



Practice of Epidemiology

Validation of a Food Frequency Questionnaire by Direct Measurement of Habitual ad Libitum Food Intake

David R. Paul, Donna G. Rhodes, Matthew Kramer, David J. Baer, and William V. Rumpler

From the Beltsville Human Nutrition Research Center, Diet and Human Performance Laboratory, Agricultural Research Service, US Department of Agriculture, Beltsville, MD.

Received for publication January 27, 2005; accepted for publication May 25, 2005.

Food frequency questionnaires are commonly used to assess habitual food intake. Although food frequency questionnaires are known to produce measurement error, the amount of error and effectiveness of correction methods are poorly understood. Twelve men from the Baltimore, MD/Washington, DC, area consumed an ad libitum diet for 16 weeks during the spring of 2001. At the end of the study period, subjects reported their food intakes with a food frequency questionnaire (Block 98). During weeks 8 and 16, subjects were dosed with doubly labeled water and maintained physical activity logs. Absolute and relative macronutrient intakes were poorly predicted by the food frequency questionnaire. The application of a single, group mean energy adjustment (using doubly labeled water or physical activity) reduced the variance of carbohydrate intake and increased the variance of fat and protein intakes, but none significantly ($p > 0.05$). Subject-specific energy adjustments reduced the variance for carbohydrate and protein intakes ($p < 0.05$). Including a body weight adjustment reduced the variance in fat intake ($p < 0.05$) when doubly labeled water was used to first correct energy intake. The application of correction methods based on energy expenditure and body weight can be used to reduce measurement error, improving the ability of the food frequency questionnaire to measure food intake.

bias (epidemiology); energy intake; energy metabolism; epidemiologic methods; nutrition assessment; questionnaires

Abbreviations: BHNRC, Beltsville Human Nutrition Research Center; SD, standard deviation; SE, standard error.

One of the most common tools used to study the relation between diet and disease is the food frequency questionnaire, because the ease of food frequency questionnaire administration and low cost allow investigators to conduct studies with large sample sizes. It has been known for some time that measurement error in the food frequency questionnaire can be large (1–5), which has led to the development of statistical procedures to reduce it (2, 3, 5–7).

There are three quantities to consider when corrections are generated for a food frequency questionnaire: 1) the food frequency questionnaire estimates, 2) the reference method, and 3) true (habitual) intake. Reference methods used in food frequency questionnaire validation studies in-

clude weighed-food records (8, 9), diet records (10–13), and 24-hour recalls (1, 9, 12, 14–16). Problems arise if the reference method is not close to true intake, for then one must understand how it differs from true intake to compensate for its deficiencies. Such deficiencies include correlated errors with the food frequency questionnaire (e.g., if there was underreporting of energy in both the food frequency questionnaire and the reference method), measurement bias (which may differ among subjects and food types), and if the relation between reference measurement error and true intake depended on the magnitude of intake (e.g., if measurement error was greater in those with higher intakes). With these limitations, the difficulty with using

Correspondence to Dr. David R. Paul, Beltsville Human Nutrition Research Center, Diet and Human Performance Laboratory, Agricultural Research Service, US Department of Agriculture, 307B Center Road, Beltsville, MD 20705 (e-mail: pauld@ba.ars.usda.gov).

food and diet records and 24-hour recall techniques for validation studies is the measurement error that is well known (4, 9, 17, 18). Furthermore, maintaining diet records tends to interfere with the daily lives of subjects, altering dietary habits so that they no longer represent true habitual intake (19).

Reference methods also include biomarkers, such as doubly labeled water and/or urinary nitrogen excretion, because they reduce the potential for correlated error and do not interfere with the measurement of food intake (1, 9, 16). However, a single 1- or 2-week mean value for energy or protein intake may not fairly represent the period of time covered by a food frequency questionnaire (several months to a year), since there is high variability in daily food intake (coefficient of variation of approximately 27 percent) (20). Moreover, doubly labeled water and nitrogen excretion can provide estimates for only energy and protein intakes, respectively, and not for the other macronutrients or specific foods.

The underlying difficulty with calibrating a food frequency questionnaire is not knowing how well the reference technique reflects true intake for the same period of time measured during the food frequency questionnaire. Not only does this lack of knowledge affect the ability of investigators to gauge the amount of measurement error, it also makes it difficult to understand the effect of applying energy corrections to uncorrected food frequency questionnaire data. Our study is based on a 16-week period when the true intake was known for all subjects, thus allowing us to determine which method or combination of methods best adjusts food frequency questionnaire estimates toward their true values. We report on biases of the different adjustment methods and estimate the reduction in variance achieved by each. Our intent is not to provide specific coefficients to adjust food frequency questionnaire results but rather to demonstrate that food frequency questionnaire correction is possible using energy expenditure estimates and body weight.

