Comparison of Food Frequency Questionnaires: The Reduced Block and Willett Questionnaires Differ in Ranking on Nutrient Intakes

A. K. Elisabet Wirfalt,1 Robert W. Jeffery,2 and Patricia J. Elmer2

Food frequency questionnaires, major tools in epidemiologic studies, are often criticized for biased and imprecise intake estimates. The aim of this study was to compare the performance of two widely used food frequency questionnaires, a reduced 60-item Block questionnaire and a 153-item Willett food frequency questionnaire, relative to three 24-hour recalls administered by telephone. The dietary data were collected in 1991 from a group of healthy women age 25-49 years (n = 101) during the baseline period of a weight-loss intervention study in Minneapolis, Minnesota. Total energy and macro- and micronutrient intakes were compared across methods by using four analytic approaches: comparison of means and correlation coefficients, regression analysis, and estimation of percent agreement between each questionnaire and recalls. The Block instrument showed an overall underestimation bias, but was more successful in categorizing individuals on percent energy from fat and carbohydrate intakes than was the Willett instrument. The Willett instrument showed no overall underestimation bias and was more successful in classifying individuals on vitamin A and calcium intakes. Diverging performance characteristics of diet assessment methods have an implication for the design of studies, interpretation of results, and comparison of findings across studies. Am J Epidemiol 1998;148:1148-56.

bias (epidemiology); diet; epidemiologic methods; nutrition assessment; nutrition surveys; questionnaires

Food frequency questionnaires, which have become the major tools in epidemiology studies, are easy to administer and inexpensive to process. These instruments are, however, commonly criticized for imprecise and biased estimates (1–3), which may contribute to the failure of epidemiologic studies to show significant associations between diet and disease (4–6). Food frequency questionnaires are used not only in studies examining the relations between diet and disease, but also in intervention studies to assess the composition of total diet and dietary change and in observational studies to compare dietary intakes between groups (7–10). The role of food frequency questionnaires in such quantitative assessment of diet is, however, debated (11, 12).

Two widely used diet assessment instruments are the Willett and the Block food frequency questionnaires. The Willett questionnaire (also known as the Health Habits and History Questionnaire) was intended for slightly broader use, i.e., to measure an individual’s relative nutrient intake as well as absolute nutrient values, to assess both nutrients and foods, and to assess diets of a variety of demographic groups (14). Studies examining the concurrent validity of these instruments relative to repeated dietary records have concluded that they are able to rank individuals on specific nutrient intakes (13–16). The Willett food frequency questionnaire was first presented as a 61-item instrument (13), but longer versions are now also available (15, 17). The original Block food frequency questionnaire included 98 food items and additional questions about food preparation practices (14). A reduced 60-item version of this instrument was later developed (16).

The current controversies concerning food frequency questionnaires and the wide use of the Block and Willett instruments in different types of studies warrant comparisons of the two types of questionnaires for their ability to assess usual diet. A substudy of the Women’s Health Trial showed a higher correlation of fat intakes between the Block instrument and multiple diet records than between the Willett instru-
event and diet records, but no information on other nutrient intakes was presented (10). Another study comparing the two instruments indicated that closed-ended response categories caused nondifferential misclassification and recommended that food frequency questionnaires should obtain exact frequencies of intake when measuring diet exposure (18).

The purpose of this study was to compare a reduced 60-item Block (14, 16) and a 153-item Willett food frequency questionnaire (13, 15, 17) relative to three 24-hour recalls, administrated to a group of healthy females aged 25–49 years on total energy, macro- and micronutrient intakes, and percent of energy from macronutrients. The paper describes 1) the comparison of absolute means between each questionnaire and recalls, 2) the examination of correlations and comparison of correlation coefficients between each questionnaire and recalls, 3) the comparison of methods by regressing the nutrient intake estimates from each food frequency questionnaire on estimates from recalls, 4) the examination of the percent agreement between each questionnaire and recalls after cross-classification of individuals by quintiles of intake distributions, and 5) the examination of energy underreporting of each method. The implications of the observed differences are discussed.

