Validity and Reproducibility of a Food Frequency Questionnaire by Cognition in an Older Biracial Sample

Martha Clare Morris¹,²,³, Christine C. Tangney⁴, Julia L. Bienias¹,², Denis A. Evans¹,²,⁵, and Robert S. Wilson¹,⁶,⁷

¹ Rush Institute for Healthy Aging, Rush University Medical Center, Chicago, IL.  
² Department of Internal Medicine, Rush University Medical Center, Chicago, IL.  
³ Department of Preventive Medicine, Rush University Medical Center, Chicago, IL.  
⁴ Department of Clinical Nutrition, Rush University Medical Center, Chicago, IL.  
⁵ Rush Alzheimer’s Disease Center, Rush University Medical Center, Chicago, IL.  
⁶ Department of Psychiatry, Rush University Medical Center, Chicago, IL.  
⁷ Department of Neurological Sciences, Rush University Medical Center, Chicago, IL.

Received for publication January 24, 2003; accepted for publication July 11, 2003.

It is not currently known how dietary assessment in older persons is affected by cognition. In a 1997–2000 study, the authors assessed the validity and reproducibility of a modified Harvard self-administered food frequency questionnaire (SFFQ) among 118 Black and 114 White randomly selected participants, aged 68–99 years, of the Chicago Health and Aging Project. Participants completed multiple 24-hour dietary recall interviews (mean = 3.6) over 12 months and two SFFQs in the first and 12th months. The average energy-adjusted intraclass correlation coefficient for 15 nutrients was 0.59 for 1-year reproducibility in nutrient intake levels assessed by the SFFQ. The average energy-adjusted Pearson correlation coefficient was 0.46 for comparative validity between nutrient intake levels on the SFFQ and the dietary recalls. SFFQ reproducibility was higher among men, and comparative validity with the dietary recalls was higher among women. There were no remarkable differences in the correlations by age, race, educational level, presence of chronic conditions, or cognitive ability.

The modified Harvard SFFQ is a reasonable method of dietary assessment even in a population of older persons, some of whom are at advanced age, have chronic health conditions, and have cognitive impairment.

Abbreviations: MMSE, Mini-Mental State Examination; SFFQ, self-administered food frequency questionnaire.
in-home interviews with 6,158 residents aged 65 years or older (79 percent of age-eligible residents), and follow-up interviews at 3- and 6-year intervals from 1997 to 2002. Dietary questionnaires were distributed to study participants for completion and mail return. The validity and reproducibility of the questionnaire were assessed in 1997–2000 in random samples selected by race from the study population. The 12-month study included self-administration of two food frequency questionnaires in the first and 12th months and six 24-hour dietary recall interviews conducted at 2-month intervals, although the average number completed per person was 3.6. Of 490 surviving residents selected in the sample, 36 were judged by the interviewer or family members to be too demented to participate, 201 declined, and 13 had incomplete dietary information. The remaining 232 persons (114 Whites and 118 Blacks) completed at least one SFFQ and one 24-hour dietary recall for validity analysis. Of these, 192 persons completed two SFFQs for reproducibility analysis. The Institutional Review Board of Rush University Medical Center approved the study, and all participants gave written consent.

Food frequency questionnaire

We used a modified version of the Harvard food frequency questionnaire that measured usual intake in the past year of 139 food items and vitamin and mineral supplements (2, 3). We selected this questionnaire over others because of its simplicity for self-administration by older persons, some of whom have limited formal education. The SFFQs were optically scanned and analyzed for daily nutrient intake using the Harvard nutrient database. For some food items, natural portion sizes (e.g., one banana) were used to determine nutrient content. Otherwise, nutrient content was based on the mean portion sizes reported by the oldest participants of national surveys (4–6) and other sources (7). We computed the daily intake of each dietary component by multiplying the nutrient content of the food item by the reported frequency of intake and summing over all food items. Nutrient values were modified on the basis of specified types of cereal, oil, and margarine, the fat content of dairy products, and the removal of meat fat or poultry skin.

24-hour dietary recalls

Dietary recall interviews were conducted in participants’ homes using computerized version 2.92 of the Minnesota Nutritional Data Systems (8). The period of dietary recall included the 24 hours up to the time of the interview, which was scheduled 1–2 hours in advance. Quantities were measured using the participant’s own dishware and standard measuring tools. A registered dietitian reviewed all dietary recalls. For analysis, we used the averaged daily nutrient intake over all recalls.

