

Integrated Spreadsheets for Nutritional Analysis of Population Diet Surveys

Sabri Bromage¹

¹ Department of Nutrition, Harvard T.H. Chan School of Public Health

Corresponding Author

Sabri Bromage

sab742@mail.harvard.edu

Citation

Bromage, S. Integrated Spreadsheets for Nutritional Analysis of Population Diet Surveys. *J. Vis. Exp.* (188), e64327, doi:10.3791/64327 (2022).

Date Published

October 18, 2022

DOI

10.3791/64327

URL

jove.com/video/64327

Abstract

It can be challenging to calculate nutrient intakes in population diet surveys because existing nutritional analysis software is generally oriented toward analyzing intakes of individuals and may not allow users to input or easily modify the food composition data used in the analysis. These are drawbacks that are more problematic in low- and middle-income country settings. While there are numerous software-assisted dietary assessment platforms that conduct onboard nutritional analysis and are appropriate for use in large surveys, they are often similarly limited, and further restrict users to specific assessment modalities. This paper presents a multifunctional system of integrated spreadsheets for nutritional analysis of population diet surveys (ISNAPDS) that provides a solution for situations in which data have been collected but cannot be adequately analyzed with existing software. The protocol involves supplying the system with fully customizable data on food composition, food group classifications, and food intake (food intake in g/day may be entered directly or calculated based on user-supplied intake frequencies and either standard or variable serving sizes). Following data entry, the user modifies a set of simple pre-populated formulas to match them to the structure of the input data and the system applies these formulas to calculate nutrient and food group intakes, and the contributions of food groups to nutrient intakes for all members of the survey population. The flexibility of the ISNAPDS system allows it to accommodate the global diversity of foods consumed and analyze quantitative, semiquantitative, and nonquantitative food consumption data collected using prospective and retrospective assessment methods employing different reference periods and portion size estimation methods. To date, the system has been applied in published and ongoing analyses of 24 h recall, diet record, food frequency, and disaggregated household consumption data from population surveys in China, Ethiopia, India, Mongolia, Thailand, and a multi-country analysis of 10 sub-Saharan African countries.

Introduction

Data on population, food, and nutrient intake are important for understanding the burden of malnutrition in populations and the relationships between diet and health, and play important roles in designing, monitoring, and evaluating evidence-based nutrition policies and programs^{1,2}.

After data on food intake have been collected, software is used to multiply the amount of each food consumed by its nutrient composition to obtain data on nutrient intake (nutritional analysis)³, a process that used to be performed manually until the advent of mainframe computers⁴. There are numerous software tools that do this, but they are generally oriented toward the analysis of individuals rather than population surveys^{5,6}. Investigators wanting to calculate nutrient intakes in large surveys may write programs using statistical software that they are not highly proficient with or resort to applying software designed for individuals to every member of the survey population and compiling the results; this is time-consuming and error-prone. Furthermore, existing nutritional analysis software may not include all the foods consumed or nutrients of interest in a particular survey or allow users to input or easily customize data on food composition, serving sizes, and food group classifications used in the analysis. While many software-assisted dietary assessment platforms exist that conduct onboard nutritional analysis and are appropriate for use in large surveys⁷, they are often similarly limited and further restrict users to specific assessment modalities (e.g., diet record, 24 h diet recall, or recall of the frequencies at which foods were typically consumed over a specified reference period).

These drawbacks are more problematic in low- and middle-income countries (LMICs) where local food composition,

recipes, and other reference data are often largely unrepresented in existing software for dietary assessment and nutritional analysis, most of which is designed for use in high-income countries^{2,7,8}. Investigators collecting data on population nutrient intakes may therefore use software-assisted dietary assessment tools that are poorly suited to their survey population or research questions, or resort to using purpose-built tools that may not contain onboard food composition data or arduous pen-and-paper methods; both of these require a separate solution for nutritional analysis⁹. Inadequacies in the existing software thus compound numerous other impediments to producing high-quality and time-relevant food and nutrient intake data needed for implementing effective nutrition strategies in LMICs². The recently developed INDDX24 Dietary Assessment Platform is a notable effort to address this data gap in LMICs^{10,11,12}. INDDX24 uses a mobile app for streamlined data collection which is linked seamlessly with the Global Food Matters Database, an extensive and growing repository of global food composition, standard recipes, and dietary reference data¹³. However, INDDX24 is limited to collecting 24 h recalls, which is considered the most broadly applicable assessment method for use in population surveys, but which may not fulfill all the research objectives, particularly if a lengthy reference period is needed (in which it may be more appropriate to collect food-frequency questionnaires).

This paper presents a multifunctional system of integrated spreadsheets for nutritional analysis of population diet surveys (ISNAPDS) that provides a solution for situations in which data have been collected but existing software is not adequately suited to analyze them. The protocol involves supplying the system with fully customizable data

on food composition, food group classifications, and food intake (food intake in g/day may be entered directly or calculated based on the user-supplied intake frequencies and either the standard or variable serving sizes). Following data entry, the user modifies a set of simple pre-populated formulas to match them to the structure of the input data. The system then applies these formulas to calculate nutrient and food group intakes and the contributions of food groups to nutrient intakes for all members of the survey population. The flexibility of the ISNAPDS system allows it to accommodate the global diversity of foods consumed and analyze quantitative, semiquantitative, and nonquantitative (i.e., qualitative) food consumption data collected using prospective and retrospective assessment methods employing different reference periods and portion size estimation methods (e.g., single, or repeated diet records or recalls, or food-frequency questionnaires).

