

Reliability and validity of a food-frequency questionnaire for Chinese postmenopausal women

Liangzhi Xu^{1,*}, Michael J Dibley² and Catherine D'Este²

¹Department of Obstetrics and Gynaecology, The Second West China Hospital, West China Medical Centre, Sichuan University, Chengdu 610041, People's Republic of China; ²Centre for Clinical Epidemiology and Biostatistics (CCEB), School of Medical Practice and Population Health, Faculty of Health, The University of Newcastle, Newcastle, New South Wales, Australia

Submitted 6 January 2003; Accepted 25 June 2003

Abstract

Objectives: (1) To determine the reliability and validity of a food-frequency questionnaire (FFQ) for use in epidemiological research in postmenopausal women; and (2) to compare the volume estimation (VE) and weight estimation (WE) method of administration of this questionnaire.

Design: An initial list of foods was derived and modified after pre-testing in 22 subjects. Test–retest reliability was assessed in 21 subjects who had repeat administrations of the questionnaire 14 days apart (FFQ1, FFQ2). The validity of the FFQ was assessed by comparing nutrient intakes with those from a 4-day food record.

Setting: Chengdu, People's Republic of China.

Subjects: Twenty-two postmenopausal women (50–70 years) were recruited from The Second University Hospital, West China University of Medical Sciences, Chengdu and participated in the pre-test. Another 21 women (50–70 years) were randomly selected from the general population of all five districts of Chengdu and participated in the reliability and validity sub-studies.

Results: Energy, protein, carbohydrate, magnesium and sodium intakes in this sample were less than the Recommended Dietary Allowances (RDAs) for 45–70-year-old women in China. Intake of non-cooking fat was higher than the Chinese RDA. Pearson correlation coefficients and intra-class correlation coefficients (ICCs) for reliability of the VE FFQ ranged from 0.51 to 0.85 and from 0.51 to 0.81, respectively; for the WE FFQ, they ranged from 0.22 to 0.86 and from 0.21 to 0.81. Correlation coefficients and ICCs for validity of the WE FFQ ranged from 0.36 to 0.69 and from 0.34 to 0.57, respectively; corresponding values for the VE FFQ were –0.30 to 0.65 and –0.14 to 0.65.

Conclusions: Both the VE and WE FFQs were reliable and valid except for sodium intake. The VE FFQ provided more valid estimates of nutrient intakes than did the WE FFQ.

Keywords
Food-frequency questionnaire
Nutrient intake
Reliability
Validity
Postmenopausal women
China

Accumulated evidence suggests that nutrient intakes are associated with the development of many diseases, especially chronic conditions. During the past 20 years, food-frequency questionnaires (FFQs) have become commonly used instruments for measuring relationships between nutrient intakes and diseases in epidemiological studies in Western countries¹. However, few FFQs have been validated for Chinese people, especially those residing in mainland China. The problem arises as to whether FFQs developed and validated in other populations or in Chinese populations outside mainland China are reliable and valid for mainland Chinese populations. For example, the food list of an FFQ developed for Australians of Chinese origin consisted of 'steak (grilled),

steak (pan-fried), roast beef/veal, crumbed veal/schnitzel (deep-fried), pork chop (grilled), pork chop (pan-fried), roast pork', which are foods rarely eaten in mainland China. Furthermore, the reference portions were expressed as big, medium, small and slice², which are food portion concepts unfamiliar to people in mainland China, particularly for meat intake. In addition, similar instruments may perform differently in different populations. Given the marked differences in social, environmental and nutritional conditions between developed and developing countries, and the obvious differences in food consumption practices, there is a need to develop new questionnaires that are more applicable to mainland China to assess nutrient intakes and their relationship with diseases.

*Corresponding author. Email liangzxi@mail.sc.cninfo.net

In China, people are familiar with estimating weights of foods based on the Liang method (1 Liang = 50 g). However, the validity of this method has not been previously reported except for calcium intake³. The present study expands on earlier work reporting the validation of a food frequency instrument to measure calcium intake in Chinese postmenopausal women³, to cover other nutrients. This study was conducted with the following aims: (1) to analyse the nutrient intakes derived from a newly developed quantitative FFQ in Chinese postmenopausal women; and (2) to compare the validity and reliability of the traditional Chinese FFQ method based on weight estimation in Liangs with those of a new FFQ method based on volume estimation.

Methods

Development of the FFQ

The development of the FFQ has previously been described in detail³. Briefly, the foods listed in the Chinese *Food Composition Table*⁴, which contain relatively high levels of calcium and/or foods containing relatively less calcium but commonly eaten in Chengdu, People's Republic of China, formed the basis for the first draft of the questionnaire. This initial draft was piloted with 22 postmenopausal women recruited from relatives of women admitted to the Second University Hospital of West China University of Medical Sciences in Chengdu. The list of foods was then revised by eliminating some items found not to be consumed or to be consumed infrequently in Chengdu. In addition, some commonly eaten foods that were not included in the original questionnaire were added. Thus the revised food list for the FFQ included 110 foods. Cooking oil, salt and soy sauce were not included because of the lack of an appropriate methodology to measure them. The food items were organised in culturally appropriate groups (e.g. rice, meat, vegetables and fruits). At the end of the questionnaire, some space was left to record other foods that were consumed but not included on the list.

