dietary data, *2*) calculated nutrient-level data, *3*) ≥2 d of dietary intake (for at least some participants), *4*) data based on a 24-h recall or diet record/food diary, and *5*) a sample size >200. If there was >1 nationally or regionally representative survey for a given country, we selected the most recent data set. To our knowledge, we included all publicly available data sets meeting our parameters at the time of this analysis.

Deriving usual intake distributions

In general, ≥2 d of dietary recall data are required to estimate the proportion of total variance that is attributable to within-individual variation (16). Although a recently developed National Cancer Institute methodology uses only 1 d of dietary information to estimate usual intake distributions (15), this approach faces several limitations in the context of global assessments. First, it requires that foods are consumed almost daily, which is often not the case in low-income countries for foods such as fruit, meat, fish, and eggs. Second, it is highly sensitive to the variance ratio selected, and large variations in these ratios have been observed across populations (16). At least some members of a population should have ≥2 d of dietary recall data to generate an unbiased estimate of usual intake, especially when populations have different demographic characteristics (16). Because one of the purposes of our analysis was to estimate usual intake distributions among diverse populations and subgroups, we chose to only include data sets with a minimum of 2 d of dietary data for at least some members of the population.

We used the Statistical Program to Assess usual Dietary Exposure (SPADE, v4.1.0) to estimate usual intakes of 21 macronutrients and 32 micronutrients from the available survey data (**Supplemental Table 2**) (20). SPADE is an R-based software package that models usual intake distributions based on repeat 24-h dietary recall data by *1*) transforming data to a normal distribution, *2*) removing within-person variability, *3*) estimating usual nutrient intake distributions as a function of age,

Inclusion criteria
Datasets had:
- Individual-level dietary data
- Calculated nutrient-level data
- Two or more days of intake (for at least some respondents)
- A 24-hour recall or diet record/food diary format
- Sample size >200
- Variables for age and sex

GDD
Canada, $n = 20,483$
Ecuador, $n = 650$
Estonia, $n = 3049$
Ethiopia, $n = 323$
Germany, $n = 921$
Guatemala, $n = 226$
Kenya, $n = 718$
Mozambique, $n = 543$
Pakistan, $n = 270$
Portugal, $n = 6424$
Sweden, $n = 4288$

FAO/WHO GIFT
Bangladesh, $n = 475$
Bulgaria, $n = 1723$
Burkina Faso, $n = 960$
Democratic Republic of Congo, $n = 214$
India, $n = 242$
Italy, $n = 3323$
Lao PDR, $n = 2045$
Malaysia, $n = 573$
Philippines, $n = 1203$
Romania, $n = 1382$
Uganda, $n = 551$
Zambia, $n = 771$

NHANES
United States, $n = 7483$

Individual Research Partners
Belgium, $n = 3146$
Bosnia and Herzegovina, $n = 853$
Brazil, $n = 46,164$
China, $n = 10,197$
Denmark, $n = 3806$
Mexico, $n = 4,343$
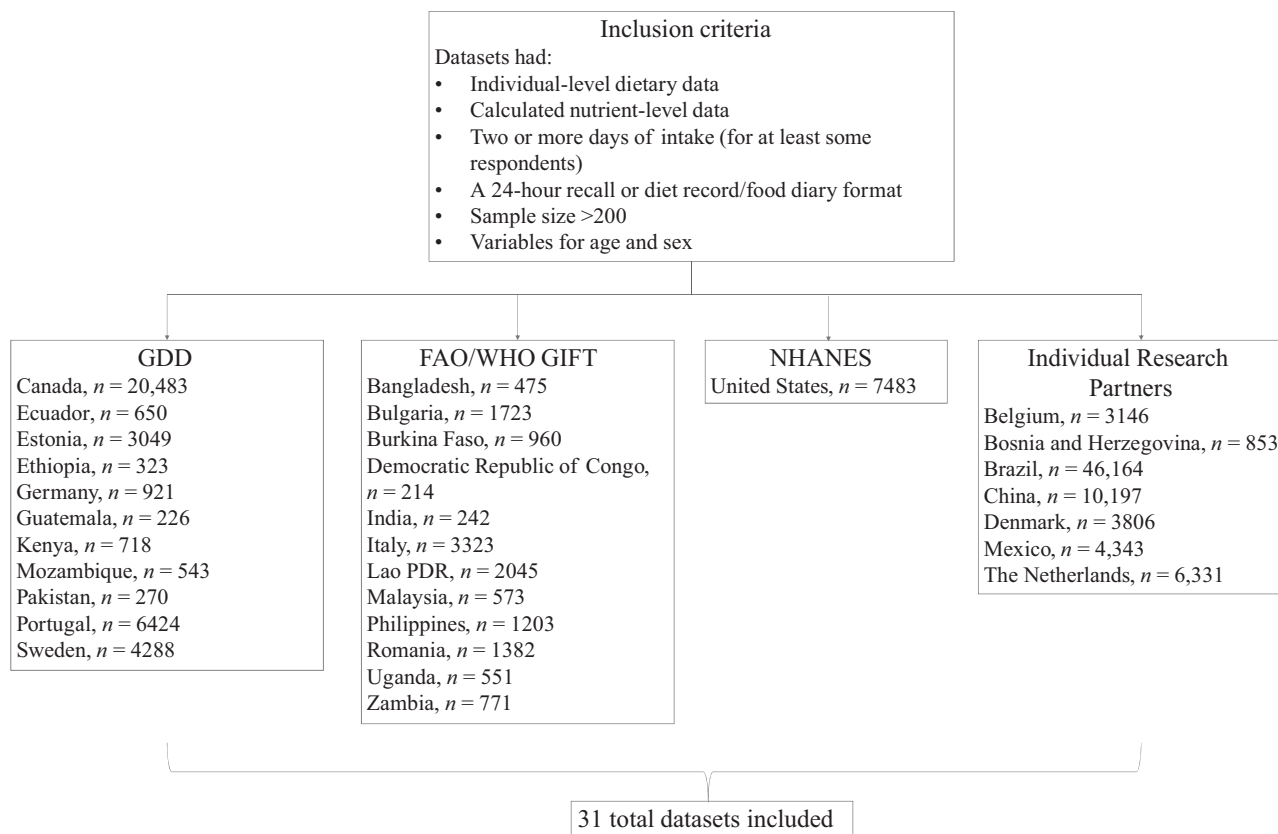The Netherlands, $n = 6,331$

31 total datasets included

FIGURE 1 Flowchart of data sets included in this analysis. For the full references and survey details, please see Supplemental Table 1. The n listed refers to the sample size of participants in each survey who had at least 1 d of recall. GDD, Global Dietary Database; GIFT, Global Individual Food consumption data Tool.

and *4*) back-transforming the data onto its original scale (21, 22)). We chose to use this statistical model instead of the raw dietary data (or the mean intakes across multiple days) because without removing within-person variation, the variance of the intake distribution would be higher than the true variance. As a result, this can lead to highly biased estimates of the proportion of the population above or below a reference value, especially when a reference is closer to the tails of a distribution (12).