MATERIALS AND METHODS

Sample and data collection

Participant information. This study was conducted as part of a weight-loss intervention program in a group of women aged 25–49 years and used dietary data collected with three assessment methods during the baseline period. The dietary assessment methods were a reduced 60-item Block food frequency questionnaire, a 153-item Willett food frequency questionnaire, and three 24-hour recalls administered by telephone. Subjects were recruited from the community through two newspaper advertisements. Participants were screened by telephone, and those who were interested and eligible were invited to an initial clinic visit. Eligibility criteria for the weight-loss intervention included an initial weight of 120–140 percent of ideal body weight, being healthy, being a nonsmoker, not being pregnant or lactating, and drinking fewer than 20 alcoholic drinks per week. During the clinic visit, information on age, education, type of work, and marital status was collected by using a standardized self-administered questionnaire. Anthropometric data were collected through direct measurements by trained personnel. A total of 122 women attended the first visit. Most were married and well educated, and the majority were white. Participants completed the two food frequency questionnaires at the beginning of the 5-week baseline period. The three 24-hour recalls were collected when all other data were complete, i.e., during the final 2 weeks of the baseline period. For this methodological study, complete data from the three diet assessment methods were available from 101 women.

The Block food frequency questionnaire. The reduced, 60-item version of the Block semiquantitative food frequency questionnaire (16) was sent to each participant prior to the clinic visit with instructions to complete it and return it at the initial clinic visit. The reference period was dietary intake during the previous 6 months. The 60-item questionnaire used in this study was in the original nonscannable format. This questionnaire version requests that respondents estimate their daily, weekly, monthly, yearly, or rarely/never consumption frequencies by indicating the exact number of times each food was eaten per day, per week, etc. Respondents were also asked to indicate whether their usual portion sizes were small, medium, or large compared with a standard. All completed questionnaires were checked by the study coordinator for accuracy and completeness. Incomplete sections were reviewed with the individual participants during the clinic visit. Daily intakes of energy and nutrients were estimated by multiplying frequency responses with the specified portion sizes and the nutrient values assigned to each food item in the nutrient database. No information on dietary supplements was collected in this study when administering the Block instrument. The dietary analysis program (DIETANAL) and food database, Version 2.2, were provided by the National Cancer Institute, Bethesda, Maryland.

The Willett food frequency questionnaire. The Willett instrument was given to participants at the end of the initial clinic visit along with a self-addressed, stamped return envelope. Participants were asked to complete this questionnaire during the following week and mail it back to the center. The study used a 153-item version of the Willett semiquantitative food frequency questionnaire, which had been expanded (for other studies examining the relation between diet and colon cancer) to include additional vegetables, fruits, and low-fat foods (18). The reference period was 1 month. This questionnaire has one standard portion size indicated for each food item. Respondents were asked to indicate their relative frequency of consumption from nine different response alternatives, ranging from never or less than once per month to six or more times per day. Daily nutrient intakes were estimated by multiplying frequency responses by the nutrient compositions of the specified portion size of each food using the Willett nutrient and database pro-
gram (HarvardSSFQ.5/93; Harvard School of Public Health, Boston, Massachusetts). For purposes of this methodological study, nutrient calculations for the Willett food frequency questionnaire included only nutrients from foods and not from supplements. All completed questionnaires were checked by the study coordinator for accuracy and completeness. Incomplete sections were reviewed with the individual participants via telephone.

