Comparison of a picture-sort food-frequency questionnaire with 24-hour dietary recalls in an elderly Utah population

Heidi J Wengreen[1,]*, Ronald G Munger[1], Siew Sun Wong[1], Nancy A West[1] and Richard Cutler[2]
[1]Utah State University, Department of Nutrition and Food Sciences, Logan, UT 84322-4450, USA:
[2]Utah State University, Department of Statistics and Mathematics, Logan, UT 84322, USA

Submitted 17 October 2000: Accepted 5 February 2001

Abstract
Objective: To evaluate the 137-item Utah Picture-sort Food-frequency Questionnaire (FFQ) in the measurement of usual dietary intake in older adults.
Design: The picture-sort FFQ was administered at baseline and again one year later. Three seasonal 24-hour dietary recall interviews were collected during the year between the two FFQs. Mean nutrient intakes were compared between methods and between administrations of the FFQ.
Setting: The FFQ interviews were administered in respondents' homes or care-centres. The 24-hour diet recalls were conducted by telephone interview on random days of the week.
Subjects: Two-hundred-and-eight men and women aged 55–84 years were recruited by random sample of controls from a case–control study of nutrition and bone health in Utah.
Results: After adjustment for total energy intake, median Spearman rank correlation coefficients between the two picture-sort FFQs were 0.69 for men aged ≤69 years, 0.66 for men aged >69 years; and 0.68 for women aged ≤69 years, 0.67 for women aged >69 years. Median correlation coefficients between methods were 0.50 for men ≤69 years old, 0.52 for men >69 years old; 0.55 for women ≤69 years old, 0.46 for women >69 years old.
Conclusions: We report intake correlations between methods and administrations comparable to those reported in the literature for traditional paper-and-pencil FFQs and one other picture-sort method of FFQ. This dietary assessment method may improve ease and accuracy of response in this and other populations with low literacy levels, poor memory skill, impaired hearing, or poor vision.

Keywords
Food-frequency questionnaire
Dietary assessment
Ageing
Epidemiology

Food-frequency questionnaires (FFQs) are commonly used in epidemiological research on diet and disease to rank individuals according to usual nutrient intake[1]. Customarily, using a paper and pencil, FFQ respondents are asked to report frequency of consumption for a list of foods[2]. However, low literacy levels along with common cognitive and physical limitations, such as poor memory and impaired vision and hearing, are often encountered when studying elderly populations and may limit the use of traditional FFQs in this population[3]. Kumanyika *et al.*[4] developed a picture-sort method of administering a food-frequency questionnaire that may help to improve accuracy of response in elderly populations by easing respondent burden. We constructed a picture-sort FFQ, based on the method developed by Kumanyika *et al.*, for use in the Utah Study of Nutrition and Bone Health (USNBH), a large state-wide case–control study of the determinants of hip fracture in elderly Utah residents.

The aim of the present study was to evaluate the 137-item Utah Picture-sort Food-frequency Questionnaire in the measurement of usual dietary intake in older adults. We tested the relative validity of the Utah Picture-sort FFQ by comparing nutrient intakes reported from the FFQ with a measure of usual dietary intake. Although dietary recalls and FFQs, both self-report methods, are likely to have some degree of correlated error, we used the average of three 24-hour dietary recall interviews, collected by telephone interview using a multiple-pass format, as a measure of usual dietary intake. This method seemed like a reasonable option to the more widely used but burdensome dietary records, especially when considering the elderly population of interest, in which we

*Corresponding author: Email hwengreen@cc.usu.edu

may encounter poor literacy skills or low motivation. To test the reproducibility of the method we compared reported nutrient intakes from two administrations of the picture-sort FFQ taken one year apart.

Although the FFQ has been widely used in diverse populations of middle-aged adults, its use in the elderly, especially in respondents over the age of 70 years, is relatively unexplored[5–7]. We examined differences in reports of nutrient intakes using the picture-sort FFQ between younger and older elderly participants by comparing agreement between nutrient estimates made from the test and reference method by gender and age strata (men and women ≤69 years of age, men and women >69 years of age).

Materials and methods

Subjects and design

The Utah Picture-sort FFQ was developed for use in the USNBH study, which includes Utah residents with and without hip fractures aged 50 to 89 years. Participants in the dietary assessment study were selected by random sample from controls of the USNBH study and were invited to participate in the dietary study after completion of the USNBH baseline interview.

Respondents completed two picture-sort FFQ interviews and three 24-hour dietary recall interviews over the course of approximately one year. Respondents began the study by completing their first FFQ between March 1998 and March 1999 (Fig. 1) at the time of their baseline interview for the USNBH. The three 24-hour dietary recall interviews were collected during the year between the first administration of the FFQ and the second administration of the FFQ. During the FFQ interview, the respondent was asked to report their frequency of consumption of selected foods during the previous year. Because of this reference time, the second FFQ was expected to give a measure of usual dietary intake in the same period of time the three 24-hour recalls were collected.

The picture-sort FFQs were administered in the home of each respondent by a trained interviewer; the interview lasted approximately 35 minutes. Trained interviewers also administered the multiple-pass 24-hour dietary recall interviews by telephone. The telephone 24-hour recall interviews lasted approximately 25 minutes each.

