

Reproducibility and validity of an FFQ developed for adults in Nanjing, China

Qing Ye¹, Xin Hong¹, Zhiyong Wang¹, Huafeng Yang¹, Xupeng Chen^{1,2}, Hairong Zhou^{1,2},
Chenchen Wang^{1,2}, Yichao Lai³, Liuyuan Sun⁴ and Fei Xu^{1,2*}

¹Department of Non-communicable Disease Prevention, Nanjing Municipal Center for Disease Control & Prevention, 2, Zizhulin, Nanjing 211166, People's Republic of China

²Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, 818 East Tianyuan Road, Nanjing 211166, People's Republic of China

³Department of Non-communicable Disease Prevention, Qinhuai District Center for Disease Control & Prevention, 2, Baiguosbu, Nanjing 210029, People's Republic of China

⁴Department of Non-communicable Disease Prevention, Liube District Center for Disease Control & Prevention, East Huancheng Road of Xiongzhou, Nanjing 211500, People's Republic of China

(Submitted 4 September 2015 – Final revision received 26 November 2015 – Accepted 4 December 2015)

Abstract

We evaluated the reproducibility and validity of an FFQ used in a Chinese community-based nutrition and health survey. A total of ninety-nine males and 104 females aged 31–80 years completed four three consecutive 24-h dietary recalls (24-HDR, served as a reference method, one three consecutive 24-HDR for each season) and two FFQ (FFQ1 and FFQ2) over a 1-year interval. The reproducibility of the FFQ was estimated with correlation coefficients, misclassification and weighted κ statistic. The validity was evaluated by comparing the data obtained from FFQ2 with the mean 24-HDR (m24-HDR). Compared with the m24-HDR, the FFQ tended to underestimate intake of most nutrients and food groups. For all nutrients and food groups, the Spearman's and intra-class correlation coefficients between FFQ1 and FFQ2 ranged from 0.66 to 0.88 and from 0.65 to 0.87, respectively. Most correlation coefficients decreased after adjusting for energy. More than 90% of the subjects were classified into the same or adjacent categories by both FFQ. For all nutrients and food groups, the crude, energy-adjusted and de-attenuated Spearman's correlation coefficients between FFQ2 and the m24-HDR ranged from 0.21 to 0.69, 0.19 to 0.58 and 0.25 to 0.71, respectively. More than 70% of the subjects were classified into the same and adjacent categories by both instruments. Both weighted κ statistic and the Bland–Altman plots showed reasonably acceptable agreement between the FFQ2 and the m24-HDR. The FFQ developed for adults in the Nanjing area can be used to reliably and validly measure usual intake of major nutrients and food groups.

Key words: FFQ; Reproducibility; Validity; China

Accumulated evidence suggests that nutrient intakes are associated with the development of many diseases, especially chronic conditions^(1,2). In order to investigate this relationship, it is important to accurately assess food and nutrient intakes. The FFQ is a useful tool for the estimation of food and nutrient consumption and has been widely used in investigating the association between diet and chronic diseases in most epidemiological studies^(3,4). FFQ are easy to administer and relatively inexpensive to use in large population studies. However, the performance of an FFQ is very sensitive to the ethnic, social and cultural backgrounds of the study population⁽⁵⁾. For this reason, the reliability and the validity of an FFQ need to be evaluated for studies conducted in a new population.

Reliability refers to the consistency of measurements on repetition, whereas validity refers to the ability to measure what

the FFQ was designed to measure. At present, there is no perfect standard for the validation of dietary intake and a superior measurement is always used for comparison; 24-h dietary recalls (24-HDR) may be superior to FFQ and have been frequently used as the reference method in many Chinese validation studies^(6–9). A critical review regarding validation of FFQ has shown that FFQ are validated against another dietary method in 75% of studies. When co-operation or literacy of the study subjects is limited, 24-h recalls may be more appropriate⁽¹⁰⁾.

As a developed economic region in China, Nanjing is experiencing a high prevalence of chronic non-communicable diseases such as hypertension, diabetes and cancer among its population. A large community-based cross-sectional nutrition and health survey has been initiated in 2014 in the Nanjing area. The main purposes of this study were to collect information on

Abbreviations: 24-HDR, 24-h dietary recalls; LOA, limits of agreement; m24-HDR, mean 24-HDR.