The calculation of inadequate intakes using the probability method (see details below) requires that intake distributions be described as probability density functions. We therefore used the *fitdistrplus* R package v1.1-6 (23) (in R v4.1.1) to fit gamma and log-normal distributions to the usual intakes of all available sex–age groups (**Supplemental Figures 2** and **3**). We defined age groups in 5-y intervals from 0–5 y to 95–99 y to align with the GDD (2) and other common demographic data sets. We selected the distribution with the best Kolmogorov–Smirnov goodness-of-fit statistic as the final distribution for each group (**Supplemental Figure 4**). We considered gamma and log-normal distributions because they are continuous probability distributions bounded between zero and positive infinity that accommodate the skewed (asymmetric) shapes common among these types of data.

Describing and comparing usual intake distributions

We quantified and compared the shape of the fitted usual intake distributions using 2 metrics: *1*) the CV, which measures the relative variability of a distribution, and *2*) the skewness, which measures the asymmetry of a distribution. Both metrics are comparable across scales and distribution types. Although we also calculated excess kurtosis—which measures the tailedness of a distribution—we omitted these results from the article because of their correlation with skewness; however, the kurtosis measures are included in the "nutriR" data package described below. Using the equations in **Supplemental Table 3**, we calculated these metrics for each usual intake distribution. These quantities were derived from the fitted distributions, rather than from the distributions of the pseudo-population modeled through the SPADE analysis, to maintain consistency throughout the analysis workflow. If distributions had extremely low variability (CV <0.01; **Supplemental Figure 5**), they were considered unrepresentative of population-wide usual intakes and omitted from the analysis.

To compare the similarity of usual intake distributions among countries, sex, and age groups, we measured the percent overlap between pairwise combinations of distributions within groups. Percent overlap between paired distributions was assessed using the Bhattacharyya coefficient (21), in which a value of 0% indicates entirely divergent distributions and a value of 100% indicates perfectly identical distributions. We evaluated the similarity in distributions among countries as the median percent overlap of all unique pairwise combinations of countries within each sex–age group. To evaluate similarity in distributions between sexes, we considered the median percent overlap in each age group across all countries. To assess similarity in distributions
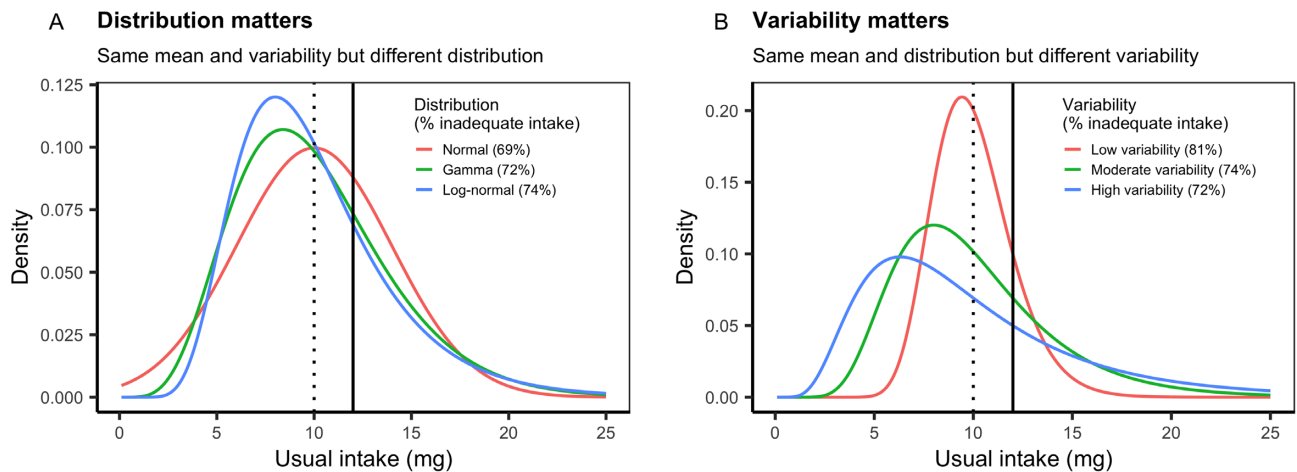
FIGURE 2 Stylized usual nutrient intake distributions illustrating the importance of intake distribution shape in determining the prevalence of inadequate nutrient intake. (A) All 3 distributions have the same mean ($\mu = 10$ mg; vertical dotted line) and CV (0.4) but are described by different probability distributions. (B) All 3 distributions are log-normal with the same mean ($\mu = 10$ mg; dotted vertical line) but different levels of variability (CV: low = 0.2, moderate = 0.4, high = 0.6). Percentages indicate the percentage of the population estimated to have inadequate nutrient intakes given an Average Requirement (solid vertical line) of 12 mg with a CV of 0.10 and calculated using the probability method (25).

among age groups, we calculated the median percent overlap of all unique pairwise combinations of age groups within each country–sex group.

Estimating prevalence of inadequate intakes

To evaluate the implications of the shape of usual intake distributions for achieving nutrient adequacy, we related metrics of distribution shape to the estimated prevalence of inadequate intakes. We estimated population-level intake inadequacy, also known as the summary exposure value (24), for 16 nutrients with estimated nutrient requirements using both the probability method (25) and the Estimated Average Requirement (EAR) cut-point method (hereafter referred to as the cut-point method) (26). The probability method compares intake distributions with a continuous relative risk curve, whereas the cut-point method is a simplification of the probability method and assumes that the prevalence of inadequacy of a certain population can be estimated by assessing the fraction whose intake is below a certain cut-point value (the EAR). The cut-point method does not require knowledge of the shape of the intake and requirement distributions, only that these distributions adhere to a set of criteria. These include that the shape of the requirement distribution be symmetrical around the EAR and that the variance of the intake distribution be larger than the variance of the requirement distribution (5, 26).