The protein, carbohydrate, fat, phosphorus, magnesium and sodium content of most foods used in the checklist were derived from the Chinese *Food Composition Table*⁴ and package labels of commercial items. Nine foods (wonton, meat and vegetable dumpling, glutinous rice dumpling, stuffed steam bun, walnut and peanut milk, boiled garden pea, white sesame candy, black sesame candy and sea cabbage) could not be located in the above sources, or in other published data sources. Samples of these foods were collected from different markets in the Chengdu area and sent to the Department of Health Testing in the School of Public Health, Sichuan University for assessment of protein, carbohydrate, fat, phosphorus, magnesium and sodium content^{5,6}.

Assessing the quantity of food consumed

The 'weight estimation (WE)' method for assessing the amount of a food consumed was determined by asking respondents to estimate how many Liangs of each food they consumed in the previous week.

The quantity of the food consumed, as determined by an assessment of the volume and portion size, was referred to as the 'volume estimation (VE)' method. For foods eaten at home, in restaurants or in cafeterias, subjects were presented with four bowls of different volume and asked to choose the appropriately sized bowl for each food, and to indicate the volume eaten by filling the bowl with an equivalent amount of rice. For fruits consumed, subjects were asked to choose the corresponding pictures representing different serving sizes of each fruit, and their corresponding estimated weight in grams was recorded. For commercial items such as sliced bread, cakes, packaged biscuits, pies and dumplings, the unit weight of the item (in grams) and the number of units consumed were recorded.

Nutrient intakes from the FFQ

The daily nutrient intakes derived from both methods were calculated as the total number of grams of each food consumed in that week multiplied by the amount of nutrient in each gram of that food, divided by seven. In the WE method, the number of Liangs was translated to grams by multiplying by 50.

Assessing the test–retest reliability and relative validity of the FFQ

The community structure of Chengdu is such that the whole city is divided into five districts. Each district is divided into several sub-districts and each sub-district has its own administration office called a 'street office'. Each sub-district is divided into several sections. Each section has an administration unit called the 'resident committee'. These are the smallest administrative units in the city and are responsible for the administration of residents who live in a small area from one to several blocks.

One street office was selected randomly from each of the five districts of Chengdu through use of a random number table. Following this, a resident committee was randomly selected from each of the five street offices. Four to five respondents were randomly selected by a random number table from female residents aged 50–70 years living in the area defined by the selected resident committee. The inclusion criteria for selecting women included postmenopausal, no history of cancer and willing to participate.

One of the authors (L.X.) interviewed all of the subjects in their home using the revised FFQ. For all foods consumed one week prior to interview, respondents were asked to estimate both the weight in Liangs and to show the volume consumed each time the food was eaten. The VE FFQ was administered after the WE FFQ at the same

visit. Two weeks after the first interview, the interviewer repeated the entire process using the same methods.

The validity of the two FFQ methods was assessed by comparing the nutrient intakes derived from the FFQs with those derived from a 4-day food record (FR). A 4-day FR was collected from the same subjects during the week after administration of the first interviewer-administered FFQ. The interviewer went to the respondent's home after each meal (breakfast, lunch and dinner) over the 4 days (i.e. a total of 12 times). The weight of each food consumed by the respondent from the last meal, recorded to the most recent meal, was determined by recording the weight of all purchased foodstuffs (as reported by the respondent or directly weighed by the interviewer), weighing the amounts of foodstuffs that remained after the meal and that were discarded during food preparation, and estimating the proportion of the consumed food eaten by the respondent (if the food was eaten by multiple family members).

Statistical analysis

Nutrient intake values were log_e-transformed to improve the normality of the distribution before further analysis. However, the distributions remained skewed for most variables so non-parametric analyses were conducted.

Pearson product-moment correlation coefficients were calculated for both reliability and validity to measure the ability to rank individuals by nutrient intake as measured by the FFQs. The intra-class correlation coefficient (ICC) is used to quantify agreement and was calculated using results from a two-way analysis of variance⁷. Subjects were also classified into tertiles of consumption of each nutrient, for both the FFQ and the 4-day FR, in order to examine their cross-classification. The overall percentages of individuals correctly classified into the same, adjacent or extreme tertiles were determined.

Results

Daily nutrient intakes

The median daily intakes of protein, carbohydrate, fat, phosphorus, magnesium and sodium derived from VE FFQ, WE FFQ and the 4-day FR are presented in Table 1. All of the absolute median values estimated by VE and WE FFQs were within ±20% and ±25%, respectively, of values from the 4-day FR, except for sodium intake.