* **Corresponding author:** F. Xu, fax +86 25 8353 8342, email frankxu@163.com

lifestyle factors including dietary habits and to observe their effects on the occurrence of chronic diseases. We developed a new FFQ to estimate the nutrient and food group intakes of this population. For validation of this FFQ, we gathered information using three consecutive 24-HDR as a reference method during each of the four seasons over a period of 1 year, totalling up to 12 d for 203 participants. In this study, we report the reliability and validity of this FFQ for use in a large community-based nutrition and health survey.

Methods

Study population

The subjects were recruited using a multi-stage stratified random sampling method (Fig. 1). First, we randomly selected two districts (one urban and one suburban). Next, three streets/towns from each chosen district were randomly selected. Finally, one community from each chosen street/town was randomly selected. This resulted in a total number of six communities. We then randomly selected 250 eligible residents from the six communities to participate in the validation study. The inclusion criteria were as

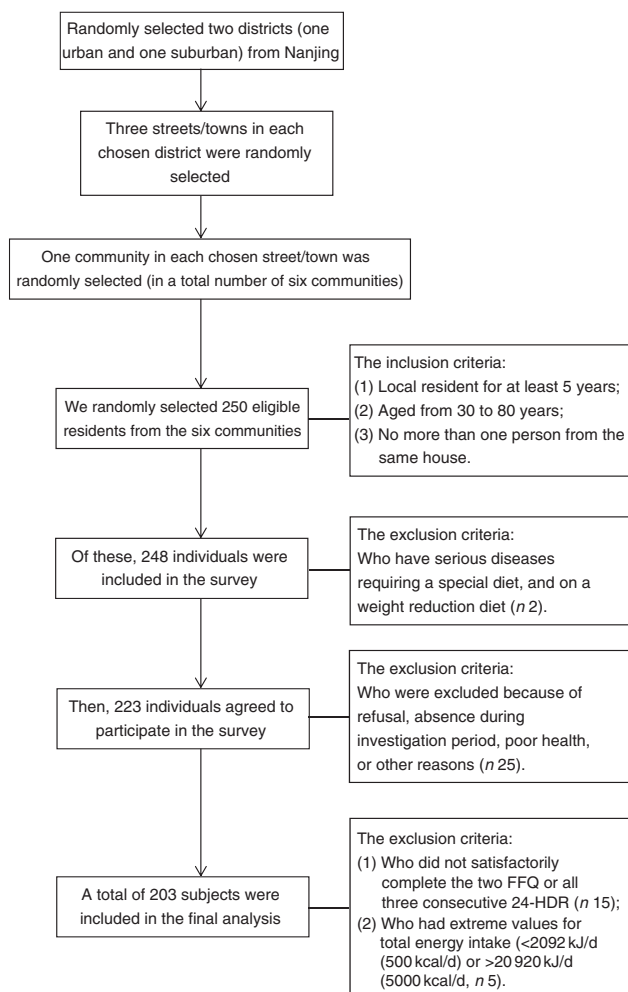


Fig. 1. Participant recruitment flow diagram. 24-HDR, 24-h dietary recalls.

follows: local resident for at least 5 years, aged between 30 and 80 years, free of serious diseases requiring a special diet and not on a weight-reduction diet. Among the 250 selected residents, 248 were eligible to participate and 223 agreed to take part in the study and completed the survey (response rate = 89.9%). The main reasons for not participating in the study included refusal, absence during the investigation period and poor health.

Ethics approval for this study was obtained from the academic and ethical committee of Nanjing Municipal Center for Disease Control and Prevention (Nanjing CDC). All the participants provided their written informed consent before the survey.

Study design

Each participant completed the same FFQ twice (FFQ1: the first FFQ and FFQ2: the second FFQ), 1 year apart. Four three consecutive 24-HDR were collected at 3-month intervals during the period of 1 year. The first three consecutive 24-HDR was obtained 1 month after the administration of FFQ1 (in June 2014), and the last three consecutive 24-HDR was obtained 1 month before the administration of FFQ2 (in May 2015). Participants who did not satisfactorily complete the two FFQ or all three consecutive 24-HDR (n 15) and those who had extreme values for total energy intake (<2092 kJ/d (500 kcal/d) or $>20\,920$ kJ/d (5000 kcal/d), n 5)⁽¹¹⁾ were excluded from this study. Thus, a total of 203 subjects (81.9%) were included in the final analysis.