MATERIALS AND METHODS

Subjects

Twelve healthy, nonsmoking men (eight Caucasians, two African Americans, two Asians) participated in this study (table 1). The educational level of the subjects included some college ($n = 1$) or associate's ($n = 1$), bachelor's ($n = 4$), master's ($n = 4$), or Ph.D. ($n = 2$) degrees. The subjects were weight stable and not using any medications known to affect food intake, appetite, or water balance. The Johns Hopkins University Bloomberg School of Public Health Committee on Human Research approved the study protocol. Subjects provided written, informed consent and received a medical evaluation by a physician that included measurement of blood pressure and analysis of fasting blood and urine samples to screen for the presence of metabolic disease. Relatively homogeneous subjects were deliberately chosen to reduce the between-person variance in reported absolute intakes (21).

TABLE 1. Characteristics of the subjects,* Maryland, spring 2001

| | Mean (SD)† |
|--------------------------------------------------------|-------------|
| Age (years) | 39 (9) |
| Height (m) | 1.81 (0.07) |
| Weight (kg) | 79.9 (8.3) |
| Body mass index (weight (kg)/height (m) ²) | 24.1 (1.4) |
| Body fat (%) | 18.1 (1.7) |

* Twelve men.

† SD, standard deviation.

Actual food intake

Voluntary food intake was studied continuously for 16 weeks, during which time subjects consumed only foods provided by the Human Studies Facility at the Beltsville Human Nutrition Research Center (BHNRC). Subjects chose foods ad libitum from defined menus and could consume any part or all of a food item, before returning the remaining portion to be weighed. BHNRC staff that came into contact with the subjects provided no guidance as to the quantities and/or types of food items chosen. During weekdays, subjects reported to the BHNRC in the morning to eat breakfast, packed selected food items for lunch, and then returned again in the evening for dinner. Any food taken from the Human Studies Facility that was subsequently not eaten (all or partial quantities) was returned the next day, weighed, and recorded. On Friday evenings, subjects were provided with coolers packed with a large amount of food for weekend meals. The weekend coolers provided a wide variety of foods in excess quantities, and subjects were allowed to request that additional food items be included. Weekend food could be consumed on either day as long as the subjects logged which day each food item was eaten. All uneaten weekend food was returned on Monday, weighed, and recorded. Although subjects were instructed to consume only the food items provided by the Human Studies Facility, they were allowed free access to beverages including caloric, noncaloric, and alcoholic beverages. Detailed records of the amount, composition, and name brand of beverages were submitted daily.

Menus

Food items offered in the morning (breakfast and lunch) were presented in a cafeteria-style setting as three different rotating menus, each lasting 7 days (table 2). Some food items remained on all three menus (e.g., milk and orange juice). In the evening, breakfast and lunch items were also available. A typical dinner was presented cafeteria style as one- or two-entrée selections, with optional gravies or sauces and a minimum of three vegetables and side dishes. A garden salad with a variety of additional toppings and dressings was also available. Fifteen different dinner menus were rotated daily during the 16-week period (table 2).

TABLE 2. Representative food offerings to 12 male subjects during breakfast and lunch (one of three weekly rotations) and one dinner (one of 15 daily rotations), Maryland, spring 2001

| Breakfast and lunch | | | | | | | Dinner |
|---------------------|------------------|--------------------|-------------------|---------------------------|---------------------|--------------|---------------------|
| Beverages | Cereals | Bread | Meat, dairy, eggs | Snack | Packaged foods | Produce | |
| 2% milk | Hot (six items*) | English muffin | Ham | Fig bars | Vegetable soup | Apple | Turkey |
| Skim milk | Cold (10 items) | Waffle | Chicken salad | Granola bar (low fat) | Beef/vegetable soup | Orange | Chicken gravy |
| Orange juice | | Honey bun | Salami | Popcorn | Clam chowder | Banana | Mashed potatoes |
| Apple juice | | Bread (four items) | Provolone cheese | Shortbread cookies | Noodle soup | Grapes | Mixed vegetables |
| Vegetable juice | | Pita bread | American cheese | Brownie | Pizza | Peaches | Citrus salad |
| | | Buttery cracker | Scrambled egg | Strawberry twist | Pocket sandwich | Dates | Cranberry sauce |
| | | Saltine cracker | Bacon | Chocolate bar (two items) | Sausage biscuit | Garden salad | Sourdough bread |
| | | | Yogurt (fat free) | Peanuts | | Lettuce | Macaroni and cheese |
| | | | Cottage cheese | Peanut butter | | Tomato | |
| | | | Parmesan cheese | | | Carrots | |
| | | | | | | Cucumber | |
| | | | | | | Celery | |

* Number of items available in a category.