Covariates

Information on all covariates was obtained from the Chicago Health and Aging Project population interviews conducted in 3-year cycles from 1993 to 2002. Age was computed from the self-reported date of birth and date of initial participation in the validation study. Participants were categorized as having a chronic condition if they reported a history of myocardial infarction, stroke, cancer, or diabetes at the population interviews up to and including the interview closest in date to the 12-month SFFQ. Cognitive scores were obtained from the population interview closest in date to the 12-month SFFQ. The tests included the East Boston Tests of Immediate Memory (score range: 0–12) and Delayed Memory (score range: 0–12) (9, 10), the Mini-Mental State Examination (score range: 0–30) (11), and the Symbol Digit Modalities Test (score range: 0–110) (12). A global cognitive score was computed as the averaged z scores of the four tests. The cognitive testing occurred a median of 4.4 months before the 12-month SFFQ (interquartile range: 20.4 months).

Statistical methods

Nutrient intake levels from both the SFFQ and the averaged dietary recalls were adjusted for total energy intake using the regression residual method (13) separately for males and females. All nutrient variables were log transformed to improve normality. Intraclass correlation coefficients were used to compare nutrient intake levels derived from the two SFFQs, and Pearson’s correlation coefficients were used to compare nutrient intake levels estimated from the SFFQ and the averaged 24-hour dietary recalls. Tests for equality of correlations were performed using Fisher’s z transformation (14). Multiple regression was used to compute partial correlations adjusted for age, sex, and race. Tertiles of the cognitive test scores were computed separately by race.

RESULTS

The stratified random sample consisted of 118 Blacks and 114 Whites. The average age of participants was 77.8 years (range: 68–99 years), with 12 percent over the age of 85 years (table 1). There were 145 females (62 percent) and 87 males. Persons with 0–8 years of formal education represented 16 percent of the sample. Nearly one-fifth of the sample (18.1 percent) had Mini-Mental State Examination (MMSE) scores below 25, a range generally considered to indicate cognitive impairment; no persons had MMSE scores below 14. A large portion of the sample (60 percent) reported at least one major chronic condition (myocardial infarction, stroke, cancer, or diabetes). The validation sample was reasonably representative of the total Chicago Health and Aging Project study population aged 68 years or older at baseline (table 1).

The SFFQ appeared to have good 1-year reproducibility. The intraclass correlations ranged from 0.70 for folate to 0.50 for vitamin B12, with an average of 0.59 over all 15 dietary components (table 2). Pearson’s correlations between levels of nutrient intake estimated by the SFFQ and dietary recalls ranged from 0.67 for vitamin E with supplements to 0.31 for protein, with an average of 0.46 for the 15 dietary components. The correlations were comparable after adjustment for age, sex, and race (average r = 0.43) (table 2).
Reproducibility and validity by cognition and other factors

Neither the intraclass correlations for SFFQ reproducibility nor the Pearson correlations between the SFFQ and the dietary recalls varied across tertiles of the combined global score of the four cognitive tests (table 2). For example, the average Pearson correlations over the 15 nutrients from lowest to highest tertiles of cognitive score were 0.48, 0.44, and 0.45, respectively (table 3). When we repeated the analyses after excluding 36 persons whose cognitive tests were administered more than 1.5 years before...
the first 24-hour recall, there were no appreciable differences in the validity \((r = 0.47, 0.44, 0.46\) for tertiles 1–3, respectively) or reproducibility \((\text{intraclass correlation coefficient} = 0.54, 0.61, 0.63\) for tertiles 1–3, respectively) correlations.

There were no remarkable differences in the reproducibility or comparative validity correlations by age (68–78 years compared with 79 years or older), race, educational level (0–8, 9–12, and ≥13 years), or presence of a chronic condition (stroke, myocardial infarction, diabetes, or cancer) (table 3). The men had lower correlations for comparative validity and higher correlations for reproducibility than the women. None of the group comparisons of the average correlations was statistically significant.

**DISCUSSION**

In this random sample of an older biracial community population, we found moderately sized correlations for nutrient intake levels assessed by a modified self-administered Harvard SFFQ and 24-hour dietary recalls and higher correlations for reproducibility between two SFFQ assessments 1 year apart. There were no substantial differences in the reproducibility and validity correlations by cognitive ability, age, race, education, or presence of a chronic condition, although the validity correlations were lower in the men and 1-year reproducibility was higher than in the women.

To our knowledge, this is the only study to report on the validity and reproducibility of dietary intake by level of cognition. Remarkably, the correlations for comparative validity were similar across level of cognitive score even though 18 percent of the sample had MMSE scores indicative of cognitive impairment. The study findings cannot be generalized to persons with severe cognitive impairment as they were not represented in the validity sample, but it is unlikely that many studies would seek response from this small subset of community residents. In the Chicago Health and Aging Project study, only 5 percent of the population aged 65 years or older had MMSE scores below 14, the lowest MMSE score in the validation sample.