Picture-sort food-frequency questionnaire

The Utah Picture-sort FFQ food list includes 138 food cards containing one or more food items that were systematically selected from the 126-item Nurses' Health Study (NHS) FFQ[8], the National Cancer Institute (NCI) FFQ[9], and a list of commonly eaten foods as identified in focus groups. A modified version of the 126-item NHS FFQ used in the Iowa Women's Health Study was administered to approximately 4500 persons in Cache Valley of Northern Utah. Food items that contributed less than 1.0% to total mean intake and had an R^2 factor of less than 0.01 for each of 25 key nutrients, indicating that a food item explained less than 1% of the variation in intake of 25 key nutrients, were dropped from the subsequent list. This list was then merged with the NCI FFQ food list and a list of commonly eaten foods as identified from focus groups. The final food list includes nearly all items from the NCI list and the NHS list, but in slightly different categories that may be more inclusive or more specific than those of other dietary assessment tools.

The picture-sort FFQ method, first developed by Kumanyika *et al.* for use in the Cardiovascular Health Study[4,10], is an adaptation of a written FFQ that engages respondents by having them sort colour picture cards into trays representing frequency of use. The Utah Picture-sort FFQ cards developed at Utah State University were laminated 4 inch by 6 inch cards with colour photographs of one or more food items. For ease in respondent sorting, the cards were separated into seven categories: (1) beverages, (2) fruit, (3) vegetables, (4) dairy products,

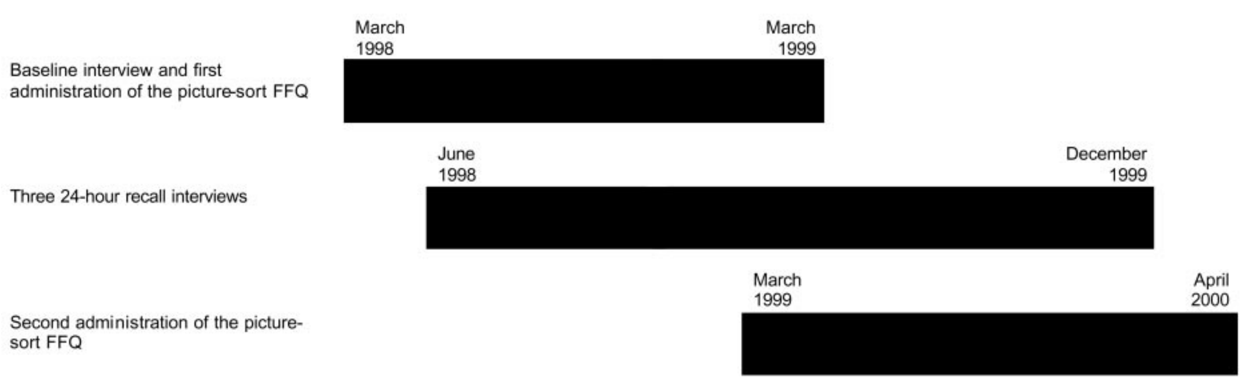

Fig. 1 Sequence of dietary interviews of controls from the Utah Study of Nutrition and Bone Health who participated in the dietary assessment study; March 1998 to April 2000

(5) meats, fish, egg and main dishes, (6) cereals and breads, and (7) snacks, and oils and other foods. In most cases the picture card depicted a standard size portion of the food or foods placed on a 10 inch dinner plate, in a 6 inch cereal bowl, or in a 12 ounce glass. Standard portion sizes were also listed on the reverse side of the card in common household measures (1 cup, 1 medium piece, etc.) for interviewer reference.

Following the picture-sort method as described by Kumanyika *et al*., each respondent was given categories of picture cards one at a time and asked to sort them into five trays according to period of use (i.e. daily, weekly, monthly, yearly or rarely/never) over the past year. After all categories of picture cards were sorted, respondents were asked about specific frequency of use for food items placed in the trays starting with the daily tray and proceeding to the rarely/never tray.

We did not ascertain information about vitamin and mineral supplement intake as part of the picture-sort FFQ method. The Utah Picture-sort FFQ was designed to ascertain information about the frequency of consumption of a list of foods only. In practice, the picture-sort should be combined with additional questions about the use of nutritional supplements, food preparation, and other additional questions that may be pertinent to the research aims.

24-hour recalls

The three 24-hour dietary recall interviews were administered approximately three months apart during the year between the two FFQ interviews to capture seasonal differences in intake. All 24-hour recalls were administered on random days of the week and included at least one weekend day (Saturday or Sunday). A two-dimensional representation of the National Health and Nutrition Examination Survey three-dimensional food models[11] was mailed to respondents before their first interview. Participants were instructed to keep the guide near their telephone for use in the interviews. Telephone interviewers used an adaptation to the multiple-pass dietary recall collection method originally developed by the US Department of Agriculture (USDA)–Human Nutrition Information Service (HNIS)[12] to obtain information about a respondent's food intake during the preceding day. Interviewers were trained to probe for detailed and accurate information on food preparation, food type and portion size for each food reported during the recall interview.

Dietary recall interviews were edited for completeness, coded, and entered into the Food Processor dietary assessment program (Food Processor Nutrition Analysis & Fitness Software, Version 7.1, ESHA Research, Salem, OR). The Food Processor program contained nutrient data from USDA as well as additional sources. Mean individual nutrient intakes were calculated from the average of the three dietary recalls. A nutrient database developed for the picture-sort FFQ food list also used the Food Processor program. Food selection for food cards that listed more than one food were made by selecting the most frequently reported food in that category from the coded 24-hour dietary recall interviews. A registered dietitian (H.W.) performed this coding and selection process. FFQ data were converted to mean daily intakes by transforming all consumption periods to consumption per day and then multiplying by frequency of use and weight of the standard serving size.