The goals of the menu design were to allow detection of macronutrient selection by offering a wide range of carbohydrate-, fat-, and/or protein-rich foods and to provide a variety of commonly available foods typically consumed by many Americans. In a research setting, it is impossible to duplicate the degree of food choice available in real life. However, more than 300 food items were used to develop menus for this study, and specific requests for food items were incorporated into the menus whenever possible.

Recording and tracking of food intake

After each subject selected his desired foods, he presented them to a staff member who recorded the identity and weight of each food item by hand and on a computer (combination of bar code recognition of the food item and hand entering of the weight). Upon termination of feeding, each subject presented his tray to a staff member who weighed any uneaten food. The accuracy of the food item recording process was verified by comparing the information on the computer with the hand-entered logs. This verification procedure was followed daily and repeated at the end of the study with all food records. Energy and macronutrient composition were determined by consultation with the US Department of Agriculture National Nutrient Database for Standard Reference (22).

Body weight and composition

Before breakfast and after voiding, body weight was determined weekly on an electronic balance to the nearest

0.01 kg. Body composition was measured by a QDR 4500A dual-energy x-ray absorptiometer (Hologic, Inc., Waltham, Massachusetts) during weeks 1, 8, and 16.

Estimated energy intake

In the absence of weight change over a defined period of time, energy intake and expenditure are assumed to be equal (23). Therefore, to obtain a surrogate estimate of energy intake, we used doubly labeled water to estimate energy expenditure. The doubly labeled water technique (24) was used to provide an estimate of energy expenditure for 7 days, twice during the study (weeks 8 and 16). Subjects reported to the BHNRC between 6:30 and 9:00 a.m., at which time they received an oral dose of H₂¹⁸O (0.16 g/kg of body weight) and ²H₂O (0.30 g/kg of body weight). Urine samples were collected immediately before the dose and on every morning (second void) for the last 7 days of the treatment period. The first sample was collected approximately 24 hours after the dose. Enrichments of ²H and ¹⁸O in urine samples were measured by infrared spectroscopy and isotope ratio mass spectrometry, respectively. The total energy expenditure and, in turn, energy intake were calculated using the equations of Weir (25).

Physical activity

Physical activity was estimated using a daily recording log method, modified from the method of Bouchard et al. (26). Briefly, subjects recorded their daily activities in a log every 15 minutes for 7 consecutive days, the same days as

TABLE 3. Macronutrient intakes measured during 16 weeks of ad libitum feedings (actual) by 12 males compared with intakes predicted from a food frequency questionnaire,* as well as group† and individual-subject‡ energy corrections based on doubly labeled water and physical activity logs, Maryland, spring 2001§

| Intake | Actual | Food frequency questionnaire | Doubly labeled water | | | Physical activity | | |
|--------------------------|--------------|------------------------------|----------------------|----------------|-----------------------------------|-------------------|----------------|--------------------------------|
| | | | Group | Individual | Body weight/doubly labeled water¶ | Group | Individual | Body weight/physical activity# |
| Carbohydrate | | | | | | | | |
| g/day | 344.5 (57.1) | 238.8** (78.8) | 303.3 (100.1) | 306.2** (65.5) | 306.2** (65.5) | 339.1 (111.9) | 339.2 (56.2) | 339.2 (56.2) |
| % of daily energy intake | 54.9 (4.7) | 48.8** (6.2) | 48.4** (16.0) | 48.7** (5.9) | 52.0 (6.1) | 48.3** (16.0) | 48.7** (5.9) | 53.8 (6.7) |
| Fat | | | | | | | | |
| g/day | 85.3 (14.6) | 83.7 (41.4) | 106.4 (52.6) | 103.2** (25.2) | 85.3 (21.9) | 118.9** (58.9) | 116.6** (35.5) | 85.3 (25.7) |
| % of daily energy intake | 30.7 (3.8) | 37.0** (6.2) | 38.2** (18.9) | 36.9** (6.1) | 32.6 (6.8) | 38.1** (18.9) | 36.9** (6.1) | 30.3 (7.2) |
| Protein | | | | | | | | |
| g/day | 87.6 (8.9) | 70.2** (23.2) | 89.2 (29.5) | 89.0 (14.9) | 89.0 (14.9) | 99.7 (33.0) | 100.6 (24.0) | 100.6 (24.0) |
| % of daily energy intake | 14.1 (2.0) | 14.5 (2.7) | 14.2 (4.7) | 14.4 (2.5) | 15.4** (3.0) | 14.2 (4.7) | 14.4 (2.5) | 15.9** (3.0) |
| Energy | | | | | | | | |
| MJ | 10.5 (1.3) | 8.28** (2.9) | 10.5 (0) | 10.5 (1.6) | 9.8** (1.5) | 11.8 (0) | 11.8 (2.0) | 10.6 (1.3) |

* Food frequency questionnaire filled out at the end of 16 weeks of ad libitum feedings.