Orientation to ISNAPDS, a summary of protocol steps, and explanation of formulas:

The ISNAPDS system (**Supplemental File 1**) is a Microsoft Excel Open XML (.xlsx) file originally developed in 2012. The version used here was developed in 2022 using Excel 365.

ISNAPDS is comprised of eight spreadsheets connected by formulas that convert input data on standard or variable serving sizes (expressed in g/serving), food composition (in units/day), food group classifications, food intake frequencies (in servings/day) into output data on food intake (in g/day, which may alternatively be supplied as input data if they are readily available), food group intake (in g/day), nutrient intake (in units/day), and the contributions of food groups to

nutrient intakes (in units/day). The system is integrated in the sense that the formulas match the foods, observations, nutrients, and food groups in each sheet to their correct counterparts in the other sheets; all the sheets are contained in a single file. The system is pre-populated with example input and output data for 10 observations that demonstrate the system's functionality (in the system and the protocol, the term observation may refer to individuals or person-days such as in the case of repeated diet records or recalls).

The protocol for using ISNAPDS involves four main steps: (1) specifying whether food intake data will be supplied by the user in g/day or must be calculated based on intake frequencies and serving sizes, and (if food intake must be calculated) whether each food is associated with a standard serving size or whether serving sizes for each food vary between observations; (2) structuring and entering data on food composition, food group classifications, and either food intake or intake frequencies accompanied by standard or variable serving sizes; (3) modifying pre-populated formulas in the output spreadsheets to match the structure of the input data; and (4) propagating formulas within the output data sheets to populate the desired results. Parts of some protocol steps can be skipped depending on the specified method for populating food intake data and whether the user is interested in calculating nutrient intakes, food group intakes, or the contributions of food groups to nutrient intakes. **Figure 1** is a flowchart that summarizes the protocol steps and input data required for a given method of populating food intake data and the desired output data.

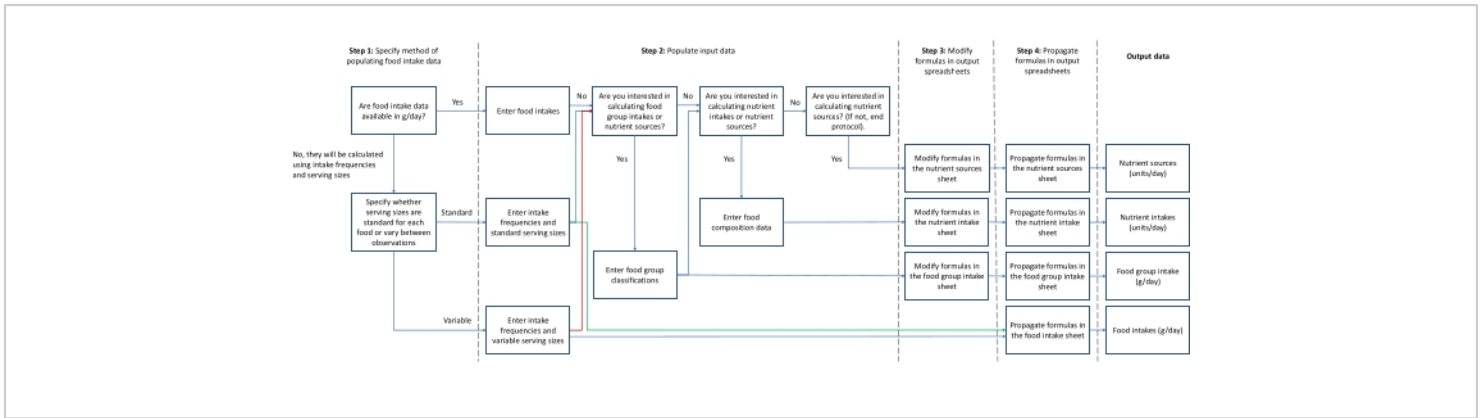


Figure 1: Protocol flowchart. Summary of the protocol steps and input data required given the method of populating food intake data and the desired output data. Step 2 (structuring and entering data) is simplified to exclude steps and parts of steps that involve structuring data. [Please click here to view a larger version of this figure.](#)

For each observation, ISNAPDS calculates the food intake in g/day (if these data are not user-supplied) by multiplying the intake frequency of each food by its serving size¹⁴. Calculating the intake of each food group involves a simple summation of the intake of all foods classified as belonging to that group. The formula for calculating nutrient intake involves multiplying the intake of each food by the concentration of each nutrient in that food per 100 g, dividing by 100, and summing the results across all foods consumed³. To calculate the contribution of each food group to the intake of each nutrient, the same formula is simply applied separately for each combination of food group and nutrient.

New users are advised to review the two-page supplemental guidelines for entering and manipulating data in the ISNAPDS spreadsheets (**Supplemental File 2**) to prevent errors and ensure that they maintain integration between the sheets. These guidelines can also be used to troubleshoot visible errors (e.g., cells displaying #REF! or #VALUE) and errors that may be detected upon running the data integrity checks described in the representative results.

Protocol

The protocol is accompanied by example results based on applying the ISNAPDS system in analysis of actual population survey data in Mongolia. The procedures followed for collecting data in this survey were in accordance with the ethical standards of the Mongolian Ministry of Health Ethical Review Board and the Harvard T.H. Chan School of Public Health Institutional Review Board (Protocol #21002). The eligible participants provided written informed consent to join the study and provide publishable data prior to enrolment and were free to withdraw from the study at any time.