Statistical analysis

Means and standard deviations for energy and nutrient intake from food and beverages, excluding nutrition supplements, estimated by each of the two picture-sort FFQs and by the average of the three 24-hour dietary recalls, were calculated. Nutrient intakes were adjusted for total energy intake using the residual method of Willett and Stampfer[13].

Agreement between administrations of the FFQ and between methods was assessed by calculating Spearman rank and Pearson product–moment correlation coefficients between the two picture-sort FFQs for both crude and energy-adjusted nutrient intakes. We report Spearman rank correlation coefficients as they correlate rank order of estimated nutrient intakes between methods and between administrations without assuming a Gaussian distribution for the nutrient intake in the population. Means, standard deviations, and ratio of means for total energy and nutrient intakes estimated by each of the FFQs and the average of the three-recalls were also calculated. Means were compared by paired *t*-tests.

In evaluating the agreement between methods, the second FFQ was compared with the average of the three recalls. We used the second FFQ to calculate correlations because this gave a measure of intake from the same period of time the three 24-hour recall interviews were obtained.

To assess variability in daily intake, intraclass correlation coefficients for nutrient estimates were calculated from the three 24-hour dietary recall interviews[2,14]. A two-way random effects analysis of variance model was fit to the data and the intraclass correlation coefficient was calculated as the ratio of between-subject variance to total variance.

All data were analysed separately by gender and by the four age–gender strata in Table 1. Analyses were performed using SPSS for Windows (Version 10, SPSS Inc., Chicago, IL).

Results

Of those respondents who agreed to participate in the dietary assessment study, 217 (75% of both men and women) completed all dietary interviews. Reasons for dropping out of the dietary assessment study included

Table 1 Selected characteristics of participants in the Utah dietary assessment study; March 1998 to April 2000

	Men		Women	
Characteristic	≤69 years (*n* = 51)	>69 years (*n* = 52)	≤69 years (*n* = 54)	>69 years (*n* = 51)
Age (years)	63* (4.5)†	77 (4.4)	61 (4.6)	77 (4.2)
BMI (kg m^{-2})	28 (5)	27 (4.0)	28 (5.2)	27 (4.4)
Weight (kg)	89 (16)	81 (12.5)	73 (13.5)	69 (12.8)
Height (cm)	180 (6.1)	175 (8.6)	163 (6.1)	160 (6.5)
White, not of hispanic origin (%)	96‡	100	91	96
High school graduate (%)	90	79	91	94
College graduate (%)	35	27	15	18
Ever taken oestrogen (%)	–	–	81	59
Currently taking oestrogen (%)	–	–	59	37
Ever regularly smoked cigarettes (%)	57	42	28	12
Currently smokes cigarettes (%)	14	2	9	2
Ever regularly drank alcohol (%)	61	48	30	18
Currently drinks alcohol (%)	25	15	17	6
Currently takes a multivitamin/mineral supplement (%)	43	54	56	69
Currently takes herbal preparations (%)	27	25	33	41

* Mean value.
† Standard deviation.
‡ Percentage of total participants in each sub-group.

refusal, death, severe illness, or interviewer inability to locate respondents. Because participants were selected from a random sample of controls who had previously completed an interview for the USNBH, they may represent a population of people more willing to participate in health-related studies than the general ageing population of Utah.

A registered dietitian (H.W.) reviewed the FFQs and 24-hour diet recalls of poor quality, as identified by comments made by the interviewer at the time of interview. Nine participants (seven men, two women) with implausible reported usual energy intake (<500 or >3500 kcal) by either of the two FFQs or by the average of the 24-hour recalls were excluded from analyses. Characteristics of the remaining 208 participants (103 men, 105 women) are reported in Table 1.

Study participants were grouped according to age and gender. The mean age of study participants was 69 years. Participants who were 69 years old or younger at the baseline interview were labelled as younger elderly men or women, and those greater than 69 years old at the baseline interview were labelled as older elderly men or women.

Most (>90%) participants in all age–gender groups were white, not of Hispanic origin, and a high percentage of respondents (>78%) in all age–gender categories had at least a high school education. Body mass index (BMI, kg m^{-2}) based on self-reported heights and weights were similar for both men and women, and older participants had lower BMI than younger respondents. Fewer women than men, and fewer older participants than younger participants, reported ever regularly or currently smoking cigarettes, or ever regularly or currently drinking alcohol. A greater percentage of women than men, and of older women than younger women, reported currently taking multivitamin mineral supplements or herbal preparations (Table 1).

We compared mean energy and nutrient intakes from food and beverages for macronutrients and selected micronutrients from the second FFQ and the average of three 24-hour dietary recalls by age–gender strata (Table 2). Mean nutrient intake estimates from our 137-item picture-sort FFQ were generally higher than nutrient intake estimates from traditional and picture-sort FFQs in other elderly populations[10,15,16]. However, in the studies referred to above, the food lists were shorter which may have an influence on the estimated intakes measured from the method. Mean intake estimates from the FFQ were consistently higher than mean intake estimates from the 24-hour dietary recalls; however, the degree of difference varied by age–gender strata. Several investigators have also reported greater mean intakes from FFQs as compared with reference methods[6,17,18].