† "Group": food frequency questionnaire intake adjusted with a single group mean.

‡ "Individual-subject": food frequency questionnaire intake adjusted with each individual's data.

§ Statistics (paired *t* test) are presented as the mean (standard deviation) for clarity.

¶ Food frequency questionnaire intake adjusted with each individual's body weight and doubly labeled water energy expenditure (correction was performed for fat intake only).

Food frequency questionnaire intake adjusted with each individual's body weight and physical activity logs (correction was performed for fat intake only).

** Different from actual: *p* < 0.05.

the doubly labeled water dosing period. Activities were recorded as a value from one to nine, corresponding to example activities listed in the log. Each activity assumed a predetermined energy expenditure score. Thus, energy expenditure was calculated as the time spent in that activity times the energy expenditure rate.

Food frequency questionnaire

The Block '98 Food Frequency Questionnaire (Berkeley Nutrition Services, Berkeley, California) was used to assess usual dietary intake during the study period. The eight-page scannable booklet asks for frequency (times per day, week, month) of consumption of 109 food items. Portion sizes were estimated with the aid of a "serving size choices" page containing photos of simple abstract three-dimensional models representing four different amounts of food. Gram amounts were then assigned on the basis of the gram weight of the volume of the chosen model, for that particular food.

During the final week of the study, subjects received instructions on how to fill out the food frequency questionnaire, which was then completed at home. Although the questionnaire was designed to capture usual intake during the past year, subjects were instead instructed to answer questions based on what they ate and drank during the study. Returned questionnaires were reviewed for omissions, implausible responses, and other errors. When necessary, forms were reviewed with the subject for clarity of answers. The completed food frequency questionnaires were sent to Berkeley Nutrition Services (now known as NutritionQuest,

Berkeley, California) for coding and analysis. Two of the questionnaires were "flagged" because the subjects reported consuming more than 17 food items per day (17.8 and 18.1 items). However, these values agreed with the actually measured number of food items. No other errors were detected.

This study was part of a broader investigation of the effects of covert manipulation on energy intake. To manipulate macronutrient composition, subjects consumed two of three drinks that provided a total of approximately 2.1 MJ/day of carbohydrate, protein/carbohydrate, and/or fat. Thus, solely for the purposes of comparing the food frequency questionnaire and actual intake with doubly labeled water and physical activity in terms of energy and macronutrient intake, 2.1 MJ were subtracted from doubly labeled water and physical activity (table 3). Subjects were unaware of the composition of the drinks and were instructed to not include them when filling out the food frequency questionnaire.

Data transformation

Preliminary analyses confirmed that the food intake data were right skewed and that variances increased as the means increased. A log transformation removed the dependency of the variance on the mean and created approximately normally distributed data (1). Although we present the data on the original scale in the tables and figure for ease of interpretation, the log scale was used to test for statistical significance.

Energy correction

Since it is well accepted that data derived from food frequency questionnaires are subject to measurement error and should be adjusted for energy intake (21) and differences in body size or physical activity (7), we based our adjustments on these variables. We applied two types of energy corrections: 1) group corrections, where a single adjustment factor based on the average of the subjects was used to correct individual subjects' food frequency questionnaires, and 2) individual corrections, where a subject-specific adjustment factor was developed and applied to each subject. Correction factors were created using energy intake estimated by either doubly labeled water or physical activity as the numerator and food frequency questionnaire intake as the denominator.

Our food frequency questionnaire measurement error analysis is based on the assumptions that 1) measurement error of actual food intake is negligible and 2) errors in the measurement of actual food intake and the food frequency questionnaire are unrelated. Potential errors in the first assumption are related to the misidentification of food items in the recording process, subjects' failing to report foods not directly measured in the Human Studies Facility, and database errors. Errors caused by misidentifying food items are assumed to be quite small, because any lack of agreement between the daily intakes recorded by hand and the bar-coding system was reconciled daily and again at the end of the study. Errors due to subjects' failing to consume all foods were minimized by direct observation of food intake in 10 of 21 weekly meals. The risk of violating the second assumption was minimized since actual food intake was not reported by the subjects (it was measured by the investigators), and there is little error in the actual food intake measure (assumption 1).

Bias analysis

We calculated the means and standard deviations for each of the three macronutrients and energy intake for actual intake, food frequency questionnaire, and the various food frequency questionnaire adjustments. We tested the reduction in bias (determining if there was a statistical difference with actual measurements) using paired *t* tests (pairing within subjects).