In both methods for estimating intake inadequacy, we used the Average Requirements (ARs) provided by Allen et al. (27) as the estimated nutrient requirements (**Supplemental Figure 6**). Allen et al. (27) reviewed alternative sources of nutrient reference values and identified which were most up-to-date and most appropriate for use in global-scale studies. We selected ARs for iron and zinc based on their bioavailability in national diets. We used the Human Development Index, a summary measure of average achievement in key dimensions of human development (28), as an indicator of likely dietary availability (**Supplemental Table 4**). Although average nutrient requirements are available for vitamin D, they assume no sun exposure, so we do not calculate inadequate vitamin D intakes here.

The continuous risk curves of the probabilistic approach have a value of 1 at low intakes, 0.5 at the relevant EAR, and zero at large intakes. These absolute risk curves are based on the cumulative normal distribution function of requirements (26) with a mean at the EAR and a CV of 0.25 for the AR for vitamin B-12 and a CV of 0.10 for the AR of all other nutrients (29).

The "nutriR" R package

We developed an open-access R package called "nutriR" to share our data and facilitate its use by other public health researchers. This package contains a series of functions that allow the user flexibility in retrieving and plotting distributions, shifting a distribution around a new mean (e.g., in response to an intervention or based on a user's own data set), estimating inadequacy of nutrient intakes based on ARs, and characterizing and comparing the shapes of distributions across populations using descriptive statistical parameters. The R package, user manual, and an interactive R Shiny web application for exploring the fitted subnational nutrient intake distributions are all available on GitHub (https://github.com/cfree14/nutriR).

Results

Our data compilation resulted in individual-level dietary intake data with calculated nutrients from 31 countries and 5 continents (Supplemental Figure 1; Supplemental Table 1). The methods of recall included a mix of food diaries and 24-h recalls, generally based on the literacy of the surveyed populations. The year of data collection varied by country, from as early as 2002 in the case of the Philippines to as recently as 2018 in the United States. Most data sets (23 of 31) employed a 24-h recall methodology. The level of representativeness varied, with 18 data sets representative at the national level and the remaining representative at a lower administrative level.

Our results show that there are large differences in the shape of usual intake distributions across nutrients, countries, sexes, and age groups and that these differences have profound impacts on
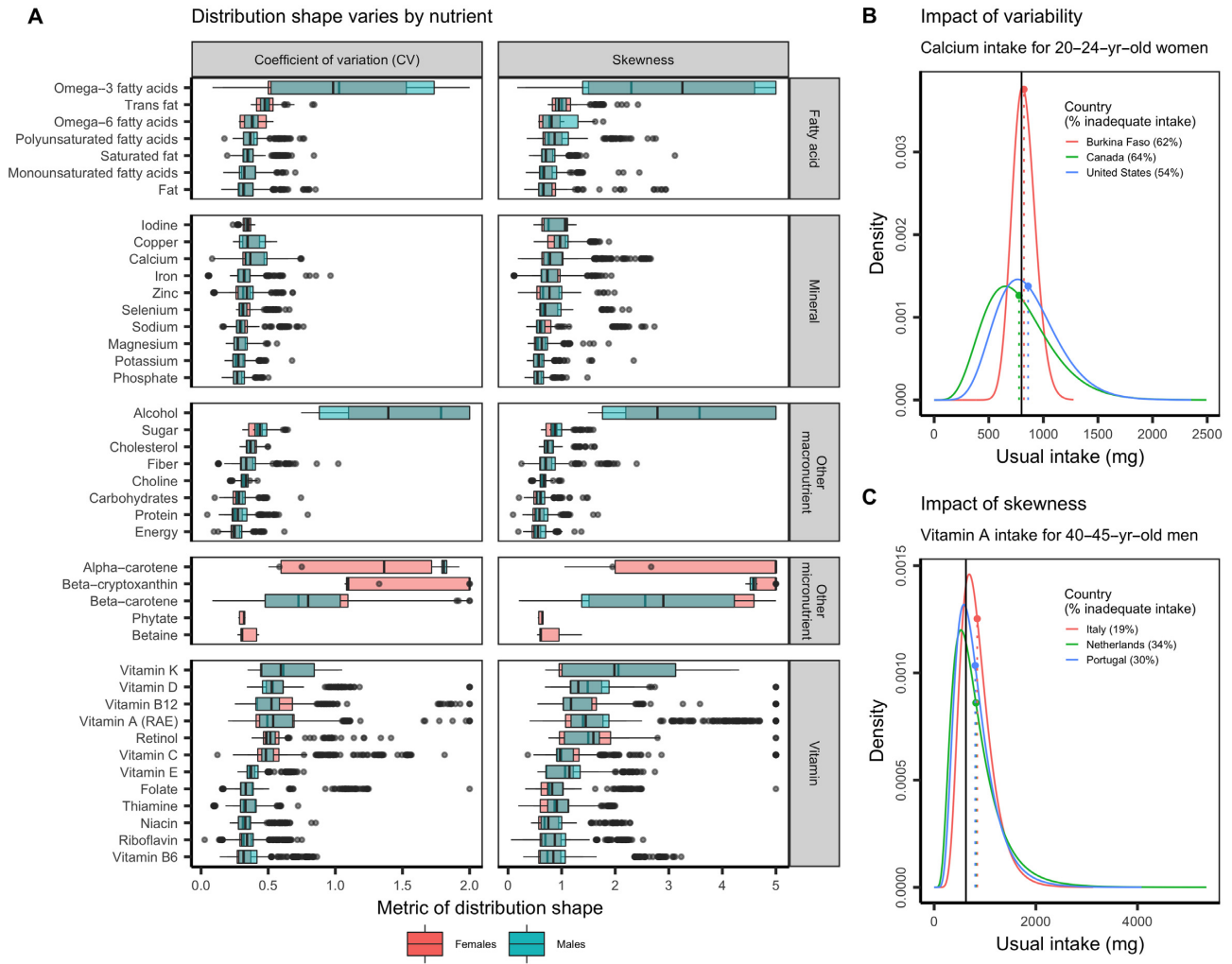
FIGURE 3 (A) The CV and skewness of usual intake distributions by nutrient. Country–sex–age group representation varies among nutrients, and only nutrients with data from ≥3 countries are shown. In the boxplots, the solid line indicates the median, the box indicates the interquartile range (IQR; 25th and 75th percentiles), the whiskers indicate 1.5 times the IQR, and the points beyond the whiskers indicate outliers. The side panels use selected distributions to illustrate the impact of the (B) variability and (C) skewness of intake distributions on the prevalence of inadequate nutrient intakes. (B) Twenty- to 24-y-old women in the selected countries exhibit usual calcium intakes with similar means but differing levels of variability. (C) Forty- to 45-y-old men in the selected countries exhibit usual vitamin A intakes with similar means but differing levels of skewness. In both panels, the dotted vertical lines indicate the mean usual intakes, the solid vertical lines indicate the Average Requirements, and the percentages indicate the prevalence of inadequate nutrient intakes within each subpopulation using the probability approach (25). RAE, retinol activity equivalents.

the prevalence of inadequate nutrient intakes. **Figure 2** provides a stylized example to illustrate the concepts of nutrient intake distributions explored in our analysis. Figure 2A shows how the type of distribution (normal, gamma, or log-normally distributed) can result in different estimates of inadequate intake, even when the mean intake is identical. Figure 2B similarly shows this concept but for the underlying assumptions about the variability of a distribution. These hypothetical examples illustrate why distributions matter when estimating usual nutrient intake for a population.