1. Specifying the method for populating food intake data

1. If food intake data are readily available in g/day, skip to step 2.
2. If food intake in g/day must be calculated based on intake frequencies and serving sizes, select cell B1 of the Food Intake Method spreadsheet and use the dropdown

menu to specify whether each food is associated with a standard serving size (the default option) or whether serving sizes for each food vary between observations (variable serving sizes).

NOTE: If food intake must be calculated but serving sizes are not readily available (e.g., in analysis of data collected using a nonquantitative food-frequency questionnaire), consider conducting a separate analysis of quantitative survey data to derive standard or variable serving sizes for different demographic strata^{15,16,17,18}.

2. Structuring and populating the input data spreadsheets using the pre-populated example data as a guide

1. In column A of the Food Composition spreadsheet, enter a list of all the unique foods reported to be consumed in the survey data (subsequently referred to as the food list). In the Food Intake sheet, enter the food list in row 1 and a list of all the unique observations in the survey data (subsequently referred to as the observation list) in column A.
 1. Ensure that the order and length of the food list in row 1 and the observation list in column A of any other sheet in which data is entered are consistent with the order and length of these lists in the Food Composition and Food Intake sheets.
2. Populate food intake in g/day by entering the necessary input data depending on the method specified in step 1.
 1. If food intake data are readily available in g/day, populate the Food Intake sheet for all foods and observations in the data and skip to step 2.3.
 2. If food intake must be calculated based on intake frequencies (expressed in servings/day) and either standard or variable serving sizes (in g/serving), in the Intake Frequencies sheet, enter the food list and the observation list and populate the sheet for all foods and observations in the data.
3. If calculating food intake using standard serving sizes, enter these data in column C of the Food Composition sheet. If using variable serving sizes, in the Variable Serving Sizes sheet, enter the food list and the observation list and populate the sheet for all foods and observations in the data.
3. If interested in calculating food group intakes or the contribution of food groups to nutrient intakes, enter food groups in column B of the Food Composition sheet. In the Food Group Intake sheet, enter a list of all the unique food groups represented in the data (subsequently referred to as the food group list) in row 1 and the observation list in column A.
 1. Ensure that the length of the food group list and the name of each food group are consistent between column B of the Food Composition sheet and row 1 of the Food Group Intake sheet (and if calculating the contribution of food groups to nutrient intakes, row 1 of the Nutrient Sources sheet). Food groups are not case-sensitive and may be entered in any order.
4. If interested in calculating nutrient intakes or the contribution of food groups to nutrient intakes, in the Food Composition sheet, enter the list of all the nutrients to be analyzed in row 1 (subsequently referred to as the nutrient list) and populate the sheet with food composition data (expressed in unit/100 g). In the Nutrient Intake sheet, enter the nutrient list in row 1 and the observation list in column A. Summary guidance and online resources on how to identify, obtain, compile,

match, and assess the quality of food composition data is provided in **Supplemental File 3**.

NOTE: Despite its name, the nutrient list can include ingredients, foods, or food groups (e.g., Flour (g), Bread (g), or Refined grains (g)) and the data populated in the Food Composition sheet can therefore include ingredient, food, or food group concentrations (expressed in g/100 g) obtained from data on recipes or equivalency factors. This enables an additional or alternative approach for calculating food or food group intakes. When using this approach, foods or food groups are treated as if they were nutrients and their intakes are populated in the Nutrient Intake sheet. This approach provides a solution for situations in which there is a need to calculate intakes of specific ingredients or foods of interest (such as fortification vehicles) that are not included in the food list, or when analyzing data collected using a dish-based dietary assessment method and there is a need to assign food groups at the level of ingredients or foods.

1. Ensure that the order and length of the nutrient list are consistent between row 1 of the Food Composition sheet and row 1 of the Nutrient Intake sheet (and if calculating the contribution of food groups to nutrient intakes, row 1 of the Nutrient Sources sheet).
5. If interested in calculating the contribution of food groups to nutrient intakes: in the Nutrient Sources spreadsheet, enter the nutrient list in row 1. In column A, enter the observation list # times by appending the list with copies of itself in a vertical fashion, where # is the number of food groups in the food group list. Populate every cell in column B with a different food group for each stack of observations.

3. Modifying the pre-populated formulas in the output spreadsheets to match the structure of the input data

1. If interested in calculating food group intakes or the contribution of food groups to nutrient intakes: in the formula in cell B2 of the Food Group Intake spreadsheet, $\{=SUM(('Food\ Composition'!B2:B21=B$1)*TRANSPOSE('Food\ Intake'!$B2:$U2))\}$, change \$B\$21 to \$B\$#, where # is the bottom row of the data entered in the Food Composition spreadsheet and change \$U2 to \$*2, where * is the rightmost column of the data entered in the Food Intake sheet. Enter the modified formula.
2. If interested in calculating nutrient intakes or the contribution of food groups to nutrient intakes: in the formula in cell B2 of the Nutrient Intake spreadsheet, $\{=SUM(('Food\ Intake'!$B2:$U2)*TRANSPOSE('Food\ Composition'!D$2:D$21))/100\}$, change \$U2 to \$#2, where # is the rightmost column of the data entered in the Food Intake spreadsheet and change \$D\$21 to \$D\$*, where * is the bottom row of the data entered in the Food Composition spreadsheet. Enter the modified formula.
3. If interested in calculating the contribution of food groups to nutrient intakes: in the formula in cell C2 of the Nutrient Sources spreadsheet, $\{=SUM(INDEX('Food\ Intake'!B1:U11,ROW()-ROUNDDOWN((ROW()-2)/(10),0)*(10),0)*(TRANSPOSE('Food\ Composition'!D$2:D$21)*($B2='Food\ Composition'!$B$2:$B$21)))/100\}$, change \$U\$11 to \$#11, where # is the rightmost column of the data entered in the Food Intake spreadsheet; change \$D\$21 and \$B\$21 to \$D\$* and \$B\$*, respectively, where * is the

bottom row of the data entered in the Food Composition spreadsheet; and change both instances of (10) to (!), where ! is the length of the observation list. Enter the modified formula.

