Presentation and interpretation of food intake data: Factors affecting comparability across studies

Mieke Faber Ph.D. a,*, Friede A.M. Wenhold Ph.D. b, Una E. Maclntyre Ph.D. b, Edelweiss Wentzel-Viljoen Ph.D. c, Nelia P. Steyn Ph.D. d, Wilna H. Oldewage-Theron Ph.D. e

a Nutritional Intervention Research Unit, Medical Research Council, South Africa
b Department of Human Nutrition, University of Pretoria, South Africa
c Centre of Excellence for Nutrition, North-West University, South Africa
d Population Health, Health Systems and Innovation, Human Sciences Research Council, South Africa
e Vaal University of Technology, South Africa

Abstract

Non-uniform, unclear, or incomplete presentation of food intake data limits interpretation, usefulness, and comparisons across studies. In this contribution, we discuss factors affecting uniform reporting of food intake across studies. The amount of food eaten can be reported as mean portion size, number of servings or total amount of food consumed per day; the absolute intake value for the specific study depends on the denominator used because food intake data can be presented as per capita intake or for consumers only. To identify the foods mostly consumed, foods are reported and ranked according to total number of times consumed, number of consumers, total intake, or nutrient contribution by individual foods or food groups. Presentation of food intake data primarily depends on a study’s aim; reported data thus often are not comparable across studies. Food intake data further depend on the dietary assessment methodology used and foods in the database consulted; and are influenced by the inherent limitations of all dietary assessments. Intake data can be presented as either single foods or as clearly defined food groups. Mixed dishes, reported as such or in terms of ingredients and items added during food preparation remain challenging. Comparable presentation of food consumption data is not always possible; presenting sufficient information will assist valid interpretation and optimal use of the presented data. A checklist was developed to strengthen the reporting of food intake data in science communication.

Article info

Article history:
Received 25 January 2013
Accepted 15 March 2013

Keywords:
Food intake data
Grouping of foods
Mixed dishes
Challenges

Rationale for food intake data

Publishing food intake data, as opposed to or in addition to nutrient data, reliably and in a format that permits comparisons among studies is important as people eat food, not nutrients. Knowledge of foods consumed is needed for nutrient profiling [1], costing and compiling the food basket [2], the food industry [3], pricing policies to promote healthier food purchases [4], addressing food insecurity [5], and developing and monitoring progress toward achieving food-based dietary guidelines [6]. Traditionally, the emphasis in nutritional epidemiology was on single nutrients or foods, but the complexities associated with diets suggest that it may be inappropriate to focus on the role of individual foods and nutrients in isolation [7]. Food consumption data are thus used to determine dietary patterns either by score-based methods (e.g., the Healthy Eating Index or Dietary Diversity Score [8,9]) or by data-driven methods such as principal component analysis, factor analysis [10], and cluster analysis [11].

Food intake data often are used to inform and substantiate policies and planning. In South Africa, for example, results of a national food consumption survey (NFCS) [12,13] showed that maize meal and bread were the most commonly consumed staple foods and, as a result, became vehicles for a national food fortification program. This decision was supported by secondary data analyses of available food intake data for adults and children, which showed that substituting fortified bread and maize meal in the diet for the unfortified equivalents significantly improved the micronutrient content of the diet [14,15].

Concerns have been raised that food intake data often are reported in a non-uniform format [16]. The need to report the details of dietary assessment more adequately [17] and develop
on healthy eating has been highlighted [18]. In Europe, initiatives collected and harmonized data, for effective evaluation of policies key indicators of food consumption, based on systematically Table 1 consumed food [21].