Differences in mean intakes are reported as ratios of mean nutrient intake estimates from the FFQ to nutrient intake estimates from the 24-hour dietary recalls, and are included in Table 2. The median ratio of mean reported intake was 1.26 in younger men, 1.25 in older men, 1.23 in younger women and 1.22 in older women.

Mean energy-adjusted nutrient intakes estimated from food and beverages from the first FFQ (administered at baseline) and the second FFQ (administered approximately one year later) were calculated. In general, lower mean energy-adjusted nutrient intakes were estimated by the second FFQ as compared with the first FFQ for all age–gender groups except younger men. Mean energy-adjusted intake estimates for younger men, older men, and older women from the first FFQ and the second FFQ were not significantly different from each other for most

https://doi.org/10.1079/PHN2001131 Published online by Cambridge University Press

Table 2 Means and standard deviations of nutrient intakes estimated from the second administration of the Utah Picture-sort Food-frequency Questionnaire (FFQ2) and means of three 24-hour dietary recalls (24-h R) by age–gender group; March 1998 to December 1999

	Males						Females					
	≤69 years			>69 years			≤69 years			>69 years		
	FFQ2	24-h R	Ratio*	FFQ2	24-h R	Ratio	FFQ2	24-h R	Ratio	FFQ2	24-h R	Ratio
Calories (kcal)	2586 (960)†	2187 (620)	1.18	2273 (749)	1883 (443)	1.21	1920 (858)	1605 (440)	1.20	1960 (486)	1520 (338)	1.29
Percentage of total energy intake as:												
Carbohydrate	46	49	0.94	49	52	0.94	48	51	0.94	51	53	0.96
Fat	38	35	1.09	37	34	1.09	36	34	1.06	36	33	1.09
Protein	16	16	1.00	15	16	0.94	16	16	1.00	15	16	0.94
Carbohydrate (g)	299 (109)	270 (100)	1.11	280 (88)	245 (65)	1.14	231 (75)	203 (64)	1.14	252 (74)	202 (56)	1.25
Fat (g)	111 (51)	85 (29)	1.31	95 (38)	72 (26)	1.32	78 (29)	61 (24)	1.28	79 (22)	55 (19)	1.44
Protein (g)	103 (39)	89 (27)	1.16	86 (37)	75 (32)	1.15	81 (30)	65 (18)	1.25	74 (22)	61 (16)	1.21
Fibre (g)	20 (8)	17 (7)	1.18	22 (8)	20 (11)	1.10	17 (8)	14 (5)	1.21	20 (8)	16 (8)	1.25
Cholesterol (mg)	372 (197)	322 (159)	1.16	309 (176)	271 (142)	1.14	263 (122)	220 (102)	1.20	242 (85)	190 (90)	1.27
Water (g)	3485 (1380)	2448 (906)	1.42	2837 (818)	1989 (591)	1.43	2996 (1046)	2109 (692)	1.42	2840 (675)	1915 (722)	1.48
Vitamin A (IU)	9424 (4820)	6651 (4272)	1.42	9935 (5411)	8355 (8344)	1.19	9550 (6095)	8724 (7211)	1.09	10 096 (4297)	8668 (6128)	1.16
Vitamin D (IU)	423 (1057)	225 (174)	1.88	272 (175)	160 (102)	1.70	235 (125)	139 (82)	1.69	257 (121)	170 (91)	1.51
Vitamin C (mg)	119 (65)	111 (80)	1.07	138 (57)	115 (66)	1.20	127 (80)	92 (50)	1.38	146 (78)	109 (46)	1.34
Calcium (mg)	1136 (630)	893 (454)	1.27	1013 (485)	814 (378)	1.24	982 (419)	721 (322)	1.36	981 (370)	752 (278)	1.30
Alcohol (g)	4 (16.6)	2.7 (12.2)	1.48	1.1 (4.1)	0.99 (4.6)	1.11	1.8 (6.4)	2.4 (8.4)	0.75	0.10 (0.32)	0.3 (1.7)	0.33
Caffeine (mg)	205 (234)	142 (301)	1.44	72 (102)	34 (61)	2.12	185 (457)	112 (192)	1.65	71 (115)	41 (88)	1.73

* Ratio of mean FFQ to mean 24-hour dietary recall.
† Standard deviation.

Table 3 Ratios of mean nutrient intakes estimated from the first administration of the Utah Picture-sort Food-frequency Questionnaire (FFQ1) and the second administration of the Utah Picture-sort Food-frequency Questionnaire (FFQ2). Spearman rank correlation coefficients (*r*) comparing energy and nutrient intake estimates from FFQ1 and FFQ2. All nutrients except calories and alcohol were adjusted for total energy intake