Test-retest reliability of the FFQ

The correlation coefficients and/or the ICCs of nutrient intake derived from VE and WE FFQs collected 2 weeks apart are shown in Table 2. Nearly all of the ICCs (0.21–0.81) were lower than the correlation coefficients (0.22–0.86). Except for sodium intake, the correlation coefficients were all above 0.6. There was good correlation and agreement for all nutrients (*r* ≥ 0.78, ICC ≥ 0.76, Table 2) between the WE FFQ and the VE

Table 1 Range of nutrient intakes of 21 subjects in the validation study based on the weight estimation (WE) and volume estimation (VE) food-frequency questionnaires (FFQ1 and FFQ2) and the 4-day food record (FR), the Recommended Dietary Allowance for 45–70-year-old women in China (RDA) and % of the RDA for the 4-day FR

Nutrient	VE FFQ1		VE FFQ2		WE FFQ1		WE FFQ2		4-day FR		RDA	% of RDA†
	Median (Q1–Q3)*	Median (Q1–Q3)	Median (Q1–Q3)	Median (Q1–Q3)	Median (Q1–Q3)	Median (Q1–Q3)	Median (Q1–Q3)	Median (Q1–Q3)	Median (Q1–Q3)	Median (Q1–Q3)		
Energy (MJ day ⁻¹)	5.90 (4.48–7.31) ^{bg}	5.38 (4.39–6.82) ^h	5.93 (4.99–7.48) ^e	5.69 (4.78–6.78) ^f	5.54 (4.10–5.97)	5.54 (4.10–5.97)	7.1–8.0	72–64				
Protein (g day ⁻¹)	46.5 (38.3–61.6) ^g	49.6 (40.4–71.7) ^h	53.5 (43.5–64.0) ^e	52.2 (43.7–64.4)	46.6 (35.6–63.3)	46.6 (35.6–63.3)	60–65	78–72				
Carbohydrate (g day ⁻¹)	209.9 (159.7–259.4) ^{ab}	194.3 (166.2–215.5)	215.5 (180.6–266.9) ^{de}	194.4 (165.9–229.3)	173.5 (152.0–201.0)	173.5 (152.0–201.0)	280–291*	62–60*				
Fat (g day ⁻¹)	39.3 (19.3–45.4) ^g	37.0 (24.0–53.4) ^h	41.9 (37.3–58.6) ^e	39.5 (29.9–59.0) ^f	37.3 (15.2–52.5)	37.3 (15.2–52.5)	20–25% of energy	138–110				
Phosphorus (mg day ⁻¹)	816.6 (660.9–1044.4) ^g	770.5 (720.3–1043.9) ^h	969.3 (736.7–1129.3) ^e	869.5 (760.8–1068.6) ^f	802.4 (582.1–931.4)	802.4 (582.1–931.4)	300	83				
Magnesium (mg day ⁻¹)	269.1 (20.98–348.3) ^{abg}	252.6 (223.1–283.7) ^h	308.6 (247.6–368.4) ^{de}	287.0 (233.6–313.5)	249.1 (190.6–281.7)	249.1 (190.6–281.7)	1100–3300	109–36				
Sodium (mg day ⁻¹)	475.9 (253.0–799.5) ^{bg}	511.5 (296.0–695.0) ^{ch}	966.1 (381.3–2767.6)	705.7 (352.7–1406.8)	1195.1 (763.5–3426.1)	1195.1 (763.5–3426.1)						

* Q1–Q3 – 25%–75% range. † RDA from carbohydrate is the amount of carbohydrate required to produce the remaining energy after deducting the energy for the fat RDA and the protein RDA from the total energy RDA. Significant difference, *P* < 0.05: ^aVE FFQ1 vs. VE FFQ2; ^bVE FFQ1 vs. 4-day FR; ^cVE FFQ2 vs. 4-day FR; ^dWE FFQ1 vs. WE FFQ2; ^eWE FFQ1 vs. 4-day FR; ^fWE FFQ2 vs. WE FFQ1; ^gWE FFQ1 vs. WE FFQ2; ^hWE FFQ2 vs. WE FFQ1.

Table 2 Reliability of the volume estimation (VE) and weight estimation (WE) food-frequency questionnaires (FFQ1 and FFQ2) in 21 postmenopausal women in China

Nutrient	VE FFQ1 vs. VE FFQ2		WE FFQ1 vs. WE FFQ2		VE FFQ1 vs. WE FFQ1		VE FFQ2 vs. WE FFQ2	
	<i>r</i>	ICC	<i>r</i>	ICC	<i>r</i>	ICC	<i>r</i>	ICC
	Energy	0.85	0.81	0.82	0.76	0.86	0.82	0.93
Protein	0.79	0.78	0.82	0.78	0.94	0.89	0.96	0.95
Carbohydrate	0.74	0.69	0.70	0.64	0.78	0.76	0.82	0.80
Fat	0.72	0.71	0.64	0.63	0.94	0.92	0.98	0.97
Phosphorus	0.78	0.75	0.78	0.75	0.94	0.90	0.78	0.92
Magnesium	0.84	0.80	0.86	0.81	0.89	0.86	0.91	0.88
Sodium	0.51	0.51	0.22	0.21	0.66	0.42	0.22	0.63

r – Pearson product–moment correlation coefficient; ICC – intra-class correlation coefficient.

FFQ, except for sodium. The exact tertile agreement between the first and the second administration ranged from 38% to 90% for nutrient intake from the WE method and from 57% to 81% for nutrient intake from the VE method (Table 3).