<table>
<thead>
<tr>
<th>Table 1 Options for ranking and identifying the “most important” foods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of times each food was recorded</td>
</tr>
<tr>
<td>• reflects the frequency of consumption;</td>
</tr>
<tr>
<td>• does not take into account the number of consumers, portion size or number of times consumed per day;</td>
</tr>
<tr>
<td>• is influenced by dietary methodology: 24-h recall/record: one person may consume a given food several times during the recording period, whereas another may not consume the item at all, leading to overestimation of consumption within the sample; food frequency questionnaire: the number of responses to an item reflects the number of individuals who consume the food, but does not differentiate between frequent and infrequent consumers (e.g., daily or monthly consumption);</td>
</tr>
<tr>
<td>• may exceed the total number of study participants (24-h recalls/records) for frequently consumed foods.</td>
</tr>
<tr>
<td>Percentage of individuals reporting consumption of a food</td>
</tr>
<tr>
<td>• reflects the percentage of consumers (and not “importance” by quantity or frequency);</td>
</tr>
<tr>
<td>• does not distinguish between true non-consumers (who never eat the specific food) and occasional consumers (who do eat the specific food, but did not do so during the reference period for dietary intake) among non-consumers;</td>
</tr>
<tr>
<td>• provides a good indication of the most commonly consumed foods making it useful to identify the most/least consumed foods;</td>
</tr>
<tr>
<td>• does not reflect the frequency of consumption or the amount consumed;</td>
</tr>
<tr>
<td>• can be misleading if amounts consumed are small or if consumption is occasional.</td>
</tr>
<tr>
<td>Total intake (e.g., in grams of food) by sample studied</td>
</tr>
<tr>
<td>• favors foods consumed in large quantities and discriminates against foods consumed in small (but sometimes nutritionally important) quantities*;</td>
</tr>
<tr>
<td>• is not a good representation of the number or percentage of consumers;</td>
</tr>
<tr>
<td>• is used to determine dietary exposure to both beneficial and harmful chemical substances at the population level [22].</td>
</tr>
<tr>
<td>Mean daily intake (e.g., in grams of food) per person</td>
</tr>
<tr>
<td>• can be used to identify the proportion of low or high consumers (e.g., for fruit and vegetable intake relative to a standard);</td>
</tr>
<tr>
<td>• is restricted to those who consumed the food item during the reference period;</td>
</tr>
<tr>
<td>• is affected by age and sex, thus affecting comparability.</td>
</tr>
<tr>
<td>Per capita intake</td>
</tr>
<tr>
<td>• includes the whole sample regardless of whether the food item was consumed or not;</td>
</tr>
<tr>
<td>• is comparable to daily intake for consumers of commonly consumed foods;</td>
</tr>
<tr>
<td>• may differ substantially from daily intake of consumers for foods consumed by only a small number of respondents.</td>
</tr>
<tr>
<td>Energy and/or nutrient contribution by foods and/or food groups</td>
</tr>
<tr>
<td>• identifies food sources of nutrients;</td>
</tr>
<tr>
<td>• to aid interpretation, total energy/nutrient intake must be given;</td>
</tr>
<tr>
<td>• must specify whether this refers to foods or dishes (e.g., if the aim is to report energy intake from potatoes, margarine added to e.g., mash potatoes should not be included).</td>
</tr>
</tbody>
</table>

* For example, in a South African study in primary school children sugar was ranked 17th based on the total amount consumed, yet it was the 3rd most frequently consumed food [21].

key indicators of food consumption, based on systematically collected and harmonized data, for effective evaluation of policies on healthy eating has been highlighted [18]. In Europe, initiatives are under way for harmonizing the conversion of national food consumption data to edible parts of raw agricultural commodities for dietary exposure assessments of chemicals [19].

A recent structured review of food intakes of South African rural households using electronic and hand searches, information provided by professional networks, unpublished reports, and personal collections [20] highlighted that the non-comparable presentation of food intake data across the study was a major challenge to identifying the most commonly used foods. The literature search and information extraction have been published [20]. The aim of this contribution thus is to discuss the factors affecting uniform reporting of food intake across studies, and to compile a checklist that might support the presentation and interpretation of food intake data.

**Presenting food intake data**

**Options of presenting food intake**

Meaningful interpretation of food intake data requires consideration of the dietary assessment methodology that was used, whether or not food intake was quantified, the reference period for the intake, context and target group of the study, and the original aim of collecting the data. The absolute value for the average amount of food consumed per person per day can be presented as per capita intake or for consumers only.

Table 1 lists options for ranking and identifying the “most important” foods. Often more than one analytical method is used to report food intake. One study [23], for example, reported absolute intake values as well as proportions of intake because the latter is magnitude independent, and underreporting is known to be common in dietary assessment of children [24] and adults [25] alike. Reporting average portion sizes and number of daily servings allow for identification of the proportion of low or high consumers (e.g., for fruit and vegetable intake relative to a standard) in the sample studied [3,26]. In food fortification programs, the number of consumers and the mean daily intake are of importance [27].