FFQ and 24-h dietary recalls

The FFQ included eighty-seven food items and ten food categories (cereal; red meat (pork, beef, mutton); poultry; fish and shrimp; eggs; dairy products; soya-based foods; vegetables; fruits; snacks/desserts), which covered about 90% of the commonly consumed foods in Nanjing. For each food item, participants were asked to recall the frequency of consumption (daily, weekly, monthly, annually or never) and the amount of consumption using a common unit of weight in China (1 liang = 50 g) over the past 12 months. Individual consumption of food items was converted to grams per day in the further analysis. According to the similarity of nutrient profiles and culinary usage among the foods and the grouping scheme used in other studies, we aggregated the eighty-seven food items into thirty pre-defined food groups^(12,13). For seasonal vegetables and fruits, participants were asked to recall how often they ate these foods during the season. The *Chinese Food Composition Table*⁽¹⁴⁾ were used to estimate the daily energy intake (kJ/d (kcal/d)) and major nutrients of study participants.

The three consecutive 24-HDR were administered for 2 weekdays and 1 weekend day in a usual week. Each participant was asked to provide the name and amount of all foods consumed during the previous 24 h. If the previous day was a special day owing to feasts or travels, food consumption data of the day before the 24 h were recorded or another day was chosen to interview the participant by telephone. The amounts of different food items that were mixed in one dish were recorded respectively. The recalled food items were assigned to the corresponding food groups as defined by the FFQ. We calculated daily mean intakes of energy (kJ/d (kcal/d)), thirteen nutrients and

ten food groups estimated by four three consecutive 24-HDR. The mean 24-HDR (m24-HDR) data were used as the standard to measure the relative validity of the FFQ.

Trained interviewers from the local CDC administered the two FFQ and the four three consecutive 24-HDR by face-to-face interviews. All diet information was collected and checked after completion. Any implausible or ambiguous information was further verified and obtained from the participants. Each participant had the same interviewer during the study period.

Statistical analysis

Medians and inter-quartile ranges for energy, nutrients and food intake were calculated for both FFQ and 24-HDR. The reproducibility was estimated using the Wilcoxon's signed-rank test, Spearman's correlation, intra-class correlation coefficient (ICC), weighted κ statistic and misclassification (quartiles method) analyses to compare the intakes from FFQ1 and FFQ2⁽¹⁵⁾. The validity of the FFQ relative to the 24-HDR was assessed by Wilcoxon's signed-rank test, Spearman's correlation, weighted κ statistic and misclassification (quartiles method) analyses. The residual method was used to exclude the possibility of variation due to energy intake⁽¹⁶⁾. To correct for within-person error in the measurement of the HDR, the observed correlation was multiplied by the de-attenuation factor $(1+\gamma/n)^{1/2}$, where γ is the ratio of the within- and between-person variances and n is the number of 24-HDR (here n 12).

Bland-Altman plots were used to examine the agreement between two dietary assessment methods across a range of intakes. The difference between the two methods was plotted against the average of the two methods. The mean difference and the 95% limits of agreement (LOA), calculated as mean difference ± 1.96 (SD of differences), were used to summarise agreement at the population level. Natural-log(ln) transformations were performed in order to narrow the LOA, as recommended by Bland & Altman⁽¹⁷⁾.

All the statistical analyses were performed using SPSS (version 20.0) and MedCalc (version 11.4). A P value < 0.05 was considered to be statistically significant.

Results

Of the 203 participants eligible for the present study, 48.8% were males. The mean age was 50.4 (SD 12.2) years (range 31–80 years); the mean BMI was 23.1 (SD 2.8) kg/m²; and 79.5% had education of junior high school or below. The proportions of current smokers and drinkers were 22.0 and 28.8%, respectively (Table 1).