Variance analysis

To measure how much adjustments to the food frequency questionnaire reduce measurement error in macronutrient intake, we looked at reductions in the variance of the difference between actual measurements and food frequency questionnaire intakes averaged over the 12 subjects. Variances were calculated for the difference between actual measurements and unadjusted food frequency questionnaire and for group and individual corrections of the food frequency questionnaire by use of doubly labeled water and the combination of doubly labeled water and body weight. Variances were calculated as $\Sigma(\text{actual} - \text{FFQ}_{\text{corr}})^2/12$, where FFQ_{corr} = energy- and possibly body weight-

corrected food frequency questionnaire macronutrient intake. The same procedure was carried out for the variance calculations based on physical activity.

Body weight adjustment

In a preliminary analysis, we found that the size of the difference between actual intake and energy-corrected food frequency questionnaire fat intake was correlated with body weight ($r = 0.67$, $p < 0.02$). Thus, the energy-corrected food frequency questionnaire estimate of fat intake could be further improved. This additional adjustment for fat intake requires estimation of two regression coefficients and also makes use of the actual fat intake measurement, which is the target of our adjustments to the food frequency questionnaire and is ordinarily not available. Still, one can conceive of a smaller, preliminary study using a sample of the population of interest to estimate these coefficients, which can then be applied to data obtained on the larger sample. An alternative is to use results from another study with a similar population from which these coefficients could be calculated. In our study, we regressed body weight on the difference between actual fat intake and the energy-corrected food frequency questionnaire fat intake, where doubly labeled water or physical activity was used for the energy correction.

Statistical tests and analysis

Comparisons between mean macronutrient intake for actual intake and food frequency questionnaire (tests of bias) were made using paired *t* tests. Differences in (ratios of) variances were tested using *F* statistics. Changes in body weight and composition were measured in a mixed-linear models framework, using the Proc Mixed procedure in SAS, version 9.1, software (SAS Institute, Inc., Cary, North Carolina). Linear relations between variables were calculated using Pearson's product-moment correlation. Attenuation factors were calculated using equation 4 in the paper by Kipnis et al. (3).

RESULTS

Body weight and composition

During the doubly labeled water measurement periods, there was no effect of study week on body weight and composition (all: $p > 0.05$). Thus, it appears that the subjects were in energy balance during the doubly labeled water measurement periods and that energy intake and expenditure were equivalent.

Energy intake

Actual energy intake was 10.5 (standard deviation (SD): 1.3) MJ/day, compared with 8.3 (SD: 2.9) MJ/day as measured by food frequency questionnaire, 10.5 (SD: 1.6) MJ/day for doubly labeled water, and 11.8 (SD: 2.0) MJ/day for physical activity. Actual energy intake and that predicted by doubly labeled water were significantly correlated ($r = 0.75$, $p < 0.005$), but the correlations between food frequency questionnaire and actual intake ($r = 0.51$, $p > 0.09$) and

between food frequency questionnaire and doubly labeled water ($r = 0.48$, $p > 0.10$) were not significant. Energy intake, as measured by physical activity, was not correlated with actual intake ($r = 0.17$, $p > 0.60$), doubly labeled water ($r = 0.37$, $p > 0.20$), or food frequency questionnaire ($r = 0.20$, $p > 0.50$).

General bias and variance results

Absolute macronutrient intakes (g/day) predicted by food frequency questionnaire were significantly lower than actual intakes for carbohydrate and protein, but not for fat (table 3). Relative macronutrient intakes as a percentage of total energy were significantly different for carbohydrate and fat, but not for protein.

Group and individual energy adjustments to the food frequency questionnaire were equivalent corrections to bias. Including the body weight adjustment to fat intake reduced the fat bias to zero (as it must, since we used the actual fat intake measure to estimate the regression coefficients). While there was little overall bias in the unadjusted food frequency questionnaire measure of fat, variances were large.

The linear relations between food frequency questionnaire macronutrient and daily energy intakes were strong and significant, while the relations observed for actual intake were not as strong (figure 1). Figure 1A shows that the true increase in energy is due primarily to increases in carbohydrate intake. Unfortunately, this relation is not preserved using the food frequency questionnaire (figure 1B). Subjects with higher energy intake tended to overreport their intakes of protein and fat, relative to subjects with lower energy intake.

Effect of energy correction

None of the energy adjustment methods corrected all the errors observed in absolute energy intake (table 3). Doubly labeled water adjustment factors appeared to appropriately correct absolute protein intake, but they undercorrected carbohydrate intake and overcorrected fat intake unless the body weight correction was used. Physical activity adjustments overcorrected protein and fat intakes unless the body weight correction was used. Body weight was useful only for correcting fat intake (see below).

Applying energy correction factors tended to reduce variances, particularly for individual corrections (table 4). Using doubly labeled water, we found that there was a reduction for all macronutrient error variances. By use of physical activity, there was a substantial reduction in error variance for carbohydrate intake (overall, the correction removed about one third of the error variance).