Intake distributions vary in shape

The shape of usual intake distributions varied widely across nutrients and between sexes. We compared usual intake distributions in terms of their variability, measured as their CV,

and their asymmetry, measured as their skewness (**Figure 3**A). Although usual intakes of minerals such as phosphate, potassium, and magnesium generally exhibited low variability (CV <0.5) and asymmetry (skewness <1.0), usual intakes of vitamins such as vitamins K, D, B-12, A, and C generally exhibited high variability and asymmetry (Figure 3A). Vitamin intake distributions also exhibited a wider variety of shapes than mineral distributions and most fatty acid and other macronutrient intake distributions (Figure 3A). The exceptions were the intake distributions for omega-3 fatty acids and β-carotene, which were highly diverse, variable, and skewed (Figure 3A), reflecting subpopulations dominated by very low usual intakes but with rare instances of very high usual intakes. Figure 3B,C illustrates how increasing the variability and skewness, respectively, of usual intake distributions affects distribution shape and estimates of intake inadequacy.
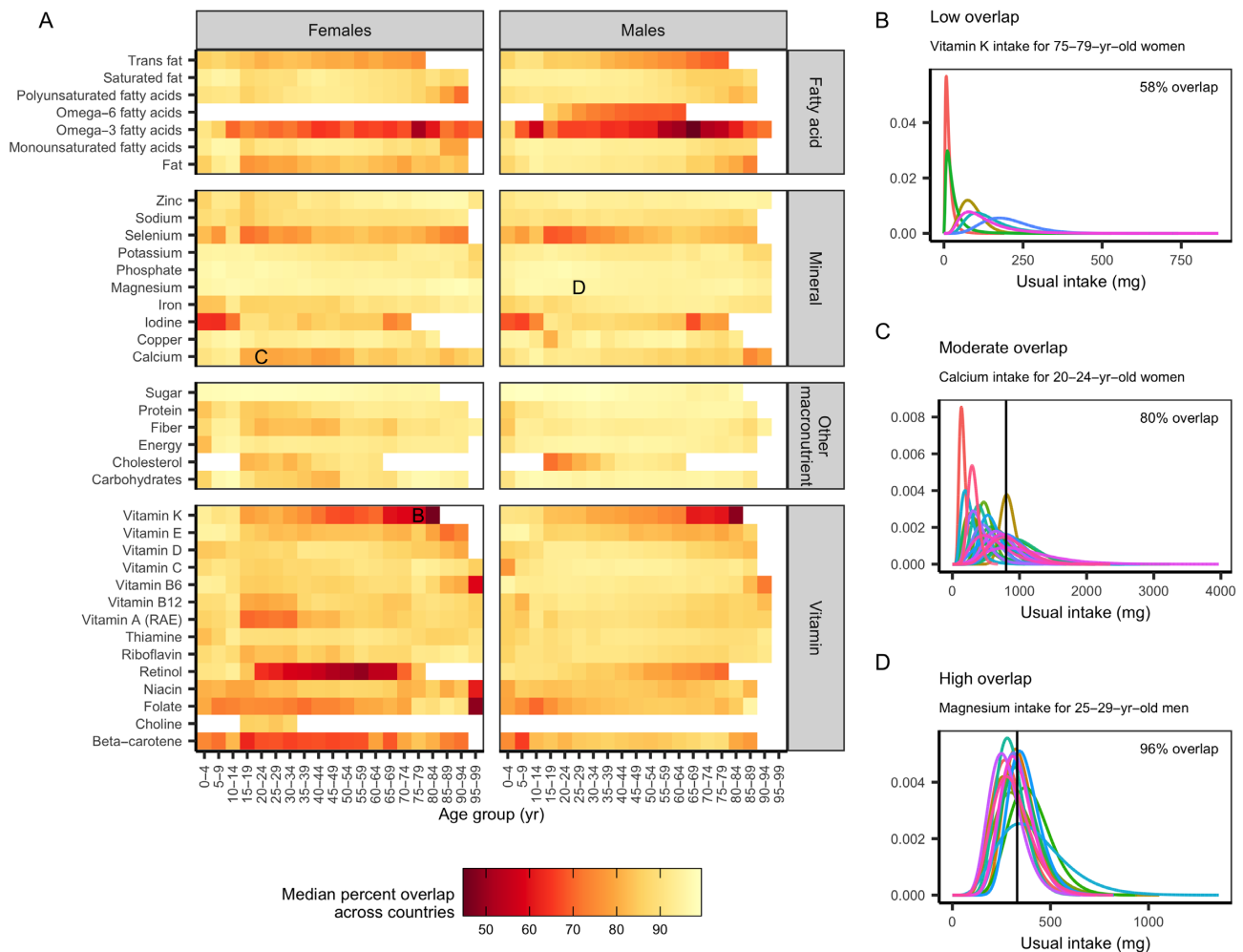
FIGURE 4   (A) The similarity in usual nutrient intake distributions across countries. Colors indicate the median percent overlap of all pairwise combinations of intake distributions from countries within each sex–age group. Country–sex–age group representation varies among nutrients, and only nutrients with data from ≥3 countries are shown. Lower overlap values indicate larger differences in nutrient distributions among countries, and higher overlap values indicate smaller differences in nutrient distributions among countries. The side panels illustrate examples of (B) low, (C) moderate, and (D) high overlap in usual intake distributions among countries. The colored lines represent usual intake distributions for different countries, solid vertical lines indicate the Average Requirements (if available), and percentages indicate the mean percent overlap. RAE, retinol activity equivalents.