Table 2 presents the median intakes of total energy, nutrients and food groups derived from the FFQ, m24-HDR, the percentage of differences and the results from the Wilcoxon's signed-rank test. The Wilcoxon's signed-rank test showed that the intakes of almost all nutrients and food groups assessed by the two FFQ were not significantly different, except for fat ($P=0.035$), retinol ($P=0.010$) and vitamin B₁ ($P=0.020$). The median intakes for all nutrients and food groups assessed using

Table 1. Characteristics of the participants in the validation study* (Mean values and standard deviations; or percentages)

Variables	<i>n</i> 203
Age at recruitment (years)	
Mean	50.4
SD	12.2
Male (%)	48.8
Weight (kg)	
Mean	63.5
SD	10.1
BMI (kg/m ²)	
Mean	23.1
SD	2.8
Waist (cm)	
Mean	84.4
SD	15.3
Marriage status (%)	
Married	92.1
Unmarried/divorced/widowed	7.9
Education (%)	
Primary school and lower	20.5
Junior high school	32.2
Senior high school	24.4
College and higher	22.9
Occupation (%)	
Manual labourers	42.4
Service staff	6.8
Mental labourers	25.9
Others	24.9
Current smokers (%)	22.0
Current drinkers (%)	28.8

* Data were from FFQ1.

FFQ2 were higher or equal to the values obtained using FFQ1, with differences in median intakes between 0 and -25.4%. There was also no significant difference between intakes of nutrients and food groups assessed by FFQ2 and m24-HDR (both assessments cover the same time period), with the exception of fibre ($P<0.001$), retinol ($P=0.014$), carotene ($P=0.005$), eggs ($P=0.022$) and soya-based foods ($P<0.001$). Compared with the m24-HDR, the FFQ tended to underestimate intakes of most nutrients and food groups.

The crude- and energy-adjusted correlation coefficients for FFQ1 and FFQ2 are presented in Table 3. For total energy, nutrients and food groups, crude Spearman's correlation coefficients ranged from 0.66 for retinol and eggs to 0.88 for vitamin C, and the crude ICC ranged from 0.65 for fat to 0.87 for vitamin C. After adjusting for energy, most correlation coefficients decreased. The energy-adjusted Spearman's correlation coefficients ranged from 0.41 (poultry) to 0.83 (carbohydrate), whereas the energy-adjusted ICC ranged from 0.45 (fruits) to 0.80 (soya-based foods). The crude, energy-adjusted and de-attenuated Spearman's correlation coefficients between the FFQ and the m24-HDR are presented in Table 3. These values enable the assessment of the relative validity of the FFQ. The crude Spearman's correlation coefficients for FFQ2 and the m24-HDR ranged from 0.21 (soya-based foods) to 0.69 (fat, retinol). The energy-adjusted coefficients ranged from 0.19 (soya-based foods) to 0.58 (fat, vitamin C), whereas the de-attenuated coefficients ranged from 0.25 (soya-based foods) to 0.71 (fat).

Table 2. Comparison of nutrient and food group intakes between two FFQ and mean 24-h dietary recall (m24-HDR) among 203 Chinese participants (Medians and 25th–75th percentiles)