24-hour recalls. The referent method of this methodological study was three 24-hour recalls. Recalls were collected via telephone using the Minnesota Nutrient Data System developed by the Nutrition Coordinating Center (NCC) at the University of Minnesota. This is an interactive data collection and automatic coding system, which uses the food descriptions and exact portion sizes solicited from the respondent during a 24-hour recall. The system leads the interviewer through a four-step process: food identification, quantity estimation, ingredient specification, and preparation method. The data collection process is followed by nutrient calculation (19). The NCC food-nutrient database is linked to the Nutrient Data System and contains about 16,000 foods. The food and nutrient information is obtained from the US Department of Agriculture, manufacturers, scientific literature, and other sources (20, 21).

Each participant received three unannounced telephone calls during a 2-week period after the initial clinic visit. Interviews were conducted on randomly selected days and included two weekdays and one weekend day. Participants were provided with a chart of two-dimensional food portions, developed to aid in the estimation of portion sizes during telephone interviews (22, 23). The interviews were conducted by nutritionists at the NCC. Recalls gathered by telephone, which are used routinely at the NCC to collect 24-hour recalls, have the advantage of being unannounced and facilitate an even distribution of recalls over different weekdays (24, 25). Studies comparing recalls gathered by telephone with diet records and in-person interviews have concluded that recalls are preferable to consecutive diet records when monitoring change in dietary intervention studies and that they give estimates of acceptable validity in population surveys (26).

Analysis

Energy and nutrient intakes were computed separately for the recalls, the reduced Block food frequency questionnaire, and the Willett food frequency questionnaire. The nutrient intake variables used in this study were total energy (kcal/MJ); total fat (g, percent energy); monounsaturated (g), polyunsaturated (g) and saturated (g) fats; protein (g and percent energy); carbohydrate (g and percent energy); alcohol (g and percent energy); cholesterol (mg); fiber (g); calcium (mg); iron (mg); vitamin A (IU); and vitamin C (mg). Prior to analysis, two measures were taken to improve normality of the data. First, one individual who reported a very high energy intake (7,684 kcal/31.7 MJ) on the Block instrument was excluded from analysis. Second, the value of the alcohol variable was set to missing in one individual who reported an extremely high consumption.

Comparison of means. The mean intakes of total energy and macro- and micronutrients were compared across the three diet assessment methods by using one-way analysis of variance and the general linear model procedure in the Statistical Analysis System, with the model sums of squares for the overall F test, and Tukey’s Studentized range test for pairwise comparisons (27). Square root-transformed data were used for significance testing of mean differences.

Comparison of correlation coefficients. Correlations of total energy and macro- and micronutrient intakes were examined between the Willett food frequency questionnaire and three 24-hour recalls ($r_w$), between the reduced Block food frequency questionnaire and the three 24-hour recalls ($r_B$), and between the two food frequency questionnaires ($r_{WB}$) by computing the Pearson product-moment correlation coefficients using the CORR procedure in SAS (27). Testing for significant differences between the two dependent correlation coefficients, $r_w$ and $r_B$, was performed using Student’s t test and the $t$ statistic obtained by the formula (28):

$$t = \frac{(r_w - r_B)(n - 3)(I + r_{WB})}{\sqrt{2(I - r_w^2) - r_B^2 - r_{WB}^2 + 2(r_{WB}r_B)}}.$$

Dietary data are commonly energy adjusted in epidemiology studies. In this study, the correlational analysis was conducted both on untransformed and on energy-adjusted and square root-transformed data to illustrate the utility of the two food frequency questionnaires in disease etiology studies. The energy-adjustment method using residuals from regressing nutrient intakes on total energy was used (29).

Regression analysis. The interchangeability of methods at group level was examined by regressing energy and nutrient estimates from each food frequency questionnaire on corresponding estimates from three 24-hour recalls (30, 31) by using the REG procedure in SAS (27). The linear association between methods (i.e., $\beta \neq 0$) was tested, and the correspondence between methods (i.e., $\beta = 1$) was examined by computing the 95 percent confidence intervals of the regression coefficients.
Relative agreement between methods. The relative agreement between methods was assessed by comparing the relative rankings of individuals on total energy and on macro- and micronutrient intakes between the two food frequency questionnaires and between each questionnaire and 24-hour recalls through cross-classification by quintiles. The percent agreement between methods was calculated.