NOTE: If using a software version prior to 2021, enter formulas as array formulas using CTRL + SHIFT + ENTER/RETURN (refer to **Supplemental File 2** for guidance).

4. Propagating the formulas within the output spreadsheets to populate the desired results

1. If food intake must be calculated based on intake frequencies and either standard or variable serving sizes: in the Food Intake spreadsheet, select cell B2 and copy and paste or drag it rightward and downward until every observation in the data is populated with intake of every food in g/day.
2. If interested in calculating food group intakes or the contribution of food groups to nutrient intakes: in the Food Group Intake spreadsheet, select cell B2 and copy and paste or drag it rightward and downward until every observation in the data is populated with intake of every food group in g/day.
3. If interested in calculating nutrient intakes or the contribution of food groups to nutrient intakes: in the Nutrient Intake spreadsheet, select cell B2 and copy and paste or drag it rightward and downward until every observation in the data is populated with the intake of every nutrient group in units/day.
4. If interested in calculating the contribution of food groups to nutrient intakes: in the Nutrient Sources spreadsheet, select cell C2 and copy and paste or drag it rightward and downward until every observation in the data is populated

with the contribution of every food group to every nutrient in units/day.

Representative Results

Checks for ensuring integrity of the output data

The checks below demonstrate the accuracy of the calculations in the ISNAPDS system using observation 1 of the pre-populated example data. To ensure the protocol is correctly adhered to when applying the system in analysis of actual survey data, it is recommended that users run each of these checks themselves for few different observations and output columns (errors in a given input spreadsheet may propagate additional errors in some output spreadsheets but not others, and errors in a given output spreadsheet may only affect sections of the data and may go unnoticed if checks are run for only one observation and output column).

Checking calculation of food intake using standard or variable serving sizes

For observation 1, intake of whole milk in g/day is equal to intake in servings/day (cell B2 of the Intake Frequencies spreadsheet: 0.14 servings/day) multiplied by either the standard serving size for whole milk (cell C2 of the Food Composition spreadsheet: 85 g/serving) or the serving size for whole milk specified for observation 1 (cell B2 of the Variable Serving Sizes spreadsheet: 130 g/serving) depending on whether standard or variable serving sizes are used:

If standard serving sizes are used: $0.14 * 85 = 12$ g/day. If variable serving sizes are used: $0.14 * 130 = 19$ g/day.

In both cases, the result is equal to the number in cell B2 of the Food Intake spreadsheet.

For simplicity, subsequent checks are based on food intakes calculated using standard serving sizes (the default method for populating food intake in the example data).

Checking calculation of food group intake

For observation 1, intake of the high fat dairy group is equal to the sum of the intakes of all foods in the high fat dairy group (cell B2 and K2 of the Food Intake spreadsheet: 12 g/day for milk and 17 g/day for yogurt):

$$12 + 17 = 29 \text{ g/day}$$

This result is equal to the number in cell B2 of the Food Group Intake spreadsheet.

Checking calculation of nutrient intake

For observation 1, energy intake is equal to the product of two vectors: intakes of all foods (cells B2 to U2 of the Food Intake spreadsheet: 12, 35, 40, 11, 2, 30, 6, 21, 14, 17, 26, 1, 80, 34, 9, 80, 10, 20, 4, and 0 g/day) and their respective energy composition (cells D2 to D21 of the Food Composition spreadsheet: 267.9, 1113.0, 623.7, 482.0, 1272.0, 63.0, 3699.0, 544.0, 96.0, 255.4, 180.0, 1619.0, 361.0, 494.0, 1252.0, 371.0, 601.0, 1276.0, 360.0, and 0.0 kcal/100g), divided by 100

$$(267.9 * 12 + 1113 * 35 + 623.7 * 40 + 482 * 11 + 1272 * 2 + 63 * 30 + 3699 * 6 + 544 * 21 + 96 * 14 + 255.4 * 17 + 180 * 26 + 1619 * 1 + 361 * 80 + 494 * 34 + 1252 * 9 + 371 * 80 + 601 * 10 + 1276 * 20 + 360 * 4 + 0 * 0) / 100 = 2420.5 \text{ kcal/day}$$

This result is slightly different than the number in cell B2 of the Nutrient Intake spreadsheet (2417.2) due to rounding. To illustrate: cell B2 of the Intake Frequencies spreadsheet (milk intake of observation 1 in servings/day) has been formatted to display as 0.14, which corresponds to a reported intake frequency of 1/week, which is more precisely expressed

as 0.142857142857143 servings/day (this is the number displayed in the formula bar when the cell is selected and is what the formula uses in its calculations). When 0.142857142857143 is multiplied by the serving size for milk (cell C2 of the Food Composition spreadsheet: 85 g), it results in an intake of 12.1428571428571 g/day, which has been formatted to display as 12 in cell B2 of the Food Intake spreadsheet. If cells B2 to U2 of the Food Intake spreadsheet were reformatted to include three decimal places and the expression above were modified accordingly and reevaluated, it would result in the number in cell B2 of the Nutrient Intake spreadsheet (this can be verified by the user).