	Men				Women			
	≤69 years		>69 years		≤69 years		>69 years	
	Ratio†	*r*	Ratio	*r*	Ratio	*r*	Ratio	*r*
Calories (kcal)	1.05	0.65	0.96	0.58	0.89***	0.72	0.95	0.70
Carbohydrates (g)	1.02	0.61	0.99	0.66	0.87**	0.61	0.95	0.57
Fat (g)	1.07	0.64	0.95	0.70	0.91*	0.72	0.96	0.67
Protein (g)	1.05	0.55	0.94	0.70	0.92***	0.26	0.91*	0.44
Fibre (g)	1.00	0.74	0.96	0.72	0.83**	0.63	0.94*	0.75
Cholesterol (mg)	1.03	0.69	0.96	0.77	0.87*	0.48	0.98	0.64
Water (g)	1.08	0.70	1.01	0.51	0.95*	0.78	0.97	0.67
Vitamin A (IU)	0.94	0.70	0.90	0.50	0.83*	0.66	0.86*	0.46
Vitamin D (IU)	0.99	0.61	0.88***	0.58	0.90	0.68	0.94	0.69
Vitamin C (mg)	0.95	0.68	0.93	0.67	0.86*	0.50	0.93	0.62
Calcium (mg)	1.04	0.69	0.95	0.66	0.89**	0.75	0.94	0.60
Alcohol (g)	1.34	0.96	0.75	0.95	1.03	0.82	0.24	0.87
Caffeine (mg)	1.11	0.88	1.05	0.73	0.96	0.88	0.90	0.85

† Mean nutrient estimates from the second administration of the FFQ/mean nutrient estimates from the first administration of the FFQ.
Statistically significant different estimated mean intakes: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. Correlations above 0.26 are significantly greater than 0 at $P = 0.05$.

nutrients. However, in younger women, mean energy-adjusted intake estimates from the first FFQ and the second FFQ were considerably different for most nutrients, and 24 of 28 nutrients tested and 10 of 13 nutrients shown in Table 3 showed statistically significant differences with $P < 0.05$.

Table 3 includes Spearman rank correlation coefficients of selected nutrient intakes estimated from the first FFQ and the second FFQ. Pearson product–moment correlations were also calculated, with similar results. Spearman rank correlation coefficients for energy-adjusted nutrients ranged from 0.26 to 0.96 for all age–gender groups. Younger men had higher correlations for nine of 13 nutrients when compared with older men, and younger women had higher correlations for seven of 13 nutrients. Median energy-adjusted Spearman rank correlation coefficients for all nutrients by gender and age–gender groups were 0.69 for younger men (range 0.55–0.96), 0.67 for older men (range 0.50–0.95), 0.68 for younger women (range 0.26–0.88) and 0.67 for older women (range 0.44–0.87).

Table 4 includes Spearman rank order correlations of selected nutrient intakes estimated from the second FFQ and the average of the three 24-hour dietary recall interviews for both crude and energy-adjusted nutrient intakes. Pearson product–moment correlations were also calculated, with similar results.

Table 4 Intraclass correlation coefficients (Intracl) of three 24-hour dietary recall interviews and Spearman rank correlation coefficients comparing energy and nutrient intake estimates from the 24-hour dietary recall interviews and the second administration of the Utah Picture-sort Food-Frequency Questionnaire (FFQ2). Correlations reported for crude and energy-adjusted (Adj) nutrient intake

	Males						Females					
	≤69 years			>69 years			≤69 years			>69 years		
	Intracl	Crude	Adj*	Intracl	Crude	Adj	Intracl	Crude	Adj	Intracl	Crude	Adj
Calories (kcal)	0.42	0.28		0.29	0.11		0.47	0.04		0.24	0.15	
Carbohydrate (g)	0.47	0.32	0.40	0.32	0.28	0.53	0.42	0.06	0.55	0.40	0.20	0.41
Fat (g)	0.25	0.43	0.32	0.30	0.08	0.53	0.43	0.11	0.48	0.15	0.30	0.46
Protein (g)	0.24	0.09	0.30	0.07	0.29	0.50	0.40	0.02	0.41	0.15	0.21	0.01
Fibre (g)	0.44	0.35	0.50	0.01	0.39	0.44	0.33	0.12	0.39	0.45	0.54	0.55
Cholesterol (mg)	0.16	0.34	0.52	0.17	0.34	0.52	0.17	0.28	0.36	0.03	0.06	0.15
Water (g)	0.30	0.51	0.63	0.30	0.53	0.56	0.37	0.55	0.71	0.31	0.43	0.58
Vitamin A (IU)	0.10	0.22	0.33	0.16	0.20	0.15	0.20	0.06	0.32	0.06	0.35	0.27
Vitamin D (IU)	0.02	0.41	0.32	0.51	0.60	0.51	0.35	0.52	0.67	0.12	0.48	0.51
Vitamin C (mg)	0.32	0.34	0.51	0.35	0.40	0.43	0.22	0.54	0.61	0.03	0.42	0.49
Calcium (mg)	0.36	0.29	0.70	0.50	0.59	0.62	0.48	0.41	0.70	0.34	0.32	0.33
Alcohol (g)	0.62	0.64	0.36	0.46	0.48	0.26	0.59	0.75	0.50	0.30	0.43	0.22
Caffeine (mg)	0.15	0.68	0.71	0.68	0.57	0.57	0.51	0.81	0.83	0.77	0.69	0.70

* Adjusted for total energy intake using the method described by Willet and Stampfer[13].
Correlations above 0.27 are significantly greater than 0 at $P \leq 0.05$.

https://doi.org/10.1079/PHN2001131 Published online by Cambridge University Press

Intraclass correlation coefficients of the three 24-hour dietary recall interviews were calculated (Table 4). Intraclass correlation coefficients of the three 24-hour dietary recall interviews were similar for all age–gender groups and were somewhat lower than anticipated, indicating more variability in the three days of dietary recalls than might be expected. Ranges of intraclass correlations for all age–gender groups were 0.01–0.47 for macronutrients, fibre and cholesterol; 0.30–0.77 for caffeine, alcohol and water; 0.02–0.51 for micronutrients. Intraclass correlations were lowest for fibre in older men (0.01) and highest for caffeine in older women (0.77).