Underreporting of energy. To examine the degree of energy underreporting in this population, we calculated the ratio of reported energy intake (EI) to computed basal metabolic rate (BMR) (EI/BMRcomp). A ratio of 1.35 has been suggested as the lower cutoff for reasonable habitual energy intake (32). The BMRcomp was computed by using the Equation for Prediction of basal metabolic rate identified by the World Health Organization for women aged 30–60 years: \[8.7 \times \text{weight (kg)} - [25 \times \text{height (m)}] + 865\] (33).

RESULTS
Comparison of means

The results of the comparison of mean energy and nutrient intakes are shown in table 1. When compared with the three 24-hour recalls, the reduced Block instrument showed an overall negative bias. The instrument underestimated mean total energy intake by 21 percent. Nutrient intakes were not uniformly affected, but carbohydrates, absolute amounts of fat, fiber, and iron were significantly lower than were the recalls. Percent energy from fat, protein, cholesterol, vitamins A and C, and alcohol were, however, similar between the Block instrument and the recalls.

The Willett instrument was similar to the 24-hour recalls on mean energy and several nutrient intakes. However, percent energy from protein, fiber, and vitamins A and C were significantly overestimated, and saturated, polyunsaturated, and monounsaturated fats were underestimated.

Correlations and comparison of correlation coefficients

The correlations between energy and nutrient estimates from the three assessment methods are shown in table 2. Nutrient correlations between the two food frequency questionnaires were overall more similar to one another than were those between either questionnaire and recalls. The median values of \(r_{WB}, r_B, r_W\) for unadjusted energy and nutrient variables were 0.60, 0.34, and 0.36, respectively. The findings also suggest that correlations with the three 24-hour recalls were stronger for fat and carbohydrate intakes with the reduced Block instrument and for vitamin A and calcium intakes with the Willett instrument.

Formal evaluation of the differences between the correlation coefficients, \(r_B\) and \(r_W\), indicate that correlations between the reduced Block instrument and 24-hour recalls were higher for percent of energy from fat and from carbohydrate. The Willett instrument showed higher correlations with 24-hour recalls for vitamin A and calcium intakes. Similar observations