Checking calculation of the contributions of food groups to nutrient intakes

For observation 1, energy intake from the high fat dairy group is equal to the product of two vectors: intakes of all foods in the high fat dairy group (cells B2 and K2 of the Food Intake spreadsheet: 12 g/day for milk and 17 g/day for yogurt) and their respective energy composition (cells D2 and D11 of the Food Composition spreadsheet: 267.9 kcal/100g for milk and 255.4 kcal/100g for yogurt), divided by 100

$$(267.9 * 12 + 255.4 * 17.1) / 100 = 75.8 \text{ kcal/day.}$$

This result is slightly different from the number in cell C22 of the Nutrient Sources spreadsheet (76.3) due to rounding, as illustrated in the previous check. If the expression above were revised to include three decimal points for food intakes, it would equal the number in cell C22: $(267.9 * 12.143 + 255.35 * 17.143) / 100 = 76.3 \text{ kcal/day.}$

As another check of the calculations in the Nutrient Sources spreadsheet, the user can filter rows to exclude all but one observation, calculate the sum of contributions from all food groups to intake a given nutrient (for energy intake of

observation 1, these contributions are in cells C2 to C112: 13.7, 458.2, 76.3, 111.7, 211.4, 404.1, 65.2, 504.7, 506.1, 50.4, 15.4, and 0.0 kcal/day, which sum to 2417.2 kcal/day), and observe that this sum is equal to the observation's total intake of that nutrient in the Nutrient Intake spreadsheet (for energy intake of observation 1, this number is in cell B2 of that spreadsheet).

Example results based on applying the ISNAPDS system in an actual population survey

Table 1, Table 2, and Table 3 provide example results based on applying ISNAPDS in the analysis of repeated weighed diet records (DRs) collected from a survey of 334 men and women living in eight regions of Mongolia. Up to 3 days of DRs were collected from each participant in both summer and winter and 1,838 person-days of DRs were collected in total (mean number of DRs per person: 5.50). References 19 and 20 provide detailed dietary assessment methods and results, respectively^{19,20}.

After entering these data in ISNAPDS and populating the output spreadsheets, each spreadsheet was imported by statistical software, R, which merged them by observation. The software also received data on each participant's sex, locality (urban vs. rural), other demographic characteristics, and the season and observation day in which each DR was collected, which were further merged with the output data. The software was used to estimate statistics within different population subgroups and seasons, including the mean intakes of food groups and nutrients (adjusted for within-person variation), and the mean contribution of food groups to nutrient intakes.

Table 1: Mean intake (g/day) of food groups among Mongolian adults in different population subgroups and seasons. Abbreviation: DRs = diet records. The table has

been modified from²⁰. [Please click here to download this Table.](#)

Table 2: Mean nutrient intake (per day) among Mongolian adults in different population subgroups and seasons. Abbreviation: DRs = diet records. The table has been modified from²⁰. [Please click here to download this Table.](#)

Table 3: Mean percentage of dietary energy contributed by consumed dishes among Mongolian adults in different population subgroups and seasons. Abbreviation: DRs = diet records. The table has been modified from²⁰. [Please click here to download this Table.](#)

Supplemental File 1: The ISNAPDS system. [Please click here to download this File.](#)

Supplemental File 2: Guidelines for entering and manipulating data in the ISNAPDS system. [Please click here to download this File.](#)

Supplemental File 3: Summary guidance and online resources for preparing food composition data for nutritional analysis of population diet surveys. [Please click here to download this File.](#)

Discussion

The ISNAPDS system presented in this paper provides a convenient starting point for numerous analyses central to nutritional surveillance and epidemiology such as: estimating intake distributions of nutrients and food groups, determining the prevalence of nutrient inadequacy and excess, identifying key food sources of each nutrient, deriving food- or nutrient-based diet metrics or adherence to dietary guidelines, analyzing relationships between diet and health outcomes, and informing the design of nutrition programs. ISNAPDS has preempted all of these analyses in population diet

surveys conducted across a range of low- and middle-income countries in Asia and Africa using a range of assessment methods. These surveys have involved 1,838 person-days of weighed diet records collected from 334 adults and 24 h recalls from 1,356 adults in Mongolia^{20,21,22,23}, quantitative food-frequency questionnaires (FFQs) and repeated 24 h recalls from 3,564 Ethiopian adults²⁴, non-quantitative FFQs from 3,171 adults in 10 African countries¹⁸, ongoing surveys involving semiquantitative FFQs from 352 Indian adults²⁵ and 4,543 Mongolian children²⁶, and other ongoing surveys in China, Thailand, Ethiopia, and Mongolia. The system has even been applied in analysis of disaggregated household food consumption data to estimate food group and nutrient intakes of 39,016 Mongolian adults and children²³.

To keep things simple, there are four main functions of ISNAPDS: calculating population intakes of foods, food groups, nutrients, and the contributions of food groups to nutrient intakes, that are broadly applicable to different studies and require relatively little tailoring to the input data. With some proficiency, the system can be expanded to accommodate more detailed functions that would be useful for some studies, such as for compiling recipes, accounting for supplement use, or applying adjustment factors for refuse, density, cooking yield, and nutrient retention^{27,28,29,30}. Some users may wish to add functions or spreadsheets that implement basic data validation, data visualization, or descriptive analyses. Spreadsheets may also be added to calculate additional indicators of specific interest to a given survey and provide a more comprehensive nutritional analysis. For example, a new spreadsheet could be supplied with harmonized average requirements or probabilities of nutrient adequacy for different combinations of intake level, age, sex, and pregnancy and lactation status, and the Nutrient Intake spreadsheet could be appended with columns

