Design and Quality Control Issues Related to Dietary Assessment, Randomized Clinical Trials and Meta-Analysis of Field-Based Studies in Developing Countries

Rosalind S. Gibson,*2 Sunil Sazawal† and Janet M. Peerson**

*Department of Human Nutrition, University of Otago, Dunedin, New Zealand, †Department of International Health, Johns Hopkins Bloomberg School of Public Health and Center for Micronutrient Research, Annamalai University, India and **Program in International Nutrition, Department of Nutrition, University of California, Davis, CA

ABSTRACT The essential design and quality control aspects of dietary assessment, randomized controlled trials (RCT) in developing countries and their subsequent meta-analysis are reviewed. Dietary assessment protocols consist of three stages: measurement of food intakes using a method appropriate for the study objectives, calculation of nutrient intakes and finally an evaluation of their nutrient adequacy. The latter may involve: 1) assessment of dietary diversity (average number of different foods consumed per day) and selected indices of dietary quality such as intakes of animal source foods (g/d); 2) percentage of energy from protein, fat and animal protein; 3) selected nutrient densities; and 4) dietary phytate:zinc molar ratio, as well as the prevalence of inadequate nutrient intakes calculated using a probability approach. To establish causality between the nutrient inadequacies identified and adverse health outcomes, RCT must be undertaken. A prerequisite of RCT is double-blind randomization, a procedure designed to eliminate biases arising from baseline confounding variables, unintended interventions and ascertainment bias. Results from existing RCT can be summarized via meta-analysis to gain a better understanding of the population relationship. Meta-analysis is a statistical technique involving linear models or generalized linear models, which can be performed after locating the individual studies, and selecting and abstracting all the necessary data. J. Nutr. 133: 1569S–1573S, 2003.

KEY WORDS: dietary assessment * probability analysis * randomized clinical trials * double blinding * meta-analysis

Until recently, the United Nations has urged that priority be given to programs in developing countries to prevent deficiencies of iodine, vitamin A and iron. There is increasing recognition, however, that in developing countries where staple diets are predominantly plant based and intakes of animal source foods are low, inadequacies of several other micronutrients may also coexist (1). The impact of such multiple micronutrient deficiencies on adverse health outcomes such as physical growth, resistance to infection, work capacity, cognitive development, school performance and physical activity in children and adults is gaining increasing recognition. Consequently, it is essential to use dietary assessment to identify those nutrients likely to be inadequate in plant-based diets, so that appropriate interventions can be designed. To establish a causal relationship between the nutrient inadequacies identified and adverse health outcomes, randomized controlled trials (RCT) must be conducted. Increasingly, the results of RCT are being combined through meta-analysis to gain a better understanding of the population relationship. Therefore, this brief article will address some essential design and quality control issues of dietary assessment protocols, RCT and meta-analysis of such studies in developing countries.

1 Published in a supplement to The Journal of Nutrition. Presented as part of the 11th meeting of the international organization, "Trace Elements in Man and Animals (TEMA)” in Berkeley, California, June 2–6, 2002. This meeting was supported by grants from the National Institutes of Health and the U.S. Department of Agriculture, and by donations from Akzo Nobel Chemicals, Singapore; California Dried Plum Board, California; Cattlemen’s Beef Board and National Cattlemen’s Beef Association, Colorado; Clinical Nutrition Research Unit, University of California, Davis; Dairy Council of California, California; GlaxoSmithKline, New Jersey; International Atomic Energy Agency, Austria; International Copper Association, New York; International Life Sciences Institute Research Foundation, Washington, D.C.; International Zinc Association, Belgium; Mead Johnson Nutritional, Indiana; Minute Maid Company, Texas; Perrier Vittel Water Institute, France; U.S. Borax Inc., California; USDA/ARS Western Human Nutrition Research Center, California; Wyeth-Ayerst Global Pharmaceuticals, Pennsylvania. Guest editors for the supplement publication were Janet C. King, USDA/ARS WHINRC and the University of California at Davis; Lindsay H. Allen, University of California at Davis; James R. Coughlin, Coughlin & Associates, Newport Coast, California; K. Michael Hambidge, University of Colorado, Denver; Carl L. Keen, University of California at Davis; Bo L. Lönnerdal, University of California at Davis and Robert B. Rucker, University of California at Davis.