Fat intake correction

The differences between actual intake and doubly labeled water- and physical activity-corrected macronutrients when regressed against body weight (BW) were significant for fat intake ($p < 0.02$ and $p < 0.01$, respectively) but not for carbohydrate or protein intake ($p > 0.05$). We estimated the regression coefficients using doubly labeled water as

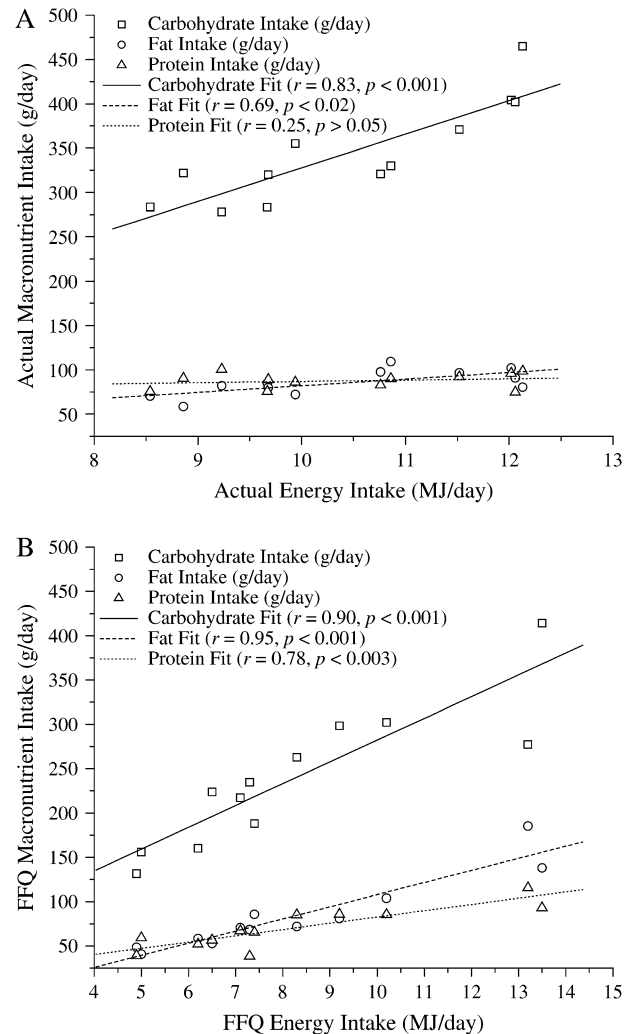


FIGURE 1. Relations between macronutrient intakes in 12 males plotted against total energy intake from ad libitum feedings (actual energy intake) (A) and food frequency questionnaire (FFQ) (B), Maryland, spring 2001. Linear relations between variables were calculated using Pearson's product-moment correlation, with significance of the difference from actual set at $p < 0.05$. "Carbohydrate fit" is the linear fit for carbohydrate intake (g/day) versus energy intake, "fat fit" is the linear fit for fat intake (g/day) versus energy intake, and "protein fit" is the linear fit for protein intake (g/day) versus energy intake.

the energy correction, with a mean of 675.72 (SE: 357.77) for the intercept and a mean of -13.18 (SE: 4.59) for the slope. For the physical activity energy adjustment, these means were 1,491.79 (SE: 525.88) for the intercept and -25.32 (SE: 6.75) for the slope. Thus, the estimate for the energy-corrected food frequency questionnaire fat intake is further improved using, for example, for physical activity, $F_{\text{new}} = 1,491.79 - 25.32 \times \text{BW} + F_{\text{PA}}$, where F is a measure of fat intake, and F_{PA} is corrected only for energy, using physical activity. These corrections reduce both the bias in (table 3) and variance of (table 4) the error.

TABLE 4. Variance* of the difference between measured macronutrient intakes and food frequency questionnaires† by 12 males, with corrections based on doubly labeled water and physical activity logs, Maryland, spring 2001‡

| Intake | Food frequency questionnaire | Doubly labeled water | | | Physical activity | | |
|--------------|------------------------------|----------------------|-------------|-----------------------------------|-------------------|-------------|---------------------------------|
| | | Group§ | Individual¶ | Body weight/doubly labeled water# | Group§ | Individual¶ | Body weight/physical activity** |
| Carbohydrate | 4.26 | 2.24 | 0.85†† | 0.85†† | 2.22 | 0.96†† | 0.96†† |
| Fat | 1.44 | 3.15 | 0.85 | 0.24†† | 4.86 | 2.49 | 0.52 |
| Protein | 0.18 | 0.17 | 0.03†† | 0.03†† | 0.25 | 0.19 | 0.19 |
| Sum | 5.89 | 5.55 | 1.73†† | 1.12†† | 7.33 | 3.63 | 1.67†† |

* Variance = (actual intake – food frequency questionnaire intake)²/12; intake is expressed as MJ/day.

† Filled out at the end of 16 weeks of ad libitum feedings.