indicating the individuals' nutrient adequacy based on the estimated average requirement cut-point or full-probability method, allowing users to estimate the prevalence of nutrient adequacy^{31,32}. As another example, the Food, Food Group Intake, Nutrient Intake, or Nutrient Sources spreadsheets could be appended with columns indicating whether or to what extent individuals satisfy different criteria for scoring diet metrics or adherence to dietary guidelines, and these could feed into new columns or a separate spreadsheet for calculating total scores^{33,34,35}. Alternatively, some users may prefer to use statistical software to perform these tasks particularly if they must be preceded by algorithmic or statistical adjustments. For surveys in specific populations or multi-country studies, the system could be pre-populated with data and metadata on food lists, food composition, standard serving sizes, and food group classifications (such as harmonized classifications useful for cross-country comparisons) to aid analysis and potentially prepare the system for operationalization as a dietary assessment tool^{13,36,37,38}.

The functionality of ISNAPDS basically comprises a series of simple matrix multiplications that could be adapted for use in other spreadsheet software with similar functionality or translated into a program for use in statistical software (see the *NutrienTracker* package in R for an example³⁹). Using a statistical program to replicate these functions would not necessarily speed up the analysis or reduce the likelihood of errors because, as in ISNAPDS, the input data would still need to be formatted for the software and the program syntax would need to be tailored to the input data. Furthermore, while Excel is the world's most popular spreadsheet software and nutritionists generally have some experience with it, proficiency with statistical software varies as does its availability. To be widely accessible, programs would need

to be written in multiple coding languages such as R, SAS, STATA, and SPSS, and the published programs would need to be maintained as the base software is updated. Importantly, as the calculations in a statistical program are conducted under-the-hood, some errors may be more difficult to identify and troubleshoot, and separation between the program and the input files may make analyses more cumbersome. By contrast, ISNAPDS provides a standalone, integrated analytical platform that allows users to observe hands-on how inputs affect outputs dynamically. For most research applications, statistical software will be helpful or necessary to conduct subsequent analyses of the outputs that ISNAPDS produces such as: identifying outliers, applying algorithms for estimating nutrient bioavailability, adjusting for within-person variation or total energy intake, computing nutrient ratios, deriving diet patterns and metrics, and associational analyses^{28,29,40,41}. However, ISNAPDS allows much of the initial analysis to be conducted in a relatively user-friendly spreadsheet interface. Owing to its transparency and ease-of-use, the system can also be a useful tool for explaining basic processes involved in nutritional analysis of diet surveys and to date has been used for this purpose in introductory workshops on population research dietetics involving 45 clinical researchers in Ethiopia and Mongolia.

To the author's knowledge, there exists only one other spreadsheet-based system in use, DietSys^{9,42}, which is also fully contained in one file, calculates food group and nutrient intakes, can accommodate user-supplied food composition and serving size data, and provides another example of the flexibility and utility of integrated spreadsheets in nutritional analysis of population diet surveys. DietSys differs from ISNAPDS in that the purpose is to facilitate data entry and nutritional analysis in 24 h recall surveys. To this end, DietSys contains a user interface programmed in Visual Basic for

Application (VBA) that guides the user through a series of simple steps for entering foods (that are automatically coded according to entries in the accompanying food composition data), recipes, number of servings, and portion sizes. It tabulates food group and nutrient intakes by the participant, day, and meal and links these data with nondietary data collected in the same survey. The system's database is currently populated with food composition data for 9,851 foods from Brazil, Argentina, and the U.S. that are pre-assigned to food groups. The serving sizes are derived from the Brazilian population and NOVA food processing classifications (allowing users to calculate the contribution of processed foods to diets)⁴³. In testing the system in Brazilian adults, using DietSys required only 27% of the time that SPSS did for data entry and analysis of the same set of 24 h recall surveys⁹.

An important limitation of ISNAPDS is that the user must supply all the data on food composition, standard serving sizes, and food group classifications used in the analysis. This allows the system to be used in any food culture; however, if the necessary input data are not readily available at hand, compiling them can be an arduous process. This is especially true for food composition data given significant global deficits in data availability, accessibility, ease-of-integration into the software, quality, recency, comprehensiveness, interpretability, and harmonization^{44,45,46}. An effort is ongoing to provide open access online repositories containing crowdsourced dietary reference data that are harmonized across numerous world regions, allow data retrieval in spreadsheet format, and provide adequate metadata to interpret and assess the quality of the data. Recent examples toward this effort include the Global Food Matters Database and WikiFCD^{13,47}. A second important limitation also applicable to other spreadsheet-based analysis systems is

that the user is primarily responsible for identifying and fixing errors that may arise due to numerous possible mistakes in entering or manipulating data or formulas. Some errors will be readily apparent to the user (e.g., in cells displaying #REF! or VALUE#), but others will simply produce erroneous values that may go unnoticed and propagate additional errors. While identifying and fixing errors is partly facilitated by the system's relative transparency and integrated design, it is recommended that users perform the data integrity checks detailed in the representative results section, troubleshoot errors using the supplemental guidelines for entering and manipulating data, and backup their work frequently.

Disclosures

The author has nothing to disclose.