The length of food lists varies. In populations with limited dietary variety, the frequency of consumption of all foods may be given [28]. Some studies list only the highest ranked foods (e.g., top 10 or 25 foods [13]), whereas others use a cut-off value for the minimum number of consumers, for example, foods consumed by at least 5% of respondents [29] or, where food lists are extensive, foods consumed by at least 85% of respondents [30].

**Factors to consider when presenting and interpreting food intake data**

When food intake data are presented and interpreted, the aim of the assessment and the resultant dietary methodology, the way in which food items were grouped, decisions regarding mixed dishes and additions during preparation and processing, and finally the inherent limitations of food databases must be considered.

**Aim of the food intake study**

Whereas the South African NFCS [13] reported the 25 food items consumed most at national and provincial levels, other
local studies focused from the outset on specific foods such as vegetables and fruit [31], dark-green leafy vegetables [32], or complementary foods [33]; or on foods rich in specific nutrients such as sodium [34]. When the aim is to describe dietary diversity, only food groups may be presented [35], whereas in some cases more detailed information of foods within food groups is given [36].

**Dietary method used and inherent limitations of dietary assessments**

The strengths, weaknesses, and validity of dietary assessment methods have been described in detail [37]. This review does not focus on the various dietary assessment methods as such, but rather on the effect of the dietary assessment methodology on the reported food intake data. For example, one study [38] reported that, for 3 y old children, sweets were the third most frequently consumed food item according to a food frequency questionnaire (FFQ), but were ranked 21st according to a 24-h recall. Because a single 24-h recall does not account for day-to-day and seasonal variability of dietary intake, this may affect the estimated number of consumers of foods not eaten regularly. Compared with the 24-h recall, food frequency data generally show a greater variety of foods, a greater percentage of consumers of a specific food, and a larger average amount consumed per food item [39]. It may well be that the “truth” lies somewhere between the two methods, yet it may differ by food or food group.

FFQs often are designed to focus on specific foods, nutrients, or other food components [40] and vast differences in cultural-specific foods between populations [41] may influence the list of foods and hence the reporting and interpretation. The challenge to obtain accurate information on foods that can be eaten either on their own or combined with other ingredients in dishes when using an FFQ is illustrated in an Australian study; consumption of milk was underestimated as many respondents did not report the milk used with cereal [42].

One study [43] highlights that consumers may vary in the categories they use to describe certain foods because of different languages, cultures, and familiarity with foods (e.g., is potato a starch or a vegetable?). Differences within a country may occur because of language (e.g., in South Africa, *imifino* is a collective term for traditional green-leafy vegetables in the Zulu vernacular, and a collective term for all vegetables in the Sepedi vernacular).

An inherent flaw of dietary data, regardless of assessment method, is that the number of consumers and the number and sizes of servings may be affected by intentional or unintentional under- or overreporting [24,25]. Misreporting typically increases as the complexity of the diet increases, and may be associated with particular target groups [37] (e.g., underreporting among overweight females) or specific foods (e.g., overreporting of vegetables and fruit).

**Grouping of foods**

Grouping of foods can range from very broad, for example, based on food-based dietary guidelines, to more specific, and is primarily guided by the aim of the study. For consistency and repeatability across studies, sufficient information needs to be reported on how foods were grouped. Systematic reviews and meta-analyses on the association between consumption of certain foods and disease are compromised when there is heterogeneity among studies because of different grouping of foods [44]. Also, when calculating dietary diversity scores, both the number and type of groups created will affect the absolute value of the score [45]. Foods can be aggregated through a series of steps [46]; can be grouped on different levels [39]; or can be categorized into groups, categories (within groups), and subcategories (within categories), using a hierarchical structure [47]. Use (e.g., *pâté* and high-fat meat for sandwiches), type of behavior (e.g., not mixing vegetarian food items with meaty food items) [48], and processing method (e.g., fresh, canned, smoked, raw, dried, juice, pickled, and frozen) [39] also may be considered.