The similarity in intake distributions varies by nutrient, country, sex, and age

The similarity of usual intake distributions varied more by country and sex than by age group. We measured the similarity of pairs of distributions as their percent overlap and the similarity of groups of distributions as the median percent overlap of all unique pairs (**Figure 4**A). The usual intake distributions of vitamin K, retinol, ω-3 fatty acids, and folate exhibited the least similarity across countries (see Figure 4B using vitamin K as an example). On the other hand, intakes of minerals—especially magnesium, phosphorus, and potassium—were very similar across countries (see Figure 4D using magnesium as an example). In general, intakes of the remaining micronutrients exhibited moderate similarity across countries (see Figure 4C using calcium as an example). Usual intakes of niacin; macronutrients such as protein, energy, and alcohol; and minerals such as selenium, sodium, zinc, and phosphorus were especially different between

sexes within countries and age groups (**Supplemental Figure 7**), whereas intakes of vitamins C, D, E, and K were generally similar between sexes (Supplemental Figure 7). Usual nutrient intake distributions were, in comparison, more similar among age groups within country and sex groups (**Supplemental Figure 8**).

Prevalence of inadequate intake by nutrient and country

We measured and compared the prevalence of inadequate intakes for the 16 nutrients with harmonized ARs (27) using both the probability method and the cut-point method (5, 25, 26). The probability method accounts for the shape of intake and requirement distributions by comparing actual intake distributions with a relative risk curve of requirement. The cut-point method simplifies the probability method by assuming that the fraction of the population's intake below a cut-point can
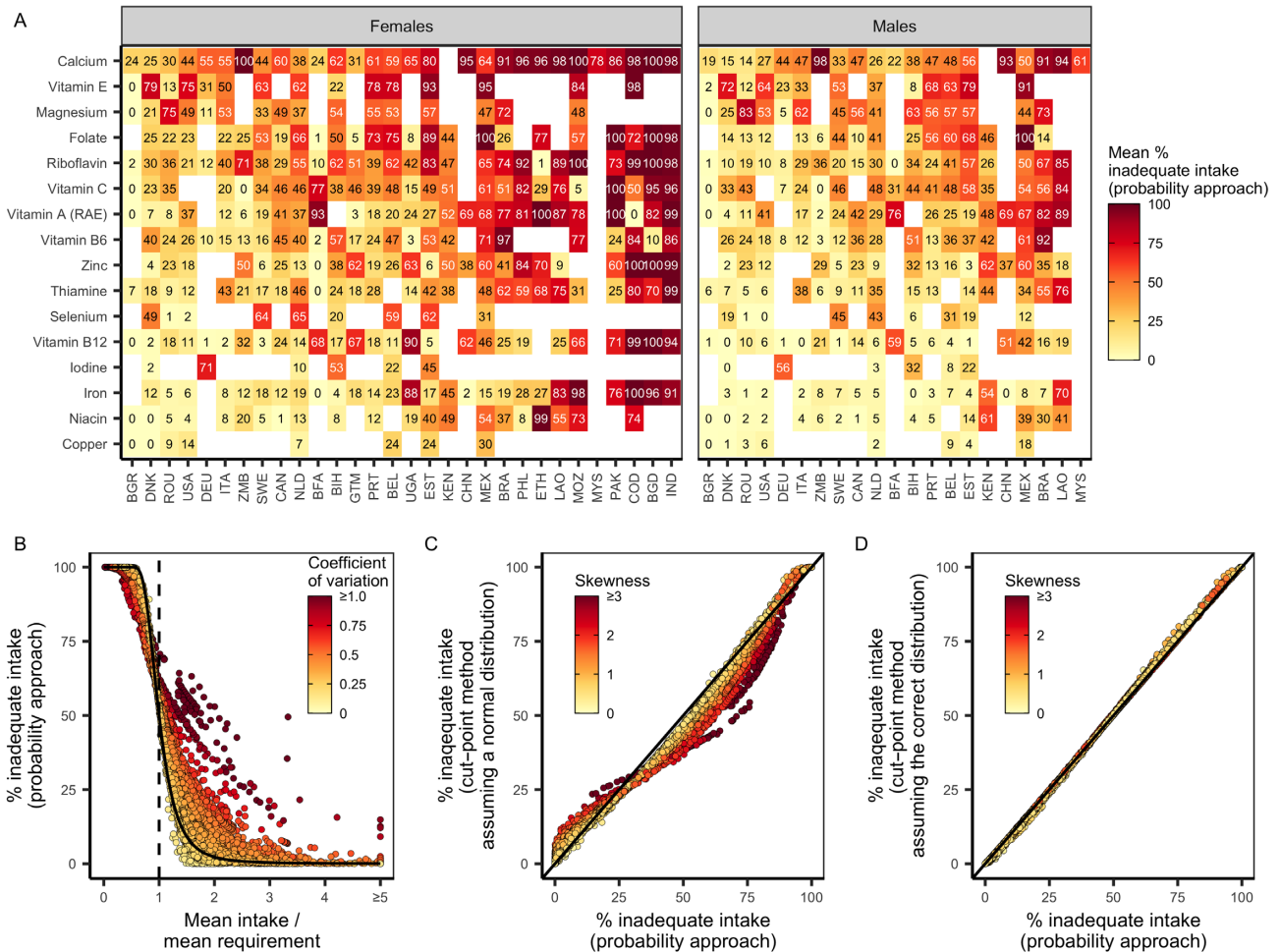
FIGURE 5 The prevalence of intake inadequacy based on (A) country, nutrient, and sex; (B) the mean and variability of usual nutrient intake distributions; and (C, D) the method for calculating intake inadequacy and its assumptions regarding the symmetry of intake distributions. In (A) and (B), intake inadequacies were calculated using the probability method. In (A), color and numbers indicate the mean percent nutrient intake inadequacy across age groups within a country. Nutrients and countries are ordered by severity of intake inadequacy. In (B) and (C), points represent usual nutrient intake distributions for each nutrient–country–sex–age group. In (B), the dashed vertical line indicates mean intakes equivalent to mean requirements. To further illustrate the impact of usual intake distribution shape on intake inadequacy, the solid curve indicates, for reference, the intake inadequacy for a normally distributed intake distribution with a CV of 0.25. In (C) and (D), points compare the prevalence of inadequate intakes estimated using the probability method and the cut-point method when assuming (C) normally distributed usual intake distributions and (D) the correct usual intake distributions. RAE, retinol activity equivalents.

determine the prevalence of inadequacy if various criteria about the requirement and intake distributions are met.