Acknowledgments

The author would like to thank Dr. Rosalind S. Gibson, Dr. Walter C. Willett, Dr. Rebecca L. Lander, Dr. Teresa T. Fung, Theresa L. Han-Markey, Dr. Guy Crosby, Dr. Megan Deitchler, Dr. Mourad Moursi, Dr. Helena Pachón, Dr. Suzanne M. Cole, Dr. Tzy-Wen L. Gong, and Laura A. Sampson for education and guidance provided over the past decade on dietary assessment and nutritional analysis; Dr. Kelvin Gorospe for advice about translating the functionality of the ISNAPDS system into a statistical program; Dr. Sinara L. Rossato for information about DietSys; and Dr. Winnie Bell for information about INDDEX24 Dietary Assessment Platform and Global Food Matters Database. The author received support from the National Institutes of Health (T32 DK 007703).

References

1. Food and Agriculture Organization of the United Nations (FAO), Intake - Center for Dietary Assessment. *Global report on the state of dietary data*. Rome, Italy (2022).
2. Micha, R., Coates, J., Leclercq, C., Charrondiere, U. R., Mozaffarian, D. Global Dietary Surveillance: Data gaps and challenges. *Food and Nutrition Bulletin*. **39** (2), 175-205 (2018).
3. Greenfield, H., Southgate, D. A. T. Food composition data and food composition databases. In *Food composition data: production, management, and use, 2nd ed.* Burlingame, B. A., Charrondiere, U. R., Eds. Food and Agriculture Organization of the United Nations (FAO) Publishing Management Service: Rome, Italy. 5-20 (2003).
4. West, C. E., Van Staveren, W. A. Food composition, nutrient intake, and the use of food composition tables. In *Design concepts in nutritional epidemiology, 2nd ed.* Margetts, B. M., Nelson, M., Eds. Oxford University Press: New York, NY, USA. 101-119 (1997).
5. Food and Agriculture Organization of the United Nations (FAO), *International Network of Food Data Systems (INFOODS)*. Software Tools. Available online: <https://www.fao.org/infoods/infoods/software-tools/en/> accessed on 17 September (2022).
6. Food and Nutrition Service (FNS), United States Department of Agriculture (USDA). *USDA Approved Nutritional analysis Software*. Available online: <https://www.fns.usda.gov/tn/usda-approved-nutrient-analysis-software>, accessed on 17 September (2022).
7. McNutt, S., Zimmerman, T. P., Brooke, C. Computer-assisted dietary assessment methods. In *Advances in the Assessment of Dietary Intake*. Schoeller, D. A.,

- Westerterp-Plantenga, M. S., Eds. CRC Press: Boca Raton, FL, USA. 45-66 (2017).
8. Coates, J. C., Colaiezzi, B., Bell, W. F., Gibson, R. S. *INDDEX Priority Technical Criteria and Review of Technology-Assisted 24-h Recall Software Programs*. Tufts University: Boston, MA, USA (2015).
 9. Rossato, S. L., Fung, T. T., Rodrigues, M. P. A Data entry system for dietary surveys based on visual basic for applications programming. *Journal of the Academy of Nutrition and Dietetics*. **117** (8), 1165-1170 (2017).
 10. Rogers, B. L. et al. Validation of the INDDEX24 mobile app vs. a pen-and-paper 24-hour dietary recall using the weighed food record as a benchmark in Burkina Faso. *The British Journal of Nutrition*. Advance online publication. 1-41 (2021).
 11. Coates, J. et al. Accuracy and cost-effectiveness of the INDDEX24 Dietary Assessment Platform in Viet Nam. *The British Journal of Nutrition*. Advance online publication. 1-38 (2022).
 12. Adams, K. P. et al. The cost and cost efficiency of conducting a 24-h dietary recall using INDDEX24, a mobile dietary assessment platform, compared with pen-and-paper interview in Viet Nam and Burkina Faso. *The British Journal of Nutrition*. Advance online publication. 1-15 (2022).
 13. International Dietary Data Expansion (INDDEX) Project. *Global Food Matters Database*. Available online: <https://inddex.nutrition.tufts.edu/global-food-matters-database>, accessed on 17 September (2022).
 14. Willett, W. C. Food Frequency Methods. In *Nutritional Epidemiology, 3rd ed.* Oxford University Press: New York, NY, USA. 70-95 (2013).
 15. Tjonneland, A. et al. Influence of individually estimated portion size data on the validity of a semiquantitative food frequency questionnaire. *International Journal of Epidemiology*. **21** (4), 770-777 (1992).
 16. Wakai, K. et al. A simple food frequency questionnaire for Japanese diet--Part I. Development of the questionnaire, and reproducibility and validity for food groups. *Journal of Epidemiology*. **9** (4), 216-226 (1999).
 17. Schlundt, D. G. et al. Separate estimates of portion size were not essential for energy and nutrient estimation: results from the Southern Community Cohort food-frequency questionnaire pilot study. *Public Health Nutrition*. **10** (3), 245-251 (2007).
 18. Bromage, S. et al. The global diet quality score is inversely associated with nutrient inadequacy, low midupper arm circumference, and anemia in rural adults in ten sub-Saharan African countries. *The Journal of Nutrition*. **151** (12 Suppl 2), 119S-129S (2021).
 19. Bromage, S. et al. Trained cohorts of University students are a useful resource for conducting dietary surveys in Mongolia. *Food and Nutrition Bulletin*. **38** (2), 267-272 (2017).
 20. Bromage, S. et al. Diet and nutrition status of Mongolian adults. *Nutrients*. **12** (5), 1514 (2020).
 21. Bromage, S. et al. Projected effectiveness of mandatory industrial fortification of wheat flour, milk, and edible oil with multiple micronutrients among Mongolian adults. *PLoS One*. **13** (8), e0201230 (2018).
 22. Jeong, C. et al. Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proceedings of the National Academy of Sciences of*