<table>
<thead>
<tr>
<th>Nutrients</th>
<th>24-hour recalls</th>
<th>Willett questionnaire</th>
<th>Block questionnaire</th>
<th>(F_{test}) (p value)</th>
<th>Pairwise comparisons†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy (kcal)</td>
<td>2,059 (545)</td>
<td>1,936 (657)</td>
<td>1,603 (913)</td>
<td>0.0001 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Fat (g)</td>
<td>80 (30)</td>
<td>73 (30)</td>
<td>69 (53)</td>
<td>0.0001 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Fat (% energy)</td>
<td>34 (6.7)</td>
<td>38 (23.8)</td>
<td>38 (7.4)</td>
<td>0.180 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Carbohydrate (g)</td>
<td>255 (77)</td>
<td>264 (89)</td>
<td>175 (87)</td>
<td>0.0001 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Carbohydrate (% energy)</td>
<td>50 (7.3)</td>
<td>55 (30)</td>
<td>45 (7.9)</td>
<td>0.0001 WQ &gt; 24HR &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Protein (g)</td>
<td>77 (21)</td>
<td>79 (30)</td>
<td>77 (21)</td>
<td>0.932 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Protein (% energy)</td>
<td>15 (2.9)</td>
<td>16 (2.6)</td>
<td>17 (3.3)</td>
<td>0.0002 WQ, BQ &gt; 24HR</td>
<td></td>
</tr>
<tr>
<td>Alcohol (g)</td>
<td>5.0 (9.9)</td>
<td>4.1 (5.7)</td>
<td>4.5 (13.3)</td>
<td>0.647 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Alcohol (% energy)</td>
<td>1.6 (2.8)</td>
<td>1.5 (2.0)</td>
<td>1.8 (3.9)</td>
<td>0.187 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Cholesterol (mg)</td>
<td>247 (131)</td>
<td>209 (110)</td>
<td>244 (144)</td>
<td>0.039 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Dietary fiber (g)</td>
<td>18 (6.4)</td>
<td>23 (9.5)</td>
<td>10 (5.1)</td>
<td>0.0001 WQ &gt; 24HR &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Saturated fat (g)</td>
<td>28.6 (11.6)</td>
<td>22.3 (12.8)</td>
<td>24.5 (20.8)</td>
<td>0.0005 24HR &gt; WQ, BQ</td>
<td></td>
</tr>
<tr>
<td>Monounsaturated fat (g)</td>
<td>29.6 (11.9)</td>
<td>23.0 (12.5)</td>
<td>24.5 (20.7)</td>
<td>0.0002 24HR &gt; WQ, BQ</td>
<td></td>
</tr>
<tr>
<td>Polyunsaturated fat (g)</td>
<td>16.0 (6.6)</td>
<td>12.5 (6.7)</td>
<td>13.8 (8.2)</td>
<td>0.0004 24HR &gt; WQ, BQ</td>
<td></td>
</tr>
<tr>
<td>Vitamin A (IU)</td>
<td>7,039 (6,585)</td>
<td>11,004 (8,154)</td>
<td>8,544 (6,100)</td>
<td>0.0001 WQ &gt; 24BQ</td>
<td></td>
</tr>
<tr>
<td>Vitamin C (mg)</td>
<td>102 (66)</td>
<td>136 (67)</td>
<td>97 (56)</td>
<td>0.0001 WQ &gt; 24HR, BQ</td>
<td></td>
</tr>
<tr>
<td>Iron (mg)</td>
<td>16 (7.8)</td>
<td>15 (6.8)</td>
<td>11 (6.3)</td>
<td>0.0001 24HR, WQ &lt; BQ</td>
<td></td>
</tr>
<tr>
<td>Calcium (mg)</td>
<td>804 (368)</td>
<td>843 (374)</td>
<td>804 (368)</td>
<td>0.704 24HR = WQ &lt; BQ</td>
<td></td>
</tr>
</tbody>
</table>

† Tukey's test at \(\alpha = 0.05\).
were made for energy-adjusted nutrient variables. Neither of the observed differences remained significant when the significance level was adjusted for multiple comparisons.

**Regression analysis**

The nutrient variables with the strongest linear association for the reduced Block instrument were total fat and protein, nutrient densities of fat and carbohydrate, and monounsaturated fat (the β-coefficients for these nutrients ranged from 0.542 to 0.471). The Willett instrument showed the strongest linear associations for vitamin A, total fat and protein, calcium, and dietary fiber (the β-coefficients for these nutrients ranged from 0.651 to 0.466). Total correspondence (i.e., confidence intervals of the β-coefficients including one) between nutrient estimates from food frequency questionnaires with 24-hour recalls was not seen for any of the examined nutrient variables or instruments.

**Quintile ranking and cross-classification**

The ability of the food frequency questionnaires to distinguish individuals with high and low energy and nutrient intakes was also examined by cross-classification on intake categories and is shown in table 3. The quintile rankings appeared to be more similar overall between the two food frequency questionnaires than between either questionnaire and 24-hour recalls. The median exact agreement was 38 percent between the two food frequency questionnaires, 26 percent between the reduced Block questionnaire and 24-hour recalls, and 25 percent between the Willett questionnaire and 24-hour recalls. However, the Block instrument appeared to rank individuals more successfully on percent energy from fat, while ranking on calcium appeared more successful with the Willett instrument. Gross dissimilarity, i.e., individuals ranked in the first quintile with one and in the fifth with the other, was similar overall between the two food frequency questionnaires and 24-hour recalls and was highest for alcohol with both questionnaires.