It may be beneficial to group vegetables and fruits based on similarity in food composition and on easily identifiable characteristics such as botanical family (e.g., citrus fruit family), color (e.g., deep orange; red fruits), and/or plant parts (e.g., leafy vegetables) [49]. When calculating dietary diversity scores as proxy for nutritional adequacy of the diet, β-carotene–rich vegetables and fruits are usually grouped together [50]. In the Diet Quality Index, all vegetables and fruit are grouped as one [51], whereas in the Dietary Approaches to Stop Hypertension-style diet adherence score, fruits and vegetables are two separate groups [52]. Comparability between studies is affected by whether vegetables and fruit consumed as part of mixed dishes are counted when calculating vegetable and fruit intake, as well as the definition of what counts as a vegetable or a fruit [53].

Potato, for example, is listed under the vegetable group in the South African food database (SAFOODS) [54], but is excluded when adherence to the World Health Organization’s recommended intake of at least 400 g/d of vegetables and fruit is determined [55]. Potatoes, sweet potatoes, and yams often are grouped as “roots and tubers.” Sweet potato, however, has a large genetic diversity of flesh color, ranging from white/cream-fleshed varieties with virtually no β-carotene, to orange-fleshed varieties with significant amounts of β-carotene [56]. Depending on the aim of the study, different sweet potato varieties can thus be grouped differently. For example, when determining dietary diversity [9], white sweet potato is grouped under “white roots and tubers,” whereas orange sweet potato, although a tuber, is grouped with the β-carotene–rich vegetables. To enhance comparability and reporting of vegetable and fruit intake, Moore and Lloyd [57] suggested that vegetables and fruit should be reported separately; consumption of potatoes and legumes should be reported separately, and that 100% fruit juice should be reported separately from fruit.

The aim of the study may necessitate the reporting of consumption of individual vegetables and fruit. This is the case, for example, in food-based interventions that promote production of specific β-carotene–rich vegetables. In a home garden project in South Africa, yellow vegetables such as pumpkin, butternut, and orange sweet potato were harvested in the study site during the first quarter of the year. As a result, participating households reported consumption of these vegetables during a dietary survey in February/March, but not in November [58]. Seasonal variability was also reflected by the consumption of spinach (a cool weather crop) during late winter/early spring [31], and wild growing leafy vegetables during the summer months [32]. For meaningful interpretation of the reported data, the season and region in which the study was conducted should thus be stated.

Foods can be grouped into major food groups according to their nutrient profile, with more detailed information of foods within the food groups [36]. Analyses of supermarket sales in New Zealand, for example, showed that high-fat milk, white bread, and butter were the most commonly chosen foods in
their corresponding food categories (milk, bread, and butter-and-margarine, respectively) [59]. When comparing seafood consumption with plasma fatty acid concentrations, Chung and co-workers [60] grouped fish into four groups (fried fish, non-fried shellfish, non-fried fish, and fish in mixed dishes), but did not distinguish between fresh, frozen, or processed fish, the different types of frying, and different cooking methods.

Differentiation by cooking method often leads to non-uniform reporting of foods consumed. For example, the NFCS [13] reported eggs as one group, regardless of preparation method, whereas smaller local studies differentiated between preparation methods such as “egg, boiled”; “egg, fried”; or “egg, fried in sunflower oil” [28,61]. Whether it is justified to present individual food items according to preparation method, or grouped together, regardless of preparation method, depends on the aim of the study.

Soft or stiff porridge prepared with maize meal is widely consumed in South Africa [13]. Some studies group all porridges made with maize meal together [13,30,38,61], whereas others differentiate between the soft and stiff versions [21,28,29]. Grouping all dishes made with maize meal is a good representation of the number of consumers, but interpretation of average portion size of the grouped dishes is problematic as the consistency substantially affects the dry matter content (and therefore energy and nutrient content) [54].