Adjusting nutrient intake by the regression method reported by Willett and Stampfer[13], to obtain a measure of nutrient intake independent of total caloric intake, improved correlation coefficients for most nutrients in all age–gender groups. Energy-adjusted correlation coefficients ranged from 0.01 to 0.55 for macronutrients, fibre and cholesterol; 0.22 to 0.83 for caffeine, alcohol and water; and 0.15 to 0.70 for micronutrients. Younger men and older men had similar energy-adjusted correlations (median = 0.50, range = 0.30–0.71 for younger men; median=0.52, range=0.15–0.62 for older men). However, younger women had stronger energy-adjusted correlations than older women (median = 0.55, range = 0.32–0.83 in younger women; median = 0.46, range = 0.01–0.70 for older women). Lower energy-adjusted correlations, defined here as correlations less than 0.30, were identified in older men for vitamin A and alcohol, and in older women for protein, cholesterol, vitamin A and alcohol. No energy-adjusted correlation coefficients less than 0.30 were identified in either younger men or younger women.

Discussion

In this study, three 24-hour dietary recalls and two picture-sort FFQs were used to assess the ability of the FFQ to discriminate among individuals' usual dietary intake, with reproducible results from an elderly population in Utah. It is well understood that no known dietary assessment method gives a perfect measure of diet, however some methods are considered more accurate than others and can be used as a comparison to test methods in validation studies[2]. Here, the average of three 24-hour dietary recalls collected by telephone approximately three months apart over one year's time were used to represent estimated usual intake. We choose the 24-hour dietary recall method for comparison because of its low cost, ease of administration by telephone, and relatively small respondent burden. Also, because an interviewer asks the questions and records each response, people with low literacy levels or problems with vision are not excluded. With these strengths in mind we expected that those agreeing to the validation study would be more representative of the total population than those who might agree to a more intense method of assessment, such as weighed food records or lengthy diet histories.

We recognise several weaknesses of the telephone 24-hour dietary recall method when used as a reference method for validation studies. First, a few days of recall may not reflect a person's true usual intake. Second, the accuracy of the interview relies heavily on the short-term memory of the respondent. Third, because it is also a self-report method of dietary assessment as is the FFQ, the within-person error may be correlated to the within-person error of the FFQ. Unidentified correlated error between methods may falsely elevate correlation coefficients and lead to the assumption of better agreement between the reference method and test method than is warranted[19].

By calculating intraclass correlation coefficients using the three 24-hour dietary recall interviews, we identified an unexpectedly high degree of variability in daily intake of some nutrients. The recall interviews were purposely spaced approximately three months apart to pick up seasonal variation in diet, so some measure of variability was anticipated; however, the degree of variability was greater than expected.

Some may argue that telephone interviews may not give acceptable results in this population because of common physical limitations such as hearing loss. However, after comparing telephone recalls of midday meals consumed by elderly people at congregate meal sites with data on actual intake for the meals, Dubois and Boivin[20] concluded that dietary recall collected by telephone is an acceptable way to obtain short-term dietary data from elderly subjects.

Mean intakes from the average of the three telephone 24-hour dietary recalls were similar to reported intakes of younger and older respondents, 66 to 100 years old, in a similar study comparing multiple 24-hour dietary recalls to picture-sort FFQs[10]. Mean intakes estimated from our 137-item picture-sort FFQ were higher than those reported by Kumanyika *et al.*[10], who used a similar 99-item picture-sort FFQ method[15]. Adjusting nutrient intake to give a measure of intake independent of total caloric intake may aid in the comparison of correlation coefficients from our study with those of other studies that compare results from different dietary assessment methods. Inaccuracies in estimated absolute nutrient intake from FFQs may come from incomplete listing of possible foods, or errors in frequency and portion size estimation.

We calculated the ratio of mean nutrient intake reported from the second FFQ to the average mean nutrient intake reported from the 24-hour dietary recall interviews, to evaluate the agreement of mean estimated energy and nutrient intakes between dietary assessment methods. Although ratios of mean intake were large for some nutrients in some age–gender groups, the percentage of calories from carbohydrate, protein and fat were similar between methods for all age–gender groups and to those reported for elderly people in the 1987 Nation Health Interview Survey[15]. Reports of ratios between our reference method and picture-sort FFQ were similar to results found by Larkin *et al.*[18], who used a 116-item FFQ that also included a sorting process. Nutrients with ratios furthest from unity in all age–gender groups were vitamin D and caffeine. This may be due to under-reporting in the 24-hour dietary recalls or over-reporting in the picture-sort FFQ of single food items such as milk and coffee. Milk provided the majority of vitamin D intake in all age–gender groups (34% in men ≤69 years old, 45% in men >69 years old, 54% in women aged ≤69 years, 54% in women aged >69 years). Similarly, coffee provided the majority of caffeine intake in all age–gender groups (73% in men ≤69 years old, 65% in men >69 years old, 84% in women aged ≤69 years, 80% in women aged >69 years). The ratio of alcohol intake was significantly less than unity for both younger and older women (0.75 and 0.33, respectively), indicating that women may under-report alcohol intake using the picture-sort FFQ.