Validity of the FFQ

Except for sodium intake, the ICCs of VE FFQ1 (ICC = 0.56–0.65) were higher than those of WE FFQ1 (ICC=0.37–0.57) (Table 4) when compared with the 4-day FR. Most of the ICCs (except for sodium) for FFQ2 (both WE and VE) versus the 4-day FR were a little higher than those of FFQ1 (both WE and VE) versus the 4-day FR (Table 4). The exact tertile agreement between the FFQs and the 4-day FR is shown in Table 5.

Discussion

Daily nutrient intakes

The median intakes of all nutrients measured by VE FFQ1 and the 4-day FR are lower than those estimated from 35–75-year-old women in Cangxi County of the same province, Sichuan, using a 3-day food weighing method, except for fat⁸. Energy, protein, carbohydrate and fat

intakes in the present study subjects are relatively lower than the nutrient intakes estimated by a self-administered FFQ in women 55 years old or more in Hong Kong⁹. The daily energy, protein, carbohydrate and magnesium intakes in menopausal women in Chengdu, People's Republic of China are less than the current Chinese Recommended Dietary Allowances (RDAs)¹⁰. One reason for these apparently lower intakes of nutrients is that it was not possible to measure intakes of cooking oil and salt precisely due to Chinese cooking methods. Furthermore, most Chinese dishes have some soup after cooking in which some oil and salt will remain uneaten. So, the amount of oil or sodium used for cooking will not be equal to the amount consumed. The proportion of total energy intake accounted for by fat intake in the study sample was higher than the Chinese RDA even though cooking oil was not included in the estimates of fat intake. Dietary modification might be necessary if this is confirmed in a sample of larger size.

Test–retest reliability of the FFQ

In the present study, the correlation coefficients and ICCs between two administrations 2 weeks apart ranged from 0.51 to 0.85 for the VE FFQ and from 0.21 to 0.86 for the WE FFQ (Table 2). It appears that the test–retest reliability of the volume estimation questionnaire is comparable to the reproducibility of other FFQs already reported in the literature. (The correlation coefficients and ICCs for energy, fat, protein, carbohydrate and magnesium observed in previous studies on the reliability of FFQs for use in adult women varied from 0.37 to 0.74 and from 0.24 to 0.74, respectively^{11–21}.) This indicates that the FFQ, developed for postmenopausal women in Chengdu, is reliable.

The percentage of agreement within tertiles between the two administrations of the FFQ ranged from 57% to 81% for the VE method and from 38% to 90% for WE. However, these results are not comparable to other studies

Table 3 Percentage of agreement within tertiles between the two administrations of the food-frequency questionnaire (FFQ1 and FFQ2) and volume estimation (VE) vs. weight estimation (WE) FFQs in 21 Chinese participants

Nutrient	VE FFQ1 vs. VE FFQ2			WE FFQ1 vs. WE FFQ2			VE FFQ1 vs. WE FFQ1			VE FFQ2 vs. WE FFQ2		
	Same tertile*	Adjacent tertile†	Opposite tertile‡	Same tertile	Adjacent tertile	Opposite tertile	Same tertile	Adjacent tertile	Opposite tertile	Same tertile	Adjacent tertile	Opposite tertile
Energy	57	38	5	67	29	4	71	29	0	90	10	0
Protein	71	19	10	67	29	4	71	29	0	90	10	0
Carbohydrate	81	19	0	71	29	0	76	19	5	57	38	5
Fat	62	38	0	57	38	5	90	10	0	90	10	0
Phosphorus	62	38	0	71	29	0	71	29	0	90	10	0
Magnesium	81	19	0	90	10	0	90	10	0	67	29	4
Sodium	62	29	10	38	48	14	48	48	4	81	19	0

* The two methods categorised nutrient intake into the same tertile.

† The two methods categorised nutrient intake into adjacent tertiles (one method categorised nutrient intake into the first tertile while the other method categorised it into the second tertile; or one method categorised nutrient intake into the second tertile while the other method categorised it into the third tertile).

‡ The two methods categorised nutrient intake into opposite tertiles (one method categorised nutrient intake into the first tertile while the other method categorised it into the third tertile).

Table 4 Validity of the volume estimation (VE) and weight estimation (WE) food-frequency questionnaires (FFQ1 and FFQ2) compared with the 4-day food record (FR) in 21 Chinese participants

Nutrient	WE FFQ1 vs. 4-day FR		WE FFQ2 vs. 4-day FR		VE FFQ1 vs. 4-day FR		VE FFQ2 vs. 4-day FR	
	<i>r</i>	ICC	<i>r</i>	ICC	<i>r</i>	ICC	<i>r</i>	ICC
	Energy	0.59	0.45	0.58	0.54	0.65	0.65	0.72
Protein	0.69	0.57	0.66	0.63	0.63	0.63	0.66	0.64
Carbohydrate	0.48	0.37	0.50	0.44	0.56	0.56	0.60	0.58
Fat	0.49	0.43	0.60	0.58	0.61	0.61	0.67	0.64
Phosphorus	0.69	0.57	0.66	0.62	0.64	0.64	0.69	0.65
Magnesium	0.56	0.44	0.56	0.51	0.58	0.59	0.60	0.56
Sodium	0.36	0.34	0.35	0.18	-0.30	-0.14	-0.07	-0.06

r – Pearson product-moment correlation coefficient; ICC – intra-class correlation coefficient.

that used quartiles or quintiles of nutrient intake because of the small sample size of the present study.