**Underreporting of energy**

Examination of the EI/BMR\textsubscript{comp} ratios indicated a tendency of underreporting with all three dietary assessment methods, but to varying degrees. The mean ratios were 1.35 with 24-hour recalls, 1.28 with the Willett questionnaire, and 1.05 with the reduced Block.
A recent study indicated that when a large number of food items represent a food group on a questionnaire, the total frequency estimate from that food group will be inflated (34). The higher estimates of fiber and vitamins A and C with the Willett instrument in this study may likely be due to the large number of fruits and vegetables included in this instrument. This overestimation with the Willett instrument and the underestimation of fiber with the reduced Block instrument relative to 24-hour recalls indicate that the estimated food sources of carbohydrates may be different with the two questionnaires compared with the recall method. Thus, even if absolute amounts of, and the percent of energy from, macronutrients were similar with the Willett instrument compared with recalls, the type of carbohydrates, and thus the meaning of relative macronutrient measures, may be substantially different across methods.

**Underreporting of energy or instrument bias?**

Underreporting of energy intake is an ongoing concern in studies using self-reports in assessment of usual diet, especially among overweight individuals. The well-educated participants of this study appear to underreport their energy intake to a lesser degree compared with US nationals of similar body composition. The National Health and Nutrition Examination Survey III collected 24-hour recalls and found a EI/BMR ratio of 1.09 among overweight women (25). Other studies have indicated that although obese individuals have a tendency to underreport energy, it is not equally common in all socioeconomic strata (35) and appears to affect quantity of diet more than quality of diet (36).

Although the current debate appears to focus on underreporting of energy in certain population subgroups, other aspects of dietary misreports may be of equal or greater importance. Misreports of usual diet are also related to instrument formats (34, 37-41) and do not appear to affect all nutrients uniformly (13, 14, 37, 39, 40).

Food frequency questionnaires are commonly used in studies examining the association between diet and disease. Therefore, the ability of the instrument to rank individuals on usual nutrient intakes should be the issue of concern rather than its ability to estimate absolute intakes at the group level. Methodological studies have indicated that specific instrument formats may be related both to over- and underreports of food consumption (39, 40) and to the ranking ability on foods and specific nutrients (37, 41). Misreports in dietary assessment should, therefore, not only be seen as a characteristic of specific population groups, but also as a consequence of the methodology used. In this study, the Block instrument showed a negative bias of

---

**DISCUSSION**

**Total diet assessment and nutrient density**

This study examined the comparability of a reduced, 60-item Block questionnaire and a 153-item Willett food frequency questionnaire in usual diet assessment. The ability of a food frequency questionnaire to estimate absolute mean nutrient intakes has implications for the ability of an instrument to describe the composition of total diet. Nutrition studies commonly use nutrient density measures, such as percent of total energy from macronutrients, which require total diet assessment. These complex measures not only depend on the assessment of one particular macronutrient, but also on how well the others are assessed. Previous studies have indicated that both the longer (98-item) and the shorter (60-item) versions of the Block instrument produced similar estimates of percent energy from fat compared with multiple diet records when questions adjusting for both fruits and vegetables and age- and gender-adjusted portion sizes were used (14, 16). Exclusion of these adjustments in this study probably contributed to the overestimation of percent energy from fat with the Block instrument.
energy and absolute nutrient intakes and the lowest EI/BMRcomp, indicating a method effect typical of short food frequency questionnaires (16, 42).