The observed correlations in the present study were based on randomly selected participants of an urban community in which the SFFQs were self-administered. There are a limited number of validity and reproducibility studies of food frequency questionnaires in older populations, and these generally assessed interviewer-administered SFFQs in volunteer or special populations (e.g., trial participants) and over short periods of assessment. For example, most examined reproducibility within 3 months and reported higher average correlations (range: 0.62–0.80) than observed in the Chicago Health and Aging Project study \((\text{intraclass correlation coefficient} = 0.59)\) (15–17). Some food frequency questionnaire validity studies of older persons were based on comparative assessment methods conducted within 1 month or less of the food frequency questionnaire, and the correlations ranged from 0.49 to 0.70 (15, 16, 18). Two studies (17, 19) examined intake over 1 year and reported correlations comparable with that of the White women \((r = 0.52)\) in the Chicago Health and Aging Project study. Few validity studies included Black participants (none of older age), and these found lower correlations for Blacks compared with Whites (20–22). We did not observe differences in the correlations by race or by educational level, which may have been because of the simplified SFFQ format.

In the absence of a true “gold standard,” we used an average of 3.6 of the 24-hour dietary recall interviews to assess the comparative validity of the SFFQ, which may have resulted in both overestimation (due to correlated recall error) and underestimation (due to measurement error of the 24-hour recalls) of the correlations with the SFFQ. Whereas a greater number of recalls likely would have resulted in higher validity correlations, this would have increased the burden on these very old participants. Most previous validation studies of older persons used from 7 to 15 days to estimate intake levels, which may partially account for their higher correlations (16, 17, 19).

Another possible source of measurement error of the 24-hour recall method in our older sample is the level of motivation required for sustained attention to minute details in dietary intake. This may explain the lower validity correlations among the men in our older sample if they had less interest in the dietary recall interviews. We did not find a

**TABLE 3. Average intraclass correlations over 15 nutrients* for reproducibility of nutrient intake levels measured by two SFFQs† at 12-month intervals and Pearson’s correlations for comparative validity between the SFFQ and 24-hour dietary recall interviews by tertile of global cognitive score, age, race, gender, educational level, and presence of a chronic condition, Chicago, Illinois, 1997–2000**

<table>
<thead>
<tr>
<th>Group</th>
<th>SFFQ-SFFQ</th>
<th>SFFQ-dietary recall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC†</td>
<td>No. of persons</td>
</tr>
<tr>
<td>Global cognitive score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowest tertile</td>
<td>0.54</td>
<td>63</td>
</tr>
<tr>
<td>Middle tertile</td>
<td>0.62</td>
<td>65</td>
</tr>
<tr>
<td>Highest tertile</td>
<td>0.60</td>
<td>64</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>68–78</td>
<td>0.60</td>
<td>106</td>
</tr>
<tr>
<td>≥79</td>
<td>0.54</td>
<td>86</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.54</td>
<td>97</td>
</tr>
<tr>
<td>White</td>
<td>0.62</td>
<td>95</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>0.55</td>
<td>118</td>
</tr>
<tr>
<td>Males</td>
<td>0.64</td>
<td>74</td>
</tr>
<tr>
<td>Education (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–8</td>
<td>0.63</td>
<td>31</td>
</tr>
<tr>
<td>9–12</td>
<td>0.55</td>
<td>77</td>
</tr>
<tr>
<td>≥13</td>
<td>0.60</td>
<td>84</td>
</tr>
<tr>
<td>Chronic conditions‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>0.57</td>
<td>110</td>
</tr>
<tr>
<td>None</td>
<td>0.60</td>
<td>82</td>
</tr>
</tbody>
</table>

* The average correlation is for the 15 nutrients listed in table 2.
† SFFQ, self-administered food frequency questionnaire; ICC, intraclass correlation coefficient.
‡ Chronic conditions include reported history of myocardial infarction, stroke, cancer, or diabetes.
difference in the correlations by sex between the SFFQ and biochemical nutrient levels, which were comparable to or higher than those reported by previous studies (Christine Tangney, Rush University Medical Center, unpublished manuscript).

The self-administered food frequency questionnaire is probably the best method of dietary assessment for use in large prospective studies. Other dietary assessment methods, such as multiple days of diet interviews or tissue sampling, are not only burdensome and invasive to older participants but also very costly. These data indicate that the modified Harvard SFFQ is a reliable and reasonably valid method of assessing overall diet in an older biracial community population.

ACKNOWLEDGMENTS

This work was supported by grants (AG13170 and AG11101) from the National Institute on Aging.

The authors gratefully acknowledge the work of study coordinators Cheryl Bibbs, Michelle Bos, Jennifer Tarpey, and Flavio Lamorticella, their staffs, and analytical programmer Woojeong Bang. The authors also acknowledge the help of Dr. Graham Colditz and Helaine Rockett from the Harvard Channing Laboratory.

REFERENCES