2 To whom correspondence should be addressed. E-mail: Rosalind.Gibson@stonebow.otago.ac.nz.
Design and implementation issues for dietary assessment protocols

Dietary assessment generally involves three phases: 1) measuring food intakes, 2) converting food intakes to observed intakes of nutrients and antinutrients, and 3) evaluating nutrient adequacy (2). Food intakes can be measured using quantitative or qualitative dietary assessment methods, depending on the study objectives. The quantitative methods consist of food records or recalls designed to measure the quantity of the individual foods consumed over a 1-d period. By increasing the number of measurement days, quantitative estimates of usual intakes of individuals can be obtained. Qualitative methods include the food frequency questionnaire (FFQ) and dietary history, each designed to obtain retrospective information on the pattern of food use during a longer, less precisely defined time period. With modification, both FFQ and dietary histories can provide data on usual nutrient intakes.

The selection of the food intake method depends primarily on the study objectives, characteristics of the study group, respondent burden, and available resources. Three categories of study objectives can be defined. The simplest involves assessing the mean intake of a group, when only a single recall or record is required, provided all days of the week are represented in the final sample. When information on the prevalence of inadequate intakes within the study group is needed, then at least two independent recalls or records on all or a subsample of the population should be completed. Finally, to examine relationships between intake and laboratory and/or functional health outcomes, multiple recalls or records are essential; the number of days depends on the intrasubject variation in nutrient intakes. An equation for estimating the number of days required per subject is available in Willett (3). Alternatively, usual nutrient intakes based on a semiquantitative FFQ or dietary history can be collected (4).

The weighed food record, completed in the households by trained dietary monitors, is the quantitative method that has been used most often in developing countries in the past (4–6). Alternatively, a modified 24-h recall, especially designed for assessing iron and zinc intakes in rural populations in developing countries, can be used (4). The modified interactive 24-h recall method involves training the respondent before the recall, where possible, and incorporates the use of plates, picture charts and food models to aid respondents in portion-size estimation and recall. The aim of these modifications is to reduce the systematic and random measurement errors by enhancing recall of foods consumed, reducing the number of memory lapses and improving the portion-size estimates.

After data collection on food intake, the data are usually converted into nutrient and antinutrient intakes, generally by using a food composition table or nutrient data base. The international dietary assessment system, WorldFood Dietary Assessment System, developed by the University of California at Berkeley, is recommended for use in developing countries (7). The system contains values for 53 nutrients (including iron, zinc, copper and manganese) and associated dietary components (including phytic acid and dietary fiber) for 1800 foods consumed in Egypt, Kenya, Mexico, Senegal, India and Indonesia, as well as a computer program for calculating energy and nutrient intakes (7). It is possible to augment this nutrient database with food composition values for additional foods, where necessary. This may be especially important for trace elements for which their content in staple foods is dependent on local soil trace element levels (i.e., Se, I, Zn) (8,9). Details on how to select foods for chemical analysis and how to compile a local nutrient data base are given in Gibson and Ferguson (4).

An added advantage of the WorldFood Dietary Assessment System is that it computes intakes of utilizable protein, as well as available iron and zinc based on the algorithms of Murphy et al. (10).

The final stage in the dietary assessment protocol is to evaluate the adequacy of the food and nutrient intakes. Several approaches can be used including the assessment of dietary diversity (i.e., total number of different foods consumed per day) and dietary quality. In developing countries, indicators of dietary quality often include intake of animal protein (g/d), the proportion of energy from animal sources (as %), intakes of key problem nutrients expressed as nutrient densities and phytate:zinc molar ratios of the diets. In addition, the prevalence of inadequate intakes of nutrients should be calculated, where appropriate, using probability analysis. This approach aims to predict the number of persons within a specific life-stage and gender group likely to have nutrient intakes below their own requirement estimates, and hence provides a better assessment of risk of nutrient inadequacies for individuals within a population (11). One of the prerequisites for conducting probability analysis is to obtain reliable data on the distribution of usual nutrient intakes among individuals within the study population. For dietary data based on 1-d recalls or records, this can be achieved by adjusting the nutrient intakes for intraindividual variation (day-to-day variation within one subject) using the National Research Council (11) method, or a refinement (12), provided 1-d recalls or records have been repeated, at least on a subsample of the population. Note that the adjustment process yields a distribution with reduced variability that preserves the shape of the original observed distribution (2). For dietary methods designed to assess usual nutrient intakes over a retrospective time frame (e.g., semiquantitative FFQ or dietary histories), no adjustment is required.