Reproducibility of the picture-sort FFQ, as evaluated by energy-adjusted Spearman rank correlation coefficients and the ratio of energy-adjusted mean nutrient estimates from repeat administrations of the FFQ, was relatively high for all age–gender groups. Spearman rank energy-adjusted correlation coefficients for all nutrients across age–gender groups ranged from 0.26 to 0.96. With the exception of protein in younger women, these correlations are generally as strong as or stronger than correlations reported by other researchers assessing reproducibility of FFQs[6,17,21,22]. Correlations evaluating the repeatability of FFQs, reported by researchers such as Willett *et al.*[6], Munger *et al.*[21], Lazarus *et al.*[22], Mares-Perlman *et al.*[23] and Klipstein-Grobusch *et al.*[17], range from 0.41 to 0.99 for nutrients we assessed in our study.

Estimates of alcohol and caffeine intake had the strongest correlations between the first and second administrations of the picture-sort FFQ. This finding is consistent with results reported by Munger *et al.*[21], Klipstein-Grobusch *et al.*[17] and Mares-Perlman *et al.*[23], and may be due to low variability in the use of foods containing alcohol and caffeine.

We used Spearman rank correlation coefficients, and ratios of the mean of three 24-hour dietary recalls and the second picture-sort FFQ, to evaluate relative validity of the Utah Picture-sort FFQ. Energy adjustment improved correlations for most nutrients, with the exception of alcohol. In this population where 79% of women and 82% of men reported no alcohol intake from any dietary assessment interview, alcohol and total energy intake were poorly correlated ($r = 0.140$ in men, $r = 0.098$ in women) and energy adjustment of alcohol attenuates the correlation. Because of the lack of association between alcohol intake and total energy intake, it may be more appropriate to use unadjusted correlations when evaluating repeatability and validity of reported alcohol intake using the picture-sort FFQ.

After energy adjustment for all nutrients with the exception of alcohol, nutrient correlations ranged from 0.01 to 0.55 (median = 0.45) for macronutrients, fibre and cholesterol; 0.43 to 0.83 (median = 0.64) for alcohol, caffeine and water; and 0.15 to 0.70 (median = 0.50) for micronutrients across all age–gender groups. In general, with the exception of protein in older women, correlation coefficients from our study were within the range of correlation coefficients reported from other studies comparing typical paper-and-pencil FFQs with a reference dietary assessment method[6,17,21,24]. Correlations evaluating the relative validity of paper-and-pencil FFQs, reported by researchers such as Willett *et al.*[6], Munger *et al.*[21], Grootenhuis *et al.*[24], Mares-Perlman *et al.*[23] and Klipstein-Grobusch *et al.*[17], range from 0.14 to 0.83 for nutrients we assessed in our study. Correlation coefficients reported by Kumanyika *et al.*[10], who also compared a picture-sort FFQ with multiple dietary recalls in an elderly population, were in the range 0.22–0.61 for macronutrients, fibre and cholesterol; and 0.18–0.58 for micronutrients. Alcohol, caffeine and water were not studied. As compared with Kumanyika *et al.*'s study, we report somewhat lower correlations for macronutrients and higher correlations for micronutrients.

Although correlations from our study did include ranges with lower limits, these lower correlations were seen only in older men and women aged >69 years. This may suggest that respondents' ability to accurately report usual intake with our picture-sort FFQ and possibly with other dietary assessment methods, including 24-hour dietary recall, may decrease progressively with age.

Calculating correlation coefficients and other statistics by age–gender strata enabled us to assess differences by gender as well as differences between respondents aged ≤69 years and those aged >69 years within gender. Differences between age–gender groups were noted for statistics used to compare mean nutrient estimates and to assess reproducibility and validity. In general, it appears that younger respondents (≤69 years) may report usual nutrient intake by using the picture-sort FFQ slightly more accurately than older respondents (>69 years). Because of known physical and cognitive limitations that occur naturally with increasing age, this is not surprising. Limitations imposed on respondents as they age may hinder their ability to accurately report nutrient intakes using any dietary assessment method. Although we report lower correlations between methods and administrations for respondents aged >69 years as compared with respondents aged ≤69 years, we believe these same differences may be identified in other elderly populations by using other dietary assessment methods. This difference in ability to accurately report nutrient intake may not be a function of the picture-sort FFQ method but rather a function of problems and limitations associated with increasing age. Further study should be conducted on the repeatability and validity of dietary assessment methods for elderly populations by extending the upper age groups to include those in their eighties and nineties, and then comparing performance between sub-groups of elderly men and women based on age.

Conclusion

Our data suggest that structuring the FFQ as a two-step process of sorting food picture cards into period of use and then reporting frequency of use per period gives a useful measure of usual dietary intake. Correlation coefficients comparing the Utah Picture-sort FFQ to a measure of usual dietary intake were comparable to correlation coefficients reported in other validation studies comparing traditional FFQs as well as other picture-sort FFQs to reference methods. Although we did not conduct formal process evaluations, we believe picture cards may allow better identification of foods in respondents with poor reading skills or impaired vision, and that the sorting process itself may improve cognitive orientation to the task of assigning accurate frequency of use to specific foods. Research on using cognitive interviewing to improve FFQs suggests that the manipulation of cards may positively influence the response process by providing visual, tactile and motor involvement[25]. In the future, it would be helpful to conduct formal tests of process and sensory evaluation between a picture-sort FFQ and a traditional paper-and-pencil FFQ. Results from such process and sensory evaluations may make it clearer whether the picture-sort FFQ offers a significant advantage over traditional FFQs in populations where poor vision or limited literacy and language skills are common.