Ranking ability

The different abilities of the two questionnaires to distinguish individuals on low- and high-fat intakes have previously been observed in a substudy of the Women’s Health Trial (10). That is, correlations of percent energy from fat with the Block instrument compared with four diet records were significantly larger than estimates from the Willett instrument. No information is available on how other nutrient estimates performed in that study. The Willett instrument has, in a large number of studies, been able to detect relations between different types of chronic diseases, specific micronutrients, and food groups and, thus, has indicated the utility of the instrument in public health research (43). Those findings may also illustrate the capability of the instrument in distinguishing different levels of micronutrient exposures (i.e., from plant and dairy foods, in particular) and that a sufficient number and quality of food sources are required to categorize individuals on these nutrients. However, other assessment methods could still be more successful in categorizing individuals on different levels of fat intake.

Previous studies have indicated that frequency of consumption is the major factor influencing the ranking of individuals on nutrient intakes (42, 44) and that portion size has a minor influence. However, by using one standard portion size when calculating nutrients, mean consumptions are generally exaggerated (45). Some of the interindividual variation is also removed (46), and, therefore, the discriminating ability is decreased. This study implies that different portion size alternatives may be more important when ranking individuals on fat intake than on vitamin A and calcium intake. Researchers validating other dietary assessment tools have made similar observations (47). In addition, the differences in ranking ability for fat in this study may partly be due to differences in frequency-response alternatives between instruments. One study has indicated that the nondifferential misclassification is reduced with open-ended frequency-response alternatives (17).

The choice of referent method may be a likely cause of the low correlations seen between the two questionnaires and the referent method in this study. Research has indicated that 7–14 daily records are needed to rank individuals on energy and most nutrients, but that vitamin A may require as many as 21 (48). However, the precision of food frequency questionnaires has been estimated to be comparable with that of three dietary recalls or records (49), and three recalls are commonly used in nutrition studies to categorize usual diets of individuals (50). This study had the advantage of being able to use data collected for another study, and thus, the performance of three widely used dietary assessment methods could be examined in a real world setting.

Interchangeability of methods

The apparent similarities in ranking ability between the two food frequency questionnaires could be due to several factors. The two food frequency questionnaires were in this study administered 1 week apart at the beginning of the 5-week assessment period, and the three recalls were administered during the last 2 weeks of this period. The order of data collection is recommended in studies calibrating or validating assessment methods of usual diet (46), but a weakness is that the different reference periods may result in differences in the actual foods consumed. Cognitive research indicates that episodic dietary memories decay rapidly and that the information reported when answering food frequency questions comes largely from generic memories of personal diet with some influence of short-term episodic memories (51). In a 24-hour recall interview, however, episodic memories from a very recent and limited time period are recalled. That is, the dietary memory retrieved when answering food frequency questionnaire items and when recalling information from the last 24-hours are different in nature. Diet reports obtained with questionnaires would, therefore, be more similar to one another than to reports from recalls. In this study, it is likely that both cognitive processing and actual food consumption differences may have contributed to the observed differences between food frequency questionnaires and recalls.

In summary, this study suggests that the specific design features of food frequency questionnaires do not affect intake estimates of macro- and micronutrients uniformly. The results imply that questionnaires, using open-ended frequency response alternatives and different portion, sizes, may more clearly distinguish between different fat and carbohydrate intakes. Less-detailed information may be sufficient when categorizing individuals on nutrients present in a few foods consumed less frequently, providing that the questionnaire includes sufficient number and quality of food sources.

The study highlights the potential danger of using dietary assessment methods for purposes other than those identified through the validation procedures. The performance characteristics of an instrument are likely to vary with specific design features and between population subgroups. Researchers therefore need to
recognize the ranges and limitations of different dietary assessment methods when evaluating the choice of methodology and interpreting results and comparing findings across studies. Future studies need to examine how the effect of specific designs varies across foods, nutrients, and population subgroups and how questionnaires should be improved to discriminate dietary exposures categories better. Further refinement and validation of current methodologies could help us explain inconsistencies across studies and to reevaluate the magnitude of relative risks and potential biases.

ACKNOWLEDGMENTS

Supported by National Institute of Diabetes and Digestive and Kidney Disorders grant DK42201 and National Cancer Institute grant P01 CA50305.

REFERENCES