Once the distribution of usual nutrient intakes has been calculated, then short-cut probability analysis can be applied to derive a more reliable estimate of the prevalence of inadequate intakes of certain nutrients in the study population (13). This approach involves calculating the proportion of individuals with usual nutrient intakes below their corresponding mean requirement estimates. Estimates of the latter derived by United Nations agencies, the World Health Organization and the Food and Agricultural Organization, are available in Bunch and Murphy (7). Alternatively, requirement estimates developed for a specific country or region can be used. For iron, full probability analyses must be applied because the distribution of iron requirements for menstruating women is highly skewed (4).

Design and implementation issues for randomized controlled trials in field-based developing country settings

RCTs are the only way to establish causality between a nutrient deficiency and an adverse functional outcome such as diarrheal diseases, stunting, or impaired cognitive function, etc. Hence, they can provide justification for implementing public health interventions aimed at preventing nutrient deficiencies. Nevertheless, RCT are expensive, time-consuming and logistically complex in developing countries, so that it is critical that they are well designed and executed to avoid biased estimates of treatment effects (14,15).

There are six steps in the design and execution of RCT. The first step involves obtaining permission from the Human Ethics Committees within the country as well as collaborating institutions. Next, support and participation at the national,
community and household level must be secured, and mechanisms established for ongoing discussions and dissemination of information to the community. In the second step, the study sample must be identified by defining first the sample universe (via mapping, visitation routes, providing household identifiers, census performance), and then the inclusion and exclusion criteria. At this stage, a data-management system must also be established, instruments designed, pretested and modified, where necessary, and strategies put in place to ensure the continuous provision of study supplies throughout the trial. The third step consists of a baseline assessment to define the study population for external validity, to compare across studies and among intervention groups, to establish any bias that may arise from withdrawals, exclusions and/or migrations and to evaluate effect modification.

The fourth step consists of the double-blind randomization procedure, a process whereby allocation must be unknown to subjects, investigators and outcome assessors (i.e., double-blinding), and any subject being enrolled has an equal probability of being allocated to any group (i.e., randomization) (16,17). Use of double-blinding avoids bias to unintended interventions and ascertainment bias, whereas use of randomization ensures that bias is prevented and that the groups at baseline are comparable with respect to both measured and immeasurable variables. For successful double-blinding, an independent person/group must be responsible for drawing the randomization code and then a two-stage process must be used for labeling the allocation of the treatment, first at the group, and then at the subject level. In addition, the color, taste and packaging of all treatments must be prepared so that they are identical, and the frequency and dose of treatments must be considered carefully. For successful randomization, simple or permuted block randomization (fixed or variable block length) is required using random number tables or appropriate computer software. Details of this process must be fully documented (15). At this stage, certain delivery issues related to the intervention must also be considered, such as storage of the treatments in the field, their packaging, transfer and labeling, quantity of supply to be left in the home, frequency of delivery, monitoring of compliance, and concerns about "contamination" and sharing among family members. For large trials, initial and ongoing consultation with a data safety and monitoring body (DSMB) is recommended. The purpose of the DSMB is to interact with the Human Ethics Committee to establish the frequency and variables to be reviewed, to determine the sample size and the early stopping rules for safety and/or efficacy and to review data from other studies (14).

The fifth step involves following the subjects throughout the RCT. This can be achieved by active or passive visitation. In the former, the subjects are actively followed in the home or in a clinic, whereas in passive visitation only severe illnesses are monitored through hospital visits or attendance at a health facility. The frequency of the visits for active assessment varies. For example, to assess the incidence and duration of morbidity, visits can be daily, twice weekly or weekly, whereas for growth and/or development assessment, the visits will be less frequent, the timing often being dependent on the age of the subjects. Quality control measures must be incorporated into each of the chosen follow-up protocols, and should include baseline and intermittent training, supervised visitations, random checks and computer-based logical checks for real-time data entry. In addition, strategies for dealing with absent subjects, and/or tracking them within the study population for the purposes of treatment or computer data management, are essential (15).