Bread is also widely consumed in South Africa. Whereas the NFCS reported the frequency of consumption for white and brown bread separately [13], other local studies grouped all types of bread together [21,29,38]. By reporting intake for all types of bread combined, as well as the different types of bread individually [34], total number of bread consumers and the different types of bread consumed are reflected. Homemade bread is commonly consumed in certain areas in the country [62] and, as it is not listed in the SAFOODS [54], it was grouped with white bread in some earlier studies [29]. Since the implementation of the national food fortification program, grouping homemade bread (made with unfortified cake flour) with white bread (made with fortified wheat flour) is no longer acceptable. This example highlights the importance of knowing the food and nutrition context when deciding on food grouping.

Valuable information may be lost if food intake data are presented as single food items only. For example, information on fast food intake will be lost if hamburgers are only reported as bread roll, meat pattty, and margarine, and hotdogs as only bread roll, Vienna sausage, and margarine. In the Bogalusa Heart Study, “burgers and sandwiches” were aggregated into one group [63]. In the food categorization system of an international collaboration project, “burgers” and “sandwiches” are two separate categories in the fast foods group, each with subcategories. The subcategories of, for example, “burgers” are beef burgers, fish burgers, chicken burgers, and vegetarian burgers [64].

Monteiro and co-workers [65] moved away from the more traditional food groupings and argued that, because of the dominance of processed foods in the global food system, foods should be classified based on the extent and purpose of the industrial processing used in their production. They proposed the following three main categories: 1) unprocessed and minimally processed foods; 2) processed culinary or food industry ingredients that are not edible or else normally not consumed by themselves, and 3) ultra-processed food products that are “ready-to-consume” or “ready-to-heat” foods. In this classification system, bread is classified as an ultra-processed food product, together with foods such as biscuits, sweets, and soft drinks.

Webb and co-workers [66] lament the lack of clear and accepted terminology to characterize energy-dense nutrient-poor foods, where terms such as unhealthy, junk food, fast food, and snack have been used, with as yet no consensus. They generated from a detailed food data set a study-specific group, the so-called “extra” foods, representing energy-dense, nutrient-poor items, out of the more traditional main groups.

**Mixed dishes versus foods eaten on their own**

Handling mixed dishes is always challenging. Beef stew with vegetables, for example, can be coded as beef plus each of the vegetables separately (assuming the type and proportions of the different components are known), or as the dish “beef-and-vegetable stew” (assuming the type and proportions of the different components of the standard dish are equivalent to that consumed). Reporting meat intake without disaggregation of meat dishes may result in an overestimation of meat consumption [67]. “Samp-and-beans” is a popular traditional South African dish typically with a 1:1 ratio of samp (coarsely crushed maize) and dried beans [54]. “Samp-and-beans” can be coded as samp (maize) and dried beans (legumes) separately, or as the dish “samp-and-beans.” Faber and co-workers [29] reported that in a rural South African village, 56% of the children ate legumes (beans) and 12% ate “samp-and-beans” during the recall period; the 56% for consumers of legumes may be an underestimation as the legumes consumed as part of the dish “samp-and-beans” are not counted.

In an Italian survey, recipes for homemade dishes were collected and then coded as single ingredients. For dishes consumed outside the house, standard recipes were used for coding [46]. Collecting recipes for home-prepared dishes from each research participant is challenging, particularly so in areas where ingredients are often not measured and depend on the food items available in the household. Standard recipes or set ingredients and amounts using information collected from key informants in the specific community, often are used in such studies [30]. New food items can be created and added to the food composition database, each dish with a unique code and calculated nutritional composition based on the recipe provided by the respondents [68].

Grouping of mixed dishes poses a challenge. In the National Diet and Nutrition Survey in the United Kingdom, mixed dishes were grouped either by meat or fish content, or by main food component (i.e., ingredient present in the highest proportion after cooking). Fitt and co-workers [69] illustrated that lasagna (containing 14% meat and 27% pasta), for example, would be a meat dish if classified according to meat or fish content, but a pasta dish if classified according to the main food component. They showed that neither of these two methods were entirely satisfactory and suggested that mixed dishes also could be grouped according to similar types of dishes e.g. soups, rice dishes, pasta dishes, egg dishes, and so on. In the Individual food consumption data and exposure assessment studies for children (EXPOCHI) project in Europe, composite dishes were indicated in the food consumption database at both the ingredient and composite dish level, with the option to present on ingredient or dish level [70]. In a U.S. study a cascaded “recipe file” consisting of four levels on which food can be reported, ranging from level 1 where a dish (e.g., “beef barbeque on bun”) is listed, to level 4, where the “ketchup” in the hamburger would be reported as “tomato sauce, vinegar and sugar” was developed [26].
Food items used during preparation of food