Acknowledgements

This study was carried out through support from the United States National Institutes of Health, National Institute of Arthritis and Musculoskeletal and Skin Diseases, grant # 5R01 AR43391-04.

References

1 Thompson FE, Byers T. Dietary assessment resource manual. *J. Nutr.* 1994; **124**: 2245S–317S.
2 Willett WC. *Nutritional Epidemiology*. New York: Oxford Press, 1990.
3 Van Staveren WA, de Groot LC, Blauw YH, van der Wielen RP. Assessing diets of elderly people: problems and approaches. *Am. J. Clin. Nutr.* 1994; **59**: 221S–3S.
4 Kumanyika SK, Tell GS, Shemanski L, Martel J, Chinchilli VM. Dietary assessment using a picture-sort approach. *Am. J. Clin. Nutr.* 1997; **65**: 1123S–9S.
5 Freedman LS, Carroll RJ, Wax Y. Estimating the relation between dietary intake obtained from a food frequency questionnaire and true average intake. *Am. J. Epidemiol.* 1991; **134**: 310–20.
6 Willett WC, Sampson L, Stampfer MJ. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol.* 1985; **122**: 51–65.
7 Willett WC, Reynolds RD, Cottrell-Hoehner S, Sampson L, Browne ML. Validation of a semi-quantitative food frequency questionnaire: comparison with a 1-year diet record. *J. Am. Diet. Assoc.* 1987; **87**: 43–7.
8 Willett WC, Sampson L, Browne ML, *et al.* The use of a self-administered questionnaire to assess diet four years in the past. *Am. J. Epidemiol.* 1988; **127**: 188–99.
9 Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L. A data-based approach to diet questionnaire design and testing. *Am. J. Epidemiol.* 1986; **124**: 453–69.
10 Kumanyika S, Tell GS, Fried L, Martel JK, Chinchilli VM. Picture-sort method for administering a food frequency questionnaire to older adults. *J. Am. Diet. Assoc.* 1996; **96**: 137–44.
11 Posner BM, Smigelski C, Duggal A, Morgan JL, Cobb J, Cupples LA. Validation of two-dimensional models for estimation of portion size in nutrition research. *J. Am. Diet. Assoc.* 1992; **92**: 738–41.
12 Johnson RK, Driscoll P, Goran MI. Comparison of multiple-pass 24-hour recall estimates of energy intake with total energy expenditure determined by the doubly labeled water method in young children. *J. Am. Diet. Assoc.* 1996; **96**: 1140–4.
13 Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *Am. J. Epidemiol.* 1986; **124**: 17–27.
14 Beaton GH, Milner J, Corey P, *et al.* Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am. J. Clin. Nutr.* 1979; **32**: 2546–9.
15 Block GSE. Estimate of nutrient intake from a food frequency questionnaire: The 1987 National Health Interview Survey. *J. Am. Diet. Assoc.* 1992; **92**: 969–77.
16 Mares-Perlman JA, Klein BE, Klein R, Ritter LL, Freudenheim JL, Luby MH. Nutrient supplements contribute to the dietary intake of middle- and older-aged adult residents of Beaver Dam, Wisconsin. *J. Nutr.* 1993; **123**: 176–88.
17 Klipstein-Grobusch K, den Breeijen JH, Goldbohm RA, *et al.* Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire. *Eur. J. Clin. Nutr.* 1998; **52**: 588–96.
18 Larkin FA, Metzner HL, Thompson FE, Flegal KM, Guire KE. Comparison of estimated nutrient intakes by food frequency and dietary records in adults. *J. Am. Diet. Assoc.* 1989; **89**: 215–23.
19 Kipnis V, Carroll RJ, Freedman LS, Li L. Implications of a new dietary measurement error model for estimation of relative risk: application to four calibration studies. *Am. J. Epidemiol.* 1999; **150**: 642–51.
20 Dubois S, Boivin JF. Accuracy of telephone dietary recalls in elderly subjects. *J. Am. Diet. Assoc.* 1990; **90**: 1680–7.
21 Munger RG, Folsom AR, Kushi LH, Kaye SA, Sellers TA. Dietary assessment of older Iowa women with a food frequency questionnaire: nutrient intake, reproducibility, and comparison with 24-hour dietary recall interviews. *Am. J. Epidemiol.* 1992; **136**: 192–200.
22 Lazarus R, Wilson A, Gliksman M, Aiken J. Repeatability of nutrient intakes estimated by a semiquantitative food frequency questionnaire in elderly subjects. *Ann. Epidemiol.* 1995; **5**: 65–8.
23 Mares-Perlman JA, Klein BE, Klein R, Ritter LL, Fisher MR, Freudenheim JL. A diet history questionnaire ranks nutrient intakes in middle-aged and older men and

women similarly to multiple food records. *J. Nutr.* 1993; **123**: 489–501.
24 Grootenhuis PA, Westenbrink S, Sie CM, de Neeling JN, Kok FJ, Bouter LM. A semiquantitative food frequency questionnaire for use in epidemiologic research among the elderly: validation by comparison with dietary history. *J. Clin. Epidemiol.* 1995; **48**: 859–68.
25 Lieberman LS, Bobroff LB, Probart CK, Kurtz CS. A visual tool for improved cross-cultural dietary assessment. *Am. J. Clin. Nutr.* 1994; **59S**: 290S (abstract).