Data analysis is the final step in the execution of a RCT. First the data are cleaned, after which data analysis, usually via "intention to treat", is performed (14). Such an analysis involves including all those subjects who were allocated to the treatment groups at the start of the study and not just those who actually received the treatment. Subgroup analysis often follows on those groups predetermined by the sample size calculation and the randomization procedure.

**Design and implementation issues for meta-analysis**

Meta-analysis examines data at the study level to determine whether effects or associations exist, to estimate their magnitude, and to identify any explanatory factors. This statistical technique has several benefits. Of these, one major advantage is its ability to consolidate results of individual studies, thereby increasing statistical power to detect effects or associations. A secondary benefit is that by identifying factors that may influence the magnitude of the effect of interest, it can identify population subgroups who may benefit from a nutrition intervention and highlight further research needs. Nevertheless, meta-analysis has some drawbacks. For example, any flaws in the individual studies analyzed become flaws in the meta-analysis, and if it is performed on observational studies only and not RCT, then no cause-and-effect relationship can be inferred. Further, if relevant studies have not been published because the results are "negative," then the meta-analysis will yield effects that appear larger than they truly are. Another drawback is that because the researcher does not have access to the raw data, it is more difficult to identify factors (e.g., sex or age) that might modify response to treatments or associations. Hence, the meta-analysis must be well designed from the outset, so that any problems encountered are dealt with appropriately.

A well-designed meta-analysis requires a statistician to assist with the data extraction and perform the complex statistical analysis involving linear models or generalized linear models (18). A principal investigator is needed to formulate the research questions and hypotheses and to interpret the results. Meta-analyses include five steps, each of which is outlined below, using examples from a meta-analysis of the effects of zinc supplementation on growth of prepubertal children (19).

The first step is to identify the study objective (e.g., to determine whether zinc supplementation improves the growth of prepubertal children and to examine sources of heterogeneity in study outcomes). In the second step, the inclusion/exclusion criteria for the studies are defined. These must be based on the methods, and on not the results of the studies. Of these criteria, the most important is whether to restrict the analysis to RCT only or to include observational studies also. Other considerations may include subject characteristics, study location and use of other treatments. In the zinc meta-analysis, only RCT in which the control and intervention groups were treated simultaneously were considered (19). In addition, subjects were all less than 12 y of age or prepubertal throughout the study, and growth measurements (weight or height) were recorded at baseline and at the end of the supplementation period. In relation to treatment criteria, only zinc intakes differed between the two groups, with the control group receiving its habitual dietary zinc intakes and the treatment group its zinc supplements for at least 8 wk, without any interruptions.

The third step in the meta-analysis is to identify the information that must be abstracted from all the chosen studies. Generally, three types of variables are abstracted: a) outcome variables (e.g., change in height, or the development of iron deficiency); b) characteristics of the study (e.g., location, amount and frequency of treatment dose, study duration, "quality" variables); and c) mean values of subject-level
characteristics (e.g., initial mean age, body mass index (BMI), biomarker concentrations). When the outcome variables are binary, then the relative risk or odds ratio is used, whereas when they are continuous, the actual difference between groups is used if the outcome variable is expressed in uniform units. For continuous variables, “effect size” is used (20). For RCT, effect size is the difference between the control and treatment group means divided by the standard deviation of the variable; essentially, the treatment mean is treated as a Z-score relative to the control mean. For observational studies, effect size is often calculated from correlation information. Generally, effect sizes of 0.2 are regarded as small, effect sizes of 0.5 as medium, and effect sizes of 0.8 or greater as large. In the zinc meta-analysis (19), all of the outcome variables (change in height for age, change in weight for age, change in weight-for-height and change in plasma or serum zinc) were continuous so effect size was used as the outcome variable. Actual change in serum zinc concentrations could also have been analyzed because serum zinc can be expressed in equivalent units for all studies.