Lists of commonly consumed foods may differ depending on whether food items used during preparation were considered. For example, the South African NFCS showed that sugar was the second most commonly consumed food item for children aged 1 to 9 y-old [12]. Sugar was not considered when drawing up the list of commonly consumed foods reported for 2 to 5 y-old South African children in a rural village and their caregivers, as it was often used during food preparation and was not coded separately [29]. Similarly, cooking oil was excluded from the list of commonly consumed foods [29]. In contrast, oil was reportedly the most commonly consumed food item for 4 to 24 mo-old children in the same area (80% consumed oil during the recall period) [71]. The lists of the most commonly consumed foods across these studies were influenced by the conscious decision to either include or exclude certain food items, making these lists not comparable.

Foods in the database

Changes in the food database can affect the reporting of food intake data. For example, the 1991 South African food composition database did not distinguish between pumpkin and butternut squash [72], and consequently these were often grouped together [31]. The updated database (SAFOODS) distinguishes between pumpkin and butternut squash [54]. Another example is African leafy vegetables, which, in the 1991 food composition database [72] were all grouped together as one food item (“imifino/marog”) with average nutrient values. SAFOODS [54] no longer includes average nutrient values, but instead, gives the nutrient values for the individual African leafy vegetables (i.e., amaranth, cat’s whiskers, black jack, etc.). On the other hand, respondents may find it difficult to specify the specific leaf that was eaten and researchers need to decide which leafy vegetable to code. This poses challenges in studies promoting vitamin A-rich foods (e.g., home gardening projects), where the consumption of individual β-carotene-rich vegetables...
(such as butternut squash, pumpkin, and green-leafy vegetables) over time is compared.

New foods enter the market continually and the range of products increases, making it practically impossible to include all food items in a food database. When a food is not available in the food database, researchers often choose an alternative. Guidelines for food matching have been published [73] but researchers may, based on their experience, knowledge and the purpose of the study, choose different alternatives, influencing the results on individual food and food group intake. Recording the assumptions made is thus important.

Conclusion

Collecting dietary intake data at the national level is costly, and national surveys, where available, often focus on the most vulnerable population groups in a country. Dietary intake data from smaller studies, although often collected with an aim of only reporting general food intake, can add valuable information about the dietary intake of specific groups in a country. Presenting food intake data uniformly without compromising scientific quality is challenging, and uniformity is not always meaningful and desirable, particularly for smaller studies. When comparing findings across studies, we need to be cognizant of the factors that influence the comparability of results of dietary surveys [74] and consider the complex population dynamics. The format in which food intake data is being presented is influenced by the amount of detail that was collected during the assessment, and should be aligned to the aim of the study. Providing sufficient information is important to aid proper interpretation of the data and ensure repeatability, yet published studies often report insufficient details of the dietary assessment [17]. Table 2 provides a general checklist that may be helpful when presenting and interpreting food intake data. The checklist is an adaptation of the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for transparent reporting of observational studies in epidemiology [75]. We call it STROFI (Strengthening the Reporting Of Food Intake) hoping that over time it may become an extension of STROBE in an attempt to improve clarity and comprehensiveness when reporting on dietary assessment, in order to help evaluate bias, confounding, and generalizability.

Acknowledgment

The concept for this paper originated from discussions during interdisciplinatory project meetings of project K5/1954/4 “Water use and nutrient content of crop and animal food products for improved food security” initiated, funded, and managed by the Water Research Commission of South Africa.

References


Faber M. Complementary foods consumed by 6-12-month-old rural infants in South Africa are inadequate in micronutrients. Public Health Nutr 2005;8:373–81.


Ruel MT. Operationalizing dietary diversity: a review of measurement issues and research priorities. J Nutr 2003;133:3911S.


Thompson FE, Steyn NP, Nel JH, Thompson OM, Yaroch AL. The meaning of “fruit” and “vegetables.” Public Health Nutr 2011;14:1222–8.