Using the probability method (25), we generated estimates of subnational nutrient intake inadequacy for the 31 evaluated data sets. There was acute and geographically widespread intake inadequacy for calcium, folate, vitamin E, and magnesium (**Figure 5**A). We found moderate and geographically dependent intake inadequacies for vitamin A (retinol activity equivalents), vitamin C, thiamin, zinc, vitamin B-12, riboflavin, and vitamin B-6 (Figure 5A). Copper, iodine, niacin, and iron exhibited comparably lower intake inadequacies. (Figure 5A). Across nutrients, we observed a high prevalence of intake inadequacy in Bangladesh, Laos, Ethiopia, the Philippines, China, and Mexico and a comparably lower prevalence of intake inadequacy in Denmark, Romania, the United States, Canada, Italy, and Sweden (Figure 5A).

The shape of intake distributions influences estimates of intake inadequacy

The shape of usual intake distributions strongly determines the prevalence of inadequate intakes estimated using the probability method (Figure 5B). Subpopulations with usual intake distributions that have identical means but different levels of variability experienced different levels of intake inadequacy (Figure 5B). When mean intakes were below the average requirement, high variability was associated with a lower prevalence of intake inadequacy, but when intakes were above the average requirement, higher variability was associated with dramatically higher intake inadequacies (Figure 5B).

The cut-point method, which makes some simplifying assumptions regarding the intake and requirement distributions, estimated levels of inadequate intake similar to the probability

method when properly accounting for the asymmetry of usual intake distributions (Figure 5D). However, when assuming symmetric usual intake distributions, the cut-point method estimated different levels of intake inadequacy relative to the probability method (Figure 5C). Differences were mostly between 0% and 25%, but in some extreme cases, differences of 50% were observed. When mean intakes were below the average requirement, the cut-point method (assuming normally distributed usual intakes) produced higher estimates of intake inadequacy compared with the probability method and painted a gloomier picture of population adequacy (Figure 5B,C). On the other hand, when mean intakes were above the average requirement, the cut-point method (assuming normally distributed usual intakes) estimated lower levels of intake inadequacy relative to the probability method and painted a rosier picture of population adequacy (Figure 5B,C).

Discussion

Although previous work has assumed similar distribution shapes across nutrients, sexes, ages, and countries, our findings show that there is notable variation in shapes across these subgroups. The degree of variation differs across nutrient types, with nutrients such as long-chain ω-3 fatty acids, vitamin K, retinol, β-carotene, and selenium exhibiting the least similarity across countries. We find evidence of high levels of inadequate intake across a number of nutrients, but more so for females compared with males.

Our analysis unpacks the relation between distribution shapes and the prevalence of inadequacy. When mean intake is below the average requirement, higher variability is associated with a lower prevalence of intake inadequacy, but when mean intake is above the average requirement, higher variability is associated with dramatically higher intake inadequacy. This variability could mean different things in different contexts; where nutrients are scarce, variability may indicate that only a few fortunate individuals achieve adequate diets, but in a context of plenty, large deviations from the mean could result from highly unbalanced and insufficient diets.

When meeting certain criteria, the cut-point method, which has been widely used in the literature due to its lower data requirements compared with the probability method, can calculate the prevalence of inadequate intake relatively accurately (5, 30–32). Yet in our analysis, the cut-point method gave higher or lower estimates of the prevalence of inadequate intakes compared with the probability approach when failing to account for distribution skewness. This trend is supported by other studies (31), but a solution has been limited by the lack of detailed information on nutrient-specific distribution shapes (26). Oftentimes, the true shape of the intake distribution is unknown and a symmetrical distribution is assumed for ease. Moreover, the cut-point method's assumption that the nutrient intake and requirement distributions are independent of each other is likely violated in many populations, especially for macronutrients (5, 26, 33), which could lead to biased estimates of inadequate intake. Calculations using the cut-point method in conjunction with real (skewed) distributions may result in more precise estimates of nutrient inadequacy.

Our findings related to the high variation in ω-3 fatty acids and selenium are consistent with previous research showing the geographical dependency of these nutrients (34, 35). The concentrations of selenium in soil and water are sensitive to environmental and anthropogenic conditions; thus, populations living in highly localized food systems are more vulnerable to deficiency and toxicity (36). Vitamin K deficiency receives little focus globally because it is thought to be rare (37); the divergent patterns of inadequate intakes observed in our findings should be further explored. A better understanding of the distributions of retinol and carotenoids has direct implications for estimating inadequacy and determining the effects of fortification and supplementation strategies. Vitamin A deficiency is widespread and affects nearly 30% of children under 5 y (38).

Inadequate iron intake was particularly prevalent among females in South and Southeast Asia, sub-Saharan Africa, and Ecuador. In some cases, our estimates of inadequate iron intake were lower than estimates of the prevalence of iron deficiency based on biomarkers, perhaps because iron deficiency biomarkers like serum ferritin may better account for iron bioavailability (39–41). But in populations with a high prevalence of inflammation and infection, estimates of iron deficiency based on anemia prevalence are overestimates, due to the multiple determinants of anemia in these settings (42). Depending on the context, estimates of inadequate iron intake from diet may either underestimate or overestimate deficiency. For example, in Bangladesh, iron deficiency is lower than intakes suggest due to the high amounts of iron in groundwater (43), but in the United States, iron deficiency is higher than what intake estimates show (44).

Inaccuracies in the estimation of intake distributions could have substantial implications for public health; underestimation of inadequate intake could result in a failure to identify and intervene on important nutrient needs, whereas overestimation could lead to an inefficient allocation of resources. Accurate methods for determining inadequate intake are critical to inform policies related to agriculture, fortification, nutrition, and dietary guidelines—especially in low- and middle-income countries, where inadequacies are most highly concentrated (45).

There are numerous applications of these distributions for nutritional epidemiology and public health. These include, for example, estimating the impacts of an intervention on nutrient intakes, assessing the proportion of populations and subpopulations at risk of nutrient deficiency or overload, understanding the consequences of distributions for nutrition-related disease risk, targeting nutrition interventions to specific sex or age groups, population-level modeling with microsimulations, forecasting future nutrient gaps, and planning fortification and supplementation policies and programs. Researchers could also use these data to further characterize patterns of covariance, such as by examining patterns of covariance between nutrients to see if the presence of one deficiency might predict the presence of another, and by understanding covariance across age–sex groups and populations. A better understanding of these trends can help to shape recommendations of dietary guidelines and patterns, to determine the need for supplementation and fortification, and to improve intervention targeting.