‡ Statistics analyzed by *F* statistic.

§ “Group”: food frequency questionnaire intake adjusted with a single group mean.

¶ “Individual-subject”: food frequency questionnaire intake adjusted with each individual's data.

Food frequency questionnaire intake adjusted with each individual's body weight and doubly labeled water energy expenditure (correction was performed for fat intake only).

** Food frequency questionnaire intake adjusted with each individual's body weight and physical activity logs (correction was performed for fat intake only).

†† Different from food frequency questionnaire: *p* < 0.05.

Overall, corrections based on the combination of individual body weight and doubly labeled water/physical activity were the best combination for reducing variances, removing more than two thirds of the error for the physical activity correction (1 – 1.67/5.89) and even more for the doubly labeled water correction.

DISCUSSION

Similar to the results from other investigations (1, 9, 27), the food frequency questionnaire underestimated energy intake when compared with measured energy intake and that estimated from doubly labeled water. Applying group and individual correction factors based on energy tended to reduce the variance of the difference between the adjusted food frequency questionnaire and actual food intake, but these energy adjustments do not correct all the measurement error.

Considering the extensive documentation of measurement error in the food frequency questionnaire, the fact that intakes of energy and macronutrients were not accurate is not surprising. Proponents of the food frequency questionnaire have described methods for correcting for measurement error since the early 1990s (5) and have criticized authors who do not perform energy corrections on raw food frequency questionnaire data (21, 28). An interesting result of this study is that, not only are absolute and relative intakes not accurate, applying energy adjustment to correct the data, while helpful, is not entirely effective. In the case of fat intake, applying the energy adjustment actually increased error. The reason for the problem using energy corrections for our subjects can be observed using figure 1. Figure 1B indicates that the relations between food frequency questionnaire macronutrient and energy intakes in our subjects were linear and strong (*r* = 0.87 – 0.94). The strength of these relations forms the basis for using energy correction. However, figure 1A demonstrates that the relations between true macronutrient and energy intakes are

strong for carbohydrate only. Thus, applying a single adjustment factor to all macronutrients does not correct all the macronutrient measurement error.

Given our homogeneous study population relative to the multitude of factors that may affect the misreporting of food intake (gender, age, physical activity, education, and body composition), there was surprisingly large subject-to-subject variation in measurement error, so one would expect even greater intersubject variation in more heterogeneous populations. The lack of effectiveness for the group versus individual corrections illustrates this. Since our subjects were relatively homogeneous, the specific results of this study (e.g., parameter estimates) are not universally applicable. Nevertheless, the discrepancies between actual intake and the food frequency questionnaire seen in this study and the adjustment methods we used to bring the food frequency questionnaire into better agreement with the actual intake should be applicable to other populations. In particular, individual energy adjustments using physical activity and a body weight correction for fat intake removed more than two thirds of the error in the food frequency questionnaire.

Another issue to be considered is that food frequency questionnaires are not intended to be used “in studies with small numbers of subjects; for surveillance and monitoring of current levels where *accurate absolute intakes* are required; . . . and in some clinical work when precise intakes are required” (19, p. 568). However, if food frequency questionnaires do not produce accurate individual food intake estimates and introducing energy correction does not correct the errors, the usefulness of food frequency questionnaires to quantitatively measure the relation between disease and food intake in any size study is questionable. One other study utilized a similar approach to ours (comparison of a food frequency questionnaire with controlled feeding in a relatively small sample of 19 persons) (27), and both studies indicate that the food frequency questionnaire was not a reliable indicator of true intake. The agreement between

the statistics we calculated on our subjects and those given in Kipnis et al. (3) and Subar et al. (1) provides additional evidence that our conclusions are not unduly influenced by sample size.

Energy was underreported by about 21 percent (table 3), similar to 20 percent in 19 subjects consuming the “usual American diet” (27) and less than 31 percent in 261 male subjects (1, 3). The overall bias for protein intake in our study and that of Schaefer et al. (27) was small compared with that reported by Subar et al. (1). Fat and carbohydrate intakes were underreported by 30 percent and 10 percent, respectively, in the study by Schaefer et al. (27) compared with 2 percent and 30 percent in our study. We found an attenuation factor of 0.192 for energy, between the 0.080 and 0.230 reported by Kipnis et al. (3) (calculated using different models); and for protein, our value was 0.212, slightly greater than the 0.156 and 0.177 they calculated. We calculated the variance of true intake as 0.025 (kcal scale), compared with 0.026 and 0.044 for their two models (3). It is likely that the differences between these studies are related to the composition of the subject pools (such as age and body composition) and/or the methods used to obtain the criterion measure (type of monitored feeding or doubly labeled water).