Validity of the FFQ

In contrast to some studies^{21,22}, the VE FFQ of the present study did not show an overestimation of most nutrients except for carbohydrate, while the WE FFQ overestimated most of the nutrients except for carbohydrate and sodium. One of the reasons for this might be that the VE FFQ used volume to estimate the amount of food consumed, thus limiting errors of perception of portion sizes. In contrast, the WE FFQ based on the respondents' direct estimation of the amount eaten was more likely to be subject to errors of perception by the respondents. However, the absolute values for nutrient intakes estimated by the FFQ in the present study were in the range of ± 20 –25% of the food record estimates. This level of error is not of importance in most epidemiological studies, where analyses are based on the ranking of individuals by level of nutrient intake, provided that the ranking has been shown to be valid²³.

Except for sodium, the correlation coefficients and ICCs for validity of the WE FFQ1 were in the range 0.48–0.69

and 0.37–0.57, respectively (Table 4). These were less than those of the VE FFQ1 (0.56–0.65, Table 4) and indicated that the weight estimation method of assessing the quantity of food consumed, which is the one usually used in China²⁴, is a little less accurate and had less agreement than the volume estimation method.

The correlations for energy, protein, fat, carbohydrate, magnesium and phosphorus observed in this study were generally of a higher level than those observed in other validation studies of adult women (correlation coefficients ranged from 0.13 to 0.73; ICCs ranged from 0.19 to 0.29)^{12,14,16,17,20–22,25–38}. There are several possible reasons for the relatively high level of validity for the volume estimation method. First, the FFQ measured the actual quantities of most home-made foods consumed by assessing the volume (ml) consumed. This is relatively more objective than estimations by respondents' perception alone.

Second, the relatively short duration of the recall period together with the interviewer-administered method would reduce memory-related problems of recall. This is because questionnaires validated against a recent period such as the past month or past week would tend to produce higher correlations than questionnaires that ask about a longer period such as the past year²⁶. The short duration of the present study and the stability of the Chinese diet also meant that there was no need to adjust for seasonal variation, so it would be logical that correlations of the present study were higher.

Third, the study was conducted in the subjects' home, so it was easy and convenient for participants to show the bowl generally used for eating. In addition, subjects in their own home would have cues to improve and facilitate the accuracy of their recall³⁹.

Fourth, it was more convenient and time-efficient to use an open-ended rather than a closed-ended frequency format because the interviewer did not need to spend time covering all the possible frequency response options. This also might provide some enhanced precision in reporting,

Table 5 Percentage of agreement within tertiles between the volume estimation (VE) and weight estimation (WE) food-frequency questionnaires (FFQ1 and FFQ2) and the 4-day food record (FR) in 21 Chinese participants

Nutrient	VE FFQ1 vs. 4-day FR			VE FFQ2 vs. 4-day FR			WE FFQ1 vs. 4-day FR			WE FFQ2 vs. 4-day FR		
	Same tertile*	Adjacent tertile†	Opposite tertile‡	Same tertile	Adjacent tertile	Opposite tertile	Same tertile	Adjacent tertile	Opposite tertile	Same tertile	Adjacent tertile	Opposite tertile
Energy	43	48	9	67	29	4	33	57	10	62	38	0
Protein	57	38	5	38	57	5	67	29	4	48	48	4
Carbohydrate	57	29	14	57	38	5	62	29	9	62	29	9
Fat	62	29	19	62	38	0	62	29	9	52	48	0
Phosphorus	52	38	10	48	48	4	62	29	9	48	48	4
Magnesium	48	48	4	43	48	9	48	48	4	48	48	4
Sodium	52	29	19	33	38	29	48	48	4	33	48	19

* The two methods categorised nutrient intake into the same tertile.

† The two methods categorised nutrient intake into adjacent tertiles (one method categorised nutrient intake into the first tertile while the other method categorised it into the second tertile; or one method categorised nutrient intake into the second tertile while the other method categorised it into the third tertile).

‡ The two methods categorised nutrient intake into opposite tertiles (one method categorised nutrient intake into the first tertile while the other method categorised it into the third tertile).

as the frequency of use is truly a continuous rather than a categorical variable. For some staple foods, such as rice, participants were asked: 'Did you eat rice every day during last week?' This was simpler and easier to recall precisely than asking 'How many times did you eat rice in the last week?' because people in Chengdu eat rice nearly every day. Furthermore, this might also prevent interviewer recording errors or biases of interviewers for given frequencies because there might be responder number preferences, which would influence the frequencies recorded.

Finally, the foods were listed in the FFQ in such a way that they followed the eating habits and meal patterns of Chinese people. That is, first with staple foods (such as cooked rice, steamed buns and dumpling) and then dishes, fruit and sweets. This culturally specific order of the food list in this FFQ might have facilitated recall of the usual dietary consumption patterns⁴⁰ and increased the quality of responses.