These findings also influence our understanding of important questions in nutrition-related research. First, these methods could help identify subgroups or nutrients that are most in need of

context-specific data or intake distributions. Because multiple days of recall data are still rare, future research could help to understand the extent to which other data sources (e.g., food supply data or consumption and expenditure surveys) can provide valid estimates of usual intake distribution shapes. Researchers could apply the methods presented here to assess the implications of assumptions about distribution shapes for the Global Burden of Disease study estimates (46). Last, future analyses that link these distributions to micronutrient biomarkers could improve our understanding of the links between nutrient intake and risk of deficiency with greater precision.

This analysis has several limitations. There are substantial differences in dietary nutrient intakes within countries—for example, due to region, crop composition, or soil factors (47)—that could result in distributions being unrepresentative of some populations. The representativeness of our underlying data sets also varies. Users should to refer to Supplemental Table 1 when drawing conclusions about representativeness. Our geographical coverage was subject to data availability. The dietary recall methods used varied across surveys, which could influence estimations of intake (48, 49). Differences in the underlying food composition databases and methodologies could reduce the external validity for comparisons across countries (50). The number of repeated measures available for each country could differentially affect precision, particularly for nutrients that have a large ratio of within- to between-person variability (32, 51). Although multiday dietary recalls measure "usual intakes," they insufficiently capture natural fluctuations in diets due to holidays, seasons, and so on (45). The inclusion of fortification in this analysis may vary based on the underlying data used for each country. Last, because we focus on food consumption, micronutrient intake from supplements is excluded. Based on these limitations, we ask users of these data and findings to take caution when extrapolating our results to other populations or across countries.

In conclusion, our article highlights how distribution shape matters when assessing a population's usual nutrient intake. To our knowledge, this is the largest compilation of nutrient intake distributions generated for public use. Our inclusion of the results in the accompanying R package ensures that these methods can be directly applied by researchers. Future work should incorporate greater nuance into assumptions about distribution shapes to more accurately describe population intakes so that research, policy, and programmatic approaches can better meet nutritional needs.

Data Availability

All code analyzing raw data will be available on a public GitHub page upon publication at https://github.com/cg0lden/subnational-distributions-extended. All raw data sets that we are allowed to share will also be available for download on that page. Code and data pertaining to the modeled distributions are available on GitHub here: https://github.com/cfree14/subnational_nutrient_distributions. Last, the distributions are publicly available through the R package that has been produced as an output of this research, which is free and available for download at https://github.com/cfree14/nutriR/.

References

1. Micha R, Coates J, Leclercq C, Charrondiere UR, Mozaffarian D. Global dietary surveillance: data gaps and challenges. Food Nutr Bull 2018;39:175–205.
2. Miller V, Singh GM, Onopa J, Reedy J, Shi P, Zhang J, et al. Global Dietary Database 2017: data availability and gaps on 54 major foods, beverages and nutrients among 5.6 million children and adults from 1220 surveys worldwide. BMJ Global Health 2021;6:e003585.
3. Del Gobbo LC, Khatibzadeh S, Imamura F, Micha R, Shi P, Smith M, et al. Assessing global dietary habits: a comparison of national estimates from the FAO and the Global Dietary Database. Am J Clin Nutr 2015;101:1038–46.
4. Murphy S, Ruel M, Carriquiry A. Should Household Consumption and Expenditures Surveys (HCES) be used for nutritional assessment and planning? Food Nutr Bull 2012;33:S235–41.
5. Carriquiry AL. Assessing the prevalence of nutrient inadequacy. Public Health Nutr 1999;2:23–34.
6. FAO. FAO/WHO GIFT: Global Individual Food consumption data Tool [Internet]. 2021 [cited September 1 2020]. Available from: http://www.fao.org/gift-individual-food-consumption/data-and-indicator/en/
7. INDDEX Project. Data4Diets: building blocks for diet-related food security analysis. Boston (MA): Tufts University; 2018.
8. Schmidhuber J, Sur P, Fay K, Huntley B, Salama J, Lee A, et al. The Global Nutrient Database: availability of macronutrients and micronutrients in 195 countries from 1980 to 2013. Lancet Planet Health 2018;2(8):e353–8.
9. Smith MR, Micha R, Golden CD, Mozaffarian D, Myers SS. Global Expanded Nutrient Supply (GENuS) model: a new method for estimating the global dietary supply of nutrients. PLoS One 2016;11:e0146976.
10. Microdata Surveys. Global Dietary Database [Internet]. 2019 [cited September 1 2020]. Available from: https://www.globaldietarydatabase.org/management/microdata-surveys
11. Thompson FE, Kirkpatrick SI, Subar AF, Reedy J, Schap TE, Wilson MM, et al. The National Cancer Institute's dietary assessment primer: a resource for diet research. J Acad Nutr Diet 2015;115:1986–95.
12. Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. J Am Diet Assoc 2006;106:1640–50.
13. Hoffmann K, Boeing H, Dufour A, Volatier JL, Telman J, Virtanen M, et al. Estimating the distribution of usual dietary intake by short-term measurements. Eur J Clin Nutr 2002;56:S53–S62.
14. Nusser SM, Carriquiry AL, Dodd KW, Fuller WA. A semiparametric transformation approach to estimating usual daily intake distributions. J. Am. Stat. Assoc 1996;91:1440–49. DOI: 10.1080/01621459.1996.10476712.
15. Luo H, Dodd KW, Arnold CD, Engle-Stone R. A new statistical method for estimating usual intakes of nearly-daily consumed foods and nutrients through use of only one 24-hour dietary recall. J Nutr 2019;149:1667–73.
16. French CD, Arsenault JE, Arnold CD, Haile D, Luo H, Dodd KW, et al. Within-person variation in nutrient intakes across populations and settings: implications for the use of external estimates in modeling usual nutrient intake distributions. Adv Nutr 2021;12(2):429–51.
17. Bezerra IN, Goldman J, Rhodes DG, Hoy MK, de Moura Souza A, Chester DN, et al. Difference in adult food group intake by sex and age groups comparing Brazil and United States nationwide surveys. Nutr J 2014;13:74.
18. Ratsavong K, van Elsacker T, Doungvichit D, Siengsounthone L, Kounnavong S, Essink D. Are dietary intake and nutritional status influenced by gender? The pattern of dietary intake in Lao PDR: a developing country. Nutr J 2020;19:31.

19. Jahns L, Carriquiry A, Arab L, Mroz TA, Popkin BM. Within- and between-person variation in nutrient intakes of Russian and U.S. children differs by sex and age. J Nutr 2004;134(11):3114–20.