In conclusion, applying energy adjustments to food frequency questionnaire data can significantly reduce overall measurement error, but this has the potential to increase errors for some macronutrients. Although individual doubly labeled water energy expenditure adjustments are clearly effective to correct overall energy intake, doubly labeled water is not practical in large epidemiologic studies. Assuming that good estimates of the regression coefficients that relate body weight to the error in fat intake are available, an additional questionnaire (similar to physical activity) given during the administration of food frequency questionnaires and body weight measurements would provide sufficient information to markedly reduce error in individual food frequency questionnaire estimates.

ACKNOWLEDGMENTS

Conflict of interest: none declared.

REFERENCES

- Subar AF, Kipnis V, Troiano RP, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* 2003;158:1–13.
- Kipnis V, Midthune D, Freedman L, et al. Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutr* 2002;5:915–23.
- Kipnis V, Subar AF, Midthune D, et al. Structure of dietary measurement error: results of the OPEN biomarker study (with discussion). *Am J Epidemiol* 2003;158:14–26.
- Trabulsi J, Schoeller DA. Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake. *Am J Physiol Endocrinol Metab* 2001;281:E891–9.
- Willett WC. *Nutritional epidemiology*. New York, NY: Oxford University Press, 1990.
- Spiegelman D, McDermott A, Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr* 1997;65(suppl):1179S–86S.
- Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies (with discussion). *Am J Clin Nutr* 1997;65:1220S–31S.
- Bonifaj C, Gerber M, Scali J, et al. Comparison of dietary assessment methods in a southern French population: use of weighed records, estimated-diet records and a food-frequency questionnaire. *Eur J Clin Nutr* 1997;51:217–31.
- Sawaya AL, Tucker K, Tsay R, et al. Evaluation of four methods for determining energy intake in young and older women: comparison with doubly labeled water measurements of total energy expenditure. *Am J Clin Nutr* 1996;63:491–9.
- Willett WC, Reynolds RD, Cottrell-Hoehner S, et al. Validation of a semi-quantitative food frequency questionnaire: comparison with a 1-year diet record. *J Am Diet Assoc* 1987;87:43–7.
- Rothenberg E. Validation of the food frequency questionnaire with the 4-day record method and analysis of 24-h urinary nitrogen. *Eur J Clin Nutr* 1994;48:725–35.
- Block G, Thompson FE, Hartman AM, et al. Comparison of two dietary questionnaires validated against multiple dietary records collected during a 1-year period. *J Am Diet Assoc* 1992;92:686–93.
- Feskanih D, Rimm EB, Giovannucci EL, et al. Reproducibility and validity of food intake measurements from a semi-quantitative food frequency questionnaire. *J Am Diet Assoc* 1993;93:790–6.
- Subar AF, Thompson FE, Kipnis V, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *Am J Epidemiol* 2001;154:1089–99.
- Johansson I, Hallmans G, Wikman A, et al. Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. *Public Health Nutr* 2002;5:487–96.
- Kroke A, Klipstein-Grobusch K, Voss S, et al. Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* 1999;70:439–47.
- Schoeller DA. How accurate is self-reported dietary energy intake? *Nutr Rev* 1990;48:373–9.
- Goris AH, Westerterp-Plantenga MS, Westerterp KR. Under-eating and underreporting of habitual food intake in obese men: selective underreporting of fat intake. *Am J Clin Nutr* 2000;71:130–4.
- Cade J, Thompson R, Burley V, et al. Development, validation and utilisation of food-frequency questionnaires—a review. *Public Health Nutr* 2002;5:567–87.
- Basiotis P, Welsh S, Cronin F, et al. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. *J Nutr* 1987;117:1638–41.
- Willett W. Commentary: dietary diaries versus food frequency questionnaires—a case of undigestible data. *Int J Epidemiol* 2001;30:317–19.
- US Department of Agriculture. USDA National Nutrient Database for Standard Reference, release 15. Washington,

- DC: US Government Printing Office, 2002. (<http://www.nal.usda.gov/fnic/foodcomp/Data/SR15/sr15.html>).
23. Schoeller DA. Validation of habitual energy intake. *Public Health Nutr* 2002;5:883–8.
 24. Speakman J. Doubly labelled water: theory and practice. London, United Kingdom: Chapman and Hall, 1997.
 25. Weir JB. New methods for calculating metabolic rate with special reference to protein metabolism. 1949. *Nutrition* 1990;6:213–21.
 26. Bouchard C, Tremblay A, Leblanc C, et al. A method to assess energy expenditure in children and adults. *Am J Clin Nutr* 1983;37:461–7.
 27. Schaefer EJ, Augustin JL, Schaefer MM, et al. Lack of efficacy of a food-frequency questionnaire in assessing dietary macronutrient intakes in subjects consuming diets of known composition. *Am J Clin Nutr* 2000;71:746–51.
 28. Willett W. Invited commentary: OPEN questions. *Am J Epidemiol* 2003;158:22–4.