Advantages of interviewer-administered FFQs over self-administered FFQs

There are a number of potential advantages for using an interviewer-administered FFQ in China. For example, the subjects should have no problem understanding the questions being asked and instructions given, as these can be clarified directly with the interviewer. The experienced interviewer can prompt subjects concerning food items not listed in the questionnaire and so reduce a source of systematic between-person error that might occur when some individuals, but not others, recall their intake of particular foods⁴¹. In addition, the experienced interviewer is able to obtain more precise and accurate estimates of the different portion sizes of each food consumed, thus minimising the error involved in having subjects try to estimate the amount of food consumed and between-person error in portion size estimation among subjects.

One of the most important problems limiting the use of self-completion methods in China is the educational level of the women likely to be involved. Studies have shown that the literacy rate among women aged 50–70 years is relatively low⁴². It would be difficult for them to complete an FFQ. With the interviewer-administered FFQ the problems of illiterate women being unable to complete the FFQ can be avoided. Another problem for older women completing self-administered questionnaires is that their eyesight is relatively poor compared with younger people, and interviewer administration of the questionnaire can overcome this.

Problems associated with the reference period and multiple dietary assessments

A reference period of 1 week was used in this study because dietary patterns among mainland Chinese are relatively stable during a specific period in specific areas²⁴. The correlation coefficient between a 3-day household

dietary survey over a single season and a simple individual questionnaire that covered all seasons of the same year ranged from 0.50 for meat to 0.93 for rice²⁴. Woo *et al.* also found that their FFQ with a 1-week reference period could provide good estimates of energy, protein and potassium in dietary surveys in the Hong Kong Chinese population⁴³. Furthermore, older people tend to have more stable eating patterns than do younger adults⁴⁴. However, validity studies with a longer reference period such as 1 year are necessary to confirm the representativeness of long-term intake with the short referent period.

In this study subjects completed two separate quantitative FFQs, 2 weeks apart, with a 4-day FR in between over a short time (2 weeks). This administration might be burdensome and imply the dependency of measurements. The way that the WE FFQ was systematically administered before the VE FFQ made dietary intakes assessed by the two versions of the FFQ likely to be dependent even though the questions used by these two methods were designed to try to avoid this. In addition, the order of administration might influence the estimates in a systematic way and increase the correlation coefficients and ICCs between the two versions of the FFQ. The sequential administration of several dietary assessments to the same individual raises the problem that completion of one method early in the sequence may influence responses to a method administered later. It is possible that the process of measuring and recording food intake could itself influence dietary habits^{12,45} and responses to the second FFQ⁴⁶. In general, the completion of diet records could have sensitised participants with respect to their food consumption so that they became more skilled at answering the FFQ and estimating portion sizes afterwards. Thus the second FFQ may be more accurate than the first one, resulting in a better validity⁴⁶. The results of the present study add a further dimension to this. The validity of both WE FFQ1 and VE FFQ1 was lower than that of FFQ2, and this was more severe with weight estimation than with volume estimation. This suggests that the relative objectivity of the VE FFQ might partially limit the dependency effects because it estimated the amount of dietary intake according to the actual volume consumed and not the subjective perception of weight.

These points need to be considered when comparing the results from other studies. However, this is unlikely to be a serious problem in practice since only one FFQ is used in one nutrient–disease study and it is generally used only once. In addition, the median values of the FFQ1 were not significantly different to those from the FFQ2 for both methods except for carbohydrate and magnesium.

Strengths and limitations of the study

The protein, fat, carbohydrate, phosphorus, magnesium and sodium contents of nine foods that were not available from published sources were measured. This information could be used as a supplement to the Chinese

Food Composition Table. However, resource limitations restricted the number of nutrients analysed for other foods that were not in the Chinese *Food Composition Table*. Thus, the FFQ remains limited in its use for investigation of other disease–nutrient relationships. This problem needs to be addressed, because this population faces bone fractures as well as cardiovascular disease, and other diseases that might be related to nutrient intakes.

This study partially reduced the over- or underestimation of nutrient intakes (compared with food records) due to memory problems in an elderly population by having an experienced interviewer prompting participants in a standardised manner, adding foods not initially present in the original questionnaire after piloting, and using a relatively short reference period for the FFQ. The present study also overcame the difficult task (in particular, for some of the more elderly women) of assessing accurately the typical size of the portions of each food consumed by using an objective and relatively accurate volume measurement together with an interviewer-assisted method to encourage complete recall.

One of the most important potential limitations of the study is the fact that the results are based on a small sample ($n = 21$) due to limited resources and time and the intensity of the investigation (two researcher-assisted administrations of the FFQ and a researcher-conducted 4-day FR for each subject). Usually, a low sample size limits representativeness and may result in widely varying estimates of within-person nutrient intakes, which result in less reliable correlation coefficients and ICCs. However, the characteristics of the sample showed they were representative of the broader group of postmenopausal women in China³ and the small sample size was sufficient to assess nutrient intake differences from the two methods of administering the FFQ (Table 1). It may be necessary to re-evaluate the performance of the questionnaire in a larger sample that is more representative of populations in mainland China.

Due to special cooking practices in China, cooking oil and salt intake data were not collected because of the difficulties in estimating the relatively small amounts used. However, this will need to be corrected for studies of diseases that might have some relationship with fat and salt, such as hypertension or heart disease.