20. FAO/WHO GIFT. Data and indicators [Internet] [cited 2021 May 4]. Available from: http://www.fao.org/gift-individual-food-consumption/data-and-indicator/en/

21. Bhattacharyya A. On a measure of divergence between two multinomial populations. Sankhyā Indian J Stat 1946;7:401–6.

22. Dekkers ALM, Verkaik-Kloosterman J, van Rossum CTM, Ocké MC. SPADE, a New Statistical Program to Estimate Habitual Dietary Intake from Multiple Food Sources and Dietary Supplements. J Nutr 2014;144(12):2083–91.

23. Delignette-Muller ML, Dutang C. fitdistrplus: an R package for fitting distributions. J Stat Softw 2015;64(4):1–34.

24. Forouzanfar MH, Afshin A, Alexander LT, Anderson HR, Bhutta ZA, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388:1659–724.

25. National Research Council (US) Subcommittee on Criteria for Dietary Evaluation. The probability approach. nutrient adequacy: assessment using food consumption surveys. [Internet] Washington (DC): National Academies Press; 1986 [cited 2022 Mar 30]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK217531/

26. Institute of Medicine (US) Subcommittee on Interpretation and Uses of Dietary Reference Intakes, Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. Using the Estimated Average Requirement for nutrient assessment of groups. DRI Dietary Reference Intakes: applications in dietary assessment [Internet]. Washington (DC): National Academies Press; 2000 [cited 2021 Oct 18]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK222898/

27. Allen LH, Carriquiry AL, Murphy SP. Perspective: proposed harmonized nutrient reference values for populations. Adv Nutr 2020;11:469–83.

28. UNDP. The next frontier: human development and the Anthropocene. New York (NY): United Nations Development Programme; 2020.

29. Food and Nutrition Board, National Academy of Sciences, Institute of Medicine. Dietary reference intakes: Estimated Average Requirements and recommended intakes [Internet]. 2020 [cited October 1 2020]. Available from: https://www.nal.usda.gov/sites/default/files/fnic_uploads//recommended_intakes_individuals.pdf

30. Institute of Medicine (US) Subcommittee on Interpretation and Uses of Dietary Reference Intakes, Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. Assessing the Performance of the EAR cut-point method for estimating prevalence. DRI Dietary Reference Intakes: applications in dietary assessment. Washington (DC): National Academies Press; 2000.

31. Lauzon B, Volatier JL, Martin A. A Monte Carlo simulation to validate the EAR cut-point method for assessing the prevalence of nutrient inadequacy at the population level. Public Health Nutr 2004;7:893–900.

32. Carriquiry AL. Estimation of usual intake distributions of nutrients and foods. J Nutr 2003;133:601S–8S.

33. Murphy SP, Vorster HH. Methods for using nutrient intake values (NIVs) to assess or plan nutrient intakes. Food Nutr Bull 2007;28:S51–S60.

34. Jones GD, Droz B, Greve P, Gottschalk P, Poffet D, McGrath SP, et al. Selenium deficiency risk predicted to increase under future climate change. Proc Natl Acad Sci U S A 2017;114:2848–53.

35. Hicks CC, Cohen PJ, Graham NAJ, Nash KL, Allison EH, D'Lima C, et al. Harnessing global fisheries to tackle micronutrient deficiencies. Nature 2019;574:95–8.

36. Fordyce FM. Selenium deficiency and toxicity in the environment. In: Selinus O, editor. Essentials of medical geology. Dordrecht: Springer Netherlands; 2013. p. 375–416.

37. Napolitano M, Mariani G, Lapecorella M. Hereditary combined deficiency of the vitamin K-dependent clotting factors. Orphanet J Rare Dis 2010;5:21.

38. Wirth JP, Petry N, Tanumihardjo SA, Rogers LM, McLean E, Greig A, et al. Vitamin A supplementation programs and country-level evidence of vitamin A deficiency. Nutrients 2017;9:190.

39. Gardner W, Kassebaum N. Global, regional, and national prevalence of anemia and its causes in 204 countries and territories, 1990–2019. Curr Dev Nutr 2020;4:830.

40. Safiri S, Kolahi A-A, Noori M, Nejadghaderi SA, Karamzad N, Bragazzi NL, et al. Burden of anemia and its underlying causes in 204 countries and territories, 1990–2019: results from the Global Burden of Disease Study 2019. J Hematol Oncol 2021;14:185.

41. Stoltzfus RJ. Iron deficiency: global prevalence and consequences. Food Nutr Bull 2003;24:S99–103.

42. Petry N, Olofin I, Hurrell RF, Boy E, Wirth JP, Moursi M, et al. The proportion of anemia associated with iron deficiency in low, medium, and high human development index countries: a systematic analysis of national surveys. Nutrients 2016;8:693.

43. Merrill RD, Shamim AA, Ali H, Jahan N, Labrique AB, Schulze K, et al. Iron status of women is associated with the iron concentration of potable groundwater in rural Bangladesh. J Nutr 2011;141:944–9.

44. Mei Z, Addo OY, Jefferds ME, Sharma AJ, Flores-Ayala RC, Brittenham GM. Physiologically based serum ferritin thresholds for iron deficiency in children and non-pregnant women: a US National Health and Nutrition Examination Surveys (NHANES) serial cross-sectional study. Lancet Haematol 2021;8:e572–82.

45. Gibson RS, Charrondiere UR, Bell W. Measurement errors in dietary assessment using self-reported 24-hour recalls in low-income countries and strategies for their prevention. Adv Nutr 2017;8:980–91.

46. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019;393:1958–72.

47. Gashu D, Nalivata PC, Amede T, Ander EL, Bailey EH, Botoman L, et al. The nutritional quality of cereals varies geospatially in Ethiopia and Malawi. Nature 2021;594:71–6.

48. De Keyzer W, Huybrechts I, De Vriendt V, Vandevijvere S, Slimani N, Van Oyen H, et al. Repeated 24-hour recalls versus dietary records for estimating nutrient intakes in a national food consumption survey. Food Nutr Res 2011;55:10.3402/fnr.v55i0.7307.

49. Bingham S. Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. Int J Epidemiol 1997;26:137S–51S.

50. Van Puyvelde H, Perez-Cornago A, Casagrande C, Nicolas G, Versele V, Skeie G, et al. Comparing calculated nutrient intakes using different food composition databases: results from the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Nutrients 2020;12:2906.

51. Nelson M, Black AE, Morris JA, Cole TJ. Between- and within-subject variation in nutrient intake from infancy to old age: estimating the number of days required to rank dietary intakes with desired precision. Am J Clin Nutr 1989;50:155–67.