- the United States of America*. **115** (48), E11248-E11255 (2018).
23. Bromage, S. et al. Comparison of methods for estimating dietary food and nutrient intakes and intake densities from household consumption and expenditure data in Mongolia. *Nutrients*. **10** (6), 703 (2018).
 24. Bromage, S. et al. The global diet quality score is associated with higher nutrient adequacy, midupper arm circumference, venous hemoglobin, and serum folate among urban and rural Ethiopian adults. *The Journal of Nutrition*. **151** (12 Suppl 2), 130S-142S (2021).
 25. Mistry, N. et al. Protocol for a case-control study of vitamin D status, adult multidrug-resistant tuberculosis disease and tuberculosis infection in Mumbai, India. *BMJ Open*. **10** (11), e039935 (2020).
 26. Ganmaa, D. et al. Prevalence and determinants of QuantiFERON-diagnosed tuberculosis infection in 9810 Mongolian schoolchildren. *Clinical Infectious Diseases*. **69** (5), 813-819 (2019).
 27. Food and Agriculture Organization of the United Nations (FAO), International Network of Food Data Systems (INFOODS). Training. Available online: <https://www.fao.org/infoods/infoods/training/en/>, accessed on 17 September (2022).
 28. Intake - Center for Dietary Assessment. Resources. Available online: <https://www.intake.org/resources>, accessed on 17 September (2022).
 29. Gibson, R. S., Ferguson E. L. *An interactive 24-hour recall for assessing the adequacy of iron and zinc intakes in developing countries*. ILSI Press, Washington, DC, USA. 98-107 (2008).
 30. Vásquez-Caicedo, A. L., Bell, S., Hartmann, B. *Report on collection of rules on use of recipe calculation procedures including the use of yield and retention factors for imputing nutrient values for composite foods*. European Food Information Resource (EuroFir): Brussels, Belgium (2008).
 31. Allen, L. H., Carriquiry, A. L., Murphy, S. P. Perspective: Proposed harmonized nutrient reference values for populations. *Advances in Nutrition*. **11** (3), 469-483 (2020).
 32. Institute of Medicine (US) Subcommittee on Interpretation and Uses of Dietary Reference Intakes. *DRI Dietary Reference Intakes: Applications in Dietary Planning*. National Academies Press: Washington, DC, USA (2000).
 33. Herforth, A. et al. A global review of food-based dietary guidelines. *Advances in Nutrition*. **10** (4), 590-605 (2019).
 34. Miller, V., Webb, P., Micha, R., Mozaffarian, D. Global Dietary Database. Defining diet quality: a synthesis of dietary quality metrics and their validity for the double burden of malnutrition. *The Lancet Planetary Health*. **4** (8), e352-e370 (2020).
 35. Bromage, S. et al. Development and validation of a novel food-based Global Diet Quality Score (GDQS). *The Journal of Nutrition*. **151** (12 Suppl 2), 75S-92S (2021).
 36. Ioannidou, S., Cascio, C., Gilsenan, M. B. European Food Safety Authority open access tools to estimate dietary exposure to food chemicals. *Environment International*. **149**, 106357 (2021).
 37. Food and Agriculture Organization of the United Nations (FAO) and International Network of Food Data Systems (INFOODS). *International*

- food composition table/database directory*. Available online: <https://www.fao.org/infoods/infoods/tables-and-databases/en/>, accessed on 17 September (2022).
38. Food and Agriculture Organization of the United Nations (FAO) and International Network of Food Data Systems (INFOODS). *Food nomenclature, terminology and classification systems*. Available online: <https://www.fao.org/infoods/infoods/standards-guidelines/food-nomenclature/en/>, accessed on 17 September (2022).
 39. Rodriguez-Martinez, A., Ayala, R., Balchunas, M., Debbabi, Y., Vidal, L. S. *NutrienTrackeR: Food Composition Information and Dietary Assessment R package version 1.2.0*. <https://CRAN.R-project.org/package=NutrienTrackeR>, accessed on 17 September (2022).
 40. Willett, W. C. *Nutritional Epidemiology, 3rd ed.* Oxford University Press: New York, NY, USA (2013).
 41. Gibson, R. S. *Principles of Nutrition Assessment, 2nd ed.* Oxford University Press, New York, NY, USA (2005).
 42. Rodrigues, S. L. et al. Development of DietSys: A comprehensive food and nutrient database for dietary surveys. *Journal of Food Composition and Analysis*. **102**, 104030 (2021).
 43. Monteiro, C. A. et al. Ultra-processed foods: what they are and how to identify them. *Public Health Nutrition*. **22** (5), 936-941 (2019).
 44. Greenfield, H., Southgate, D. A. T. Guidelines for the use of food composition data. In *Food composition data: production, management, and use, 2nd ed.* Burlingame, B. A., Charrondiere, U. R., Eds. Food and Agriculture Organization of the United Nations (FAO) Publishing Management Service: Rome, Italy. 187-197 (2003).
 45. Sumbo, P., Beecher, G. Strengths and Limitations of Food Composition Databases. In *Advances in the Assessment of Dietary Intake*. Schoeller, D. A., Westerterp-Plantenga, M. S., Eds. CRC Press: Boca Raton, FL, USA. 337-352 (2017).
 46. Kapsokefalou, M. et al. Food composition at present: New challenges. *Nutrients*. **11** (8), 1714 (2019).
 47. Thornton, K., Seals-Nutt, K., Matsuzaki, M. Introducing WikFCD: Many food composition tables in a single knowledge base. *CEUR Workshop Proceedings*. **2969** (2021).