Conclusion

In summary, the reliability and validity of the volume estimation method for crude energy, protein, fat, carbohydrate, phosphorus and magnesium intakes are relatively higher than for the weight estimation method. Both methods could be used in epidemiological studies in Chinese postmenopausal women to estimate the nutrient intakes listed above. There is a need to assess other nutrients and re-assess the reliability and validity in other

populations, and to evaluate other nutrient-specific and general FFQs for use in China.

Acknowledgements

This work was supported by a research development grant by the International Clinical Epidemiology Network (INCLIN). The authors would like to thank Drs Guangjian Wang and Maoyu Zhang for their assistance in nutrient content measurement, Dr Xiaoping Pan for the biostatistical analyses, and Jane McDonald, Jenny Byrne and Sonia Freeman for their help in word processing during manuscript preparation.

References

- 1 Willett W. Food-frequency methods. In: Willett W, ed. *Nutritional Epidemiology*, 2nd ed. New York: Oxford University Press, 1998; 74–100.
- 2 Hsu-Hage BH, Wahlqvist ML. A food frequency questionnaire for use in Chinese populations and its validation. *Asia Pacific Journal of Clinical Nutrition* 1992; **1**: 211–23.
- 3 Xu L, Porteous JE, Phillips MR, Zheng S. Development and validation of a calcium intake questionnaire for the Chinese postmenopausal population. *Annals of Epidemiology* 2000; **10**(3): 169–75.
- 4 Chinese Academy of Preventive Medicine, Institute of Nutrition and Food Hygiene. *Food Composition Table*. Beijing: People's Health Press, 1992 [in Chinese].
- 5 Bureau of Standardisation of the People's Republic of China. *National Standardisation of Food Hygiene (2)*, 2nd ed. Beijing: Chinese Standardisation Press, 1995.
- 6 Bureau of Standardisation of the People's Republic of China. *National Standardisation of Food Hygiene (4)*, 1st ed. Beijing: Chinese Standardisation Press, 1997.
- 7 Morton AP, Dobson AJ. Assessing agreement. *Medical Journal of Australia* 1989; **150**(7): 384–7.
- 8 Hu JF, Zhao XH, Jia JB, Parpia B, Campbell TC. Dietary calcium and bone density among middle-aged and elderly women in China. *American Journal of Clinical Nutrition* 1993; **58**(2): 219–27.
- 9 Woo J, Leung SS, Ho SC, Lam TH, Janus ED. Dietary intake and practices in the Hong Kong Chinese population. *Journal of Epidemiology and Community Health* 1998; **52**(10): 631–7.
- 10 Chinese Nutrition Society. Chinese recommended dietary allowance. *Acta Nutrimenta Sinica* 1989; **11**(1): 93–6.
- 11 Rohan TE, Potter JD. Retrospective assessment of dietary intake. *American Journal of Epidemiology* 1984; **120**(6): 876–87.
- 12 Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *American Journal of Epidemiology* 1985; **122**(1): 51–65.
- 13 Byers T, Marshall J, Anthony E, Fiedler R, Zielezny M. The reliability of dietary history from the distant past. *American Journal of Epidemiology* 1987; **125**(6): 999–1011.
- 14 Martin-Moreno JM, Boyle P, Gorgojo L, Maisonneuve P, Fernandez-Rodriguez JC, Salvini S, et al. Development and validation of a food frequency questionnaire in Spain. *International Journal of Epidemiology* 1993; **22**(3): 512–9.
- 15 Goldbohm RA, van't Veer P, van den Brandt PA, van't Hof MA, Brants HA, Sturmans F, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. *European Journal of Clinical Nutrition* 1995; **49**(6): 420–9.