In the fourth step, the literature is reviewed in an effort to locate all studies meeting the inclusion criteria. Large data bases, such as Medline and the Cochrane library, should be searched first, then bibliographies of earlier published studies, especially reviews, and finally word of mouth. Of the studies identified, many may be rejected immediately based on their title or abstract (e.g., observational trials, studies in the wrong species). Of the remainder, selection should be performed by a minimum of two independent observers, at least one of whom should be an “objective” observer with no emotional or scientific interest in the meta-analysis results in an effort to avoid unconscious bias. The observers must assess the methods of each study independently, meeting together later to reconcile any differences. Reasons for each rejection should be recorded in a data base for subsequent inclusion in the published analysis (21,22).

In the fifth step, each study is reviewed independently by at least two people (they may be the same people as those above) to record information on the three types of variables defined in step three. In particular, for the continuous variables, requests for raw or summarized data from the authors may be necessary.

Effect sizes can be calculated from p-values, t-values, or F-values. Standard deviations can be calculated from standard errors or estimated from interquartile ranges, ANOVA tables or bar graphs with error bars. In some cases, results must be pooled if they are presented separately for subgroups (e.g., males and females). Some of these techniques result in imprecise estimates of standard deviations and effect sizes. In cases where the effect of treatment is not statistically significant, relevant information is often omitted so that contact with the author is necessary. To avoid results that are biased against the null hypothesis of zero effect, an effect size of zero can be assumed.

The sixth and final step in the meta-analysis is the analysis and interpretation of the results. Because the variance of the sampling error is known, statistical programs, such as the SAS MIXED procedure, which allow the analyst to specify exactly the variance of the term, should be used. The first decision in the analysis is whether to estimate the mean effect size using a fixed-effects approach (i.e., the only source of variation is sampling variation) or a random-effects approach, whereby the underlying effect size of the studies naturally vary from each other (21). Of the two, the random-effects model is preferred because it allows statistical inference to extend to all possible studies that examine the relationship of interest. However, the random-effects model has less power to detect significant effects. The variance of the sampling error term depends on the form of the outcome variable, whereas the variance of the random effect is usually thought to be unknown but constant.

A specialized type of regression analysis, “meta-regression,” is used to examine the relationship between effect sizes (or other outcome variables) and potential effect modifiers. A fixed-effects or random-effects model can be used, as discussed earlier, although the latter is recommended (22). In meta-regression, the underlying “true” effect size of each individual study is fitted to a regression equation, which can have one or more explanatory terms; the number depends on the number of studies in the meta-analysis.

A typical analysis may include the following stages: 1) assess publication bias by examining the relationship between effect size and sample size, and 2) calculate a 95% confidence interval for the mean effect size using a random effects approach. Based on these results, use Rosenthal’s “file drawer” method for estimating the number of nonsignificant studies needed to overturn the result, if the effect is significantly greater than zero (23). Alternatively, if a fixed-effect model is used, assess whether there is significant heterogeneity between studies using Hedges’ χ² test (24). If the latter, switch to a random-effects model or use meta-regression to identify sources of heterogeneity. If desired, 3) examine relationships between effect size and study characteristics using meta-regression, and 4) present the results with graphs. Commonly used graphs include: 1) individual confidence intervals for each study, along with an overall confidence interval for all studies combined; 2) confidence intervals for specific subgroups (e.g., those with low versus normal mean plasma zinc concentrations); 3) “cumulative” confidence intervals that show chronologically the contribution of each study to the overall knowledge about the effect size; and 4) scatterplots with individual study effect sizes on the vertical axis and a covariate of interest on the horizontal axis (e.g., effect size for change in height versus initial mean height-for-age of the study subjects).

In summary, the first stage of a dietary assessment protocol involves carrying out the most appropriate method for measuring food intake, depending on the study objectives. Nutrient intakes are generally calculated from the food intake data with the aim of evaluating nutrient adequacy, provided reliable data on the distribution of usual nutrient intakes are obtained. With such data, interventions employing a RCT design can be performed so that causal relationships between nutrient deficiencies and adverse health outcomes can be confirmed. Results from individual RCT that meet predefined inclusion criteria can then be combined through meta-analysis to gain a better understanding of the population relationship and explain any diversity in results.

LITERATURE CITED