- 16 Mannisto S, Virtanen M, Mikkonen T, Pietinen P. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. *Journal of Clinical Epidemiology* 1996; **49**(4): 401–9.
- 17 Friis S, Kruger Kjaer S, Stripp C, Overvad K. Reproducibility and relative validity of a self-administered semiquantitative food frequency questionnaire applied to younger women. *Journal of Clinical Epidemiology* 1997; **50**(3): 303–11.
- 18 EPIC Group of Spain. Relative validity and reproducibility of a diet history questionnaire in Spain. II. Nutrients. *International Journal of Epidemiology* 1997; **26**(Suppl. 1): S100–9.
- 19 Riboli E, Toniolo P, Kaaks R, Shore RE, Casagrande C, Pasternack BS. Reproducibility of a food frequency questionnaire used in the New York University Women's Health Study: effect of self-selection by study subjects. *European Journal of Clinical Nutrition* 1997; **51**(7): 437–42.
- 20 Egami I, Wakai K, Kato K, Lin Y, Kawamura T, Tamakoshi A, *et al.* A simple food frequency questionnaire for Japanese diet – Part II. Reproducibility and validity for nutrient intakes. *Journal of Epidemiology* 1999; **9**(4): 227–34.
- 21 Shimizu H, Ohwaki A, Kurisu Y, Takatsuka N, Ido M, Kawakami N, *et al.* Validity and reproducibility of a quantitative food frequency questionnaire for a cohort study in Japan. *Japanese Journal of Clinical Oncology* 1999; **29**(1): 38–44.
- 22 Larkin FA, Metzner HL, Thompson FE, Flegal KM, Guire KE. Comparison of estimated nutrient intakes by food frequency and dietary records in adults. *Journal of the American Dietetic Association* 1989; **89**(2): 215–23.
- 23 Beaton G. Interpretation of results from diet history studies. In: Kohlmeier L, ed. *The Diet History Method*. London: Smith-Gordon, 1991; 15–38.
- 24 Chen J, Campbell TC, Li J, Peto R. *Diet, Life-style and Mortality in China. A Study of the Characteristics of 65 Chinese Counties*, 1st ed. Oxford: Oxford University Press, 1990.
- 25 Willett WC, Sampson L, Browne ML, Stampfer MJ, Rosner B, Hennekens CH, *et al.* The use of a self-administered questionnaire to assess diet four years in the past. *American Journal of Epidemiology* 1988; **127**(1): 188–99.
- 26 Block G, Woods M, Potosky A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. *Journal of Clinical Epidemiology* 1990; **43**(12): 1327–35.
- 27 Tjonneland A, Overvad K, Haraldsdottir J, Bang S, Ewertz M, Jensen OM. Validation of a semiquantitative food frequency questionnaire developed in Denmark. *International Journal of Epidemiology* 1991; **20**(4): 906–12.
- 28 Callmer E, Riboli E, Saracci R, Akesson B, Lindgarde F. Dietary assessment methods evaluated in the Malmo food study. *Journal of Internal Medicine* 1993; **233**(1): 53–7.
- 29 Horwath CC. Validity of a short food frequency questionnaire for estimating nutrient intake in elderly people. *British Journal of Nutrition* 1993; **70**(1): 3–14.
- 30 Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, *et al.* Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24h recalls, food-frequency questionnaires and estimated-diet records. *British Journal of Nutrition* 1994; **72**(4): 619–43.
- 31 Goldbohm RA, van den Brandt PA, Brants HA, van't Veer P, Al M, Sturmans F, *et al.* Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *European Journal of Clinical Nutrition* 1994; **48**(4): 253–65.
- 32 Jain M, Howe GR, Rohan T. Dietary assessment in epidemiology: comparison of food frequency and a diet history questionnaire with a 7-day food record. *American Journal of Epidemiology* 1996; **143**(9): 953–60.
- 33 Riboli E, Elmstahl S, Saracci R, Gullberg B, Lindgarde F. The Malmo Food Study: validity of two dietary assessment methods for measuring nutrient intake. *International Journal of Epidemiology* 1997; **26**(1): S161–73.
- 34 Baumgartner KB, Gilliland FD, Nicholson CS, McPherson RS, Hunt WC, Pathak DR, *et al.* Validity and reproducibility of a food frequency questionnaire among Hispanic and non-Hispanic white women in New Mexico. *Ethnicity & Disease* 1998; **8**(1): 81–92.
- 35 Klipstein Grobusch K, den Breeijen JH, Goldbohm RA, Geleijnse JM, Hofman A, Grobbee DE, *et al.* Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire. *European Journal of Clinical Nutrition* 1998; **52**(8): 588–96.
- 36 Wolk A, Ljung H, Vessby B, Hunter D, Willett WC. Effect of additional questions about fat on the validity of fat estimates from a food frequency questionnaire. *European Journal of Clinical Nutrition* 1998; **52**(3): 186–92.
- 37 Jain M, McLaughlin J. Validity of nutrient estimates by food frequency questionnaires based either on exact frequencies or categories. *Annals of Epidemiology* 2000; **10**(6): 354–60.
- 38 Brunner E, Stallone D, Juneja M, Bingham S, Marmot M. Dietary assessment in Whitehall II: comparison of 7 d diet diary and food-frequency questionnaire and validity against biomarkers. *British Journal of Nutrition* 2001; **86**(3): 405–14.
- 39 Parent M, Kronld M, Chow R. Reconstruction of past calcium intake patterns during adulthood. *Journal of the American Dietetic Association* 1993; **93**(6): 649–52.
- 40 Barrett Connor E. Nutrition epidemiology: how do we know what they ate? *American Journal of Clinical Nutrition* 1991; **54**(Suppl. 1): 182S–7S.
- 41 Eck LH, Klesges RC, Hanson CL, Slawson D, Portis L, Lavasque ME. Measuring short-term dietary intake: development and testing of a 1-week food frequency questionnaire. *Journal of the American Dietetic Association* 1991; **91**(8): 940–5.
- 42 The World Bank. World development indicators. In: *World Development Report 1993: Investing in Health*. New York: Oxford University Press, 1993; 227–304.
- 43 Woo J, Leung S. A food frequency questionnaire for use in the Chinese population in Hong Kong: description and examination of validity. *Nutrition Research* 1997; **17**: 1633–41.
- 44 Hankin J. Development of a diet history questionnaire for studies of older persons. *American Journal of Clinical Nutrition* 1989; **50**(Suppl. 5): 1121–7.
- 45 Pietinen P, Hartman AM, Haapa E, Rasanen L, Haapakoski J, Palmgren J, *et al.* Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. *American Journal of Epidemiology* 1988; **128**(3): 655–66.
- 46 Willett W, Lenart E. Reproducibility and validity of food-frequency questionnaires. In: Willett W, ed. *Nutritional Epidemiology*, 2nd ed. New York: Oxford University Press, 1998; 101–47.