Daily intakes	FFQ1		FFQ2		m24-HDR		Wilcoxon's test (<i>P</i>)		Percentages of median difference	
	Median	25th–75th percentile	Median	25th–75th percentile	Median	25th–75th percentile	FFQ1 v. FFQ2	FFQ2 v. m24-HDR	FFQ1 v. FFQ2	FFQ2 v. m24-HDR
Energy intake (kcal)	1309	872, 1768	1382	1013, 1905	1482	1061, 1938	0.055	0.448	–5.3	–6.7
Protein (g)	54.9	37.8, 92.0	59.7	43.0, 96.7	57.3	40.3, 97.7	0.087	0.602	–8.0	4.2
Fat (g)	33.7	21.6, 47.3	35.8	25.3, 52.2	40.5	29.5, 58.3	0.035	0.059	–5.9	–11.6
Carbohydrate (g)	176.2	131.7, 251.4	192.5	142.0, 251.9	187.3	137.7, 261.9	0.156	0.939	–8.5	2.8
Fibre (g)	12.0	7.5, 25.4	12.4	8.5, 26.8	8.1	5.6, 27.5	0.407	<0.001	–3.2	53.1
Ca (mg)	558	392, 917	595	432, 950	591	403, 991	0.140	0.931	–6.2	0.7
P (mg)	799	575, 1087	870	655, 1161	840	628, 1063	0.083	0.377	–8.2	3.6
Retinol (µg)	104.7	55.2, 158.8	121.2	68.6, 182.5	137.7	98.0, 194.8	0.010	0.014	–13.6	–12.0
Carotene (µg)	665.9	389.6, 3860.4	776.0	469.6, 5998.3	615.0	349.2, 8096.5	0.070	0.005	–14.2	26.2
Vitamin B ₁ (mg)	0.61	0.47, 0.83	0.67	0.50, 0.87	0.74	0.52, 1.00	0.020	0.109	–9.0	–9.5
Vitamin B ₂ (mg)	0.82	0.57, 1.22	0.89	0.68, 1.33	0.94	0.68, 1.45	0.056	0.486	–7.9	–5.3
Niacin (mg)	13.8	10.2, 22.9	15.5	11.3, 24.5	15.1	11.4, 25.9	0.057	0.691	–11.0	2.6
Vitamin C (mg)	99.0	54.1, 160.7	103.5	65.6, 162.1	102.0	78.7, 164.3	0.160	0.149	–4.3	1.5
Vitamin E (mg)	12.5	8.0, 23.1	13.2	9.2, 25.5	12.3	7.0, 26.2	0.207	0.153	–5.3	7.3
Cereal (g)	342.9	282.1, 437.5	353.6	304.5, 430.4	333.3	300.0, 400.0	0.601	0.091	–3.1	6.1
Red meat (g)	38.1	21.4, 57.1	42.9	23.2, 78.6	50.0	25.0, 400.0	0.135	0.756	–12.6	–14.2
Poultry (g)	21.2	15.7, 39.3	21.4	21.4, 50.0	25.0	16.7, 42.9	0.087	0.248	–0.9	–14.4
Fish and shrimp (g)	21.4	7.1, 32.1	21.4	7.1, 32.1	21.4	7.1, 33.1	0.186	0.678	0.0	0.0
Eggs (g)	25.0	10.7, 50.0	25.0	10.7, 50.0	33.3	16.7, 50.0	0.051	0.022	0.0	–24.9
Dairy products (g)	52.9	5.7, 114.3	53.6	7.1, 160.7	100.0	0, 200.0	0.438	0.081	–1.3	–46.4
Soya-based foods (g)	39.3	17.9, 68.9	39.3	23.2, 71.4	21.4	0, 50.0	0.295	<0.001	0.0	83.6
Vegetables (g)	253.6	153.6, 503.6	289.3	173.2, 500.0	300	217.3, 416.7	0.269	0.315	–14.0	–3.6
Fruits (g)	95.0	32.1, 150.0	100.0	32.1, 150.0	100	50.0, 150.0	0.523	0.146	–5.3	0.0
Snacks/desserts (g)	7.1	2.7, 17.9	8.9	3.6, 18.6	11.2	0, 66.7	0.188	0.445	–25.4	–20.5

Table 3. Correlation coefficients for nutrient and food group intakes between the two FFQ and the second FFQ with mean 24-h dietary recall (m24-HDR)

Variables	FFQ1 v. FFQ2				FFQ2 v. m24-HDR		
	Spearman		ICC		Spearman		
	Crude	Energy adjusted	Crude	Energy adjusted	Crude	Energy adjusted	De-attenuated
Energy intake (kcal)	0.77	–	0.77	–	0.66	–	0.64
Protein (g)	0.79	0.55	0.81	0.54	0.61	0.44	0.60
Fat (g)	0.68	0.60	0.65	0.61	0.69	0.58	0.71
Carbohydrate (g)	0.85	0.83	0.81	0.74	0.60	0.41	0.63
Fibre (g)	0.87	0.72	0.79	0.72	0.48	0.36	0.53
Ca (mg)	0.80	0.69	0.79	0.70	0.53	0.38	0.51
P (mg)	0.79	0.62	0.79	0.58	0.65	0.40	0.61
Retinol (µg)	0.66	0.44	0.72	0.49	0.69	0.48	0.69
Carotene (µg)	0.83	0.65	0.80	0.62	0.46	0.33	0.48
Vitamin B ₁ (mg)	0.74	0.54	0.73	0.52	0.66	0.54	0.64
Vitamin B ₂ (mg)	0.74	0.59	0.82	0.58	0.59	0.39	0.60
Niacin (mg)	0.84	0.72	0.81	0.66	0.60	0.37	0.59
Vitamin C (mg)	0.88	0.62	0.87	0.73	0.66	0.58	0.69
Vitamin E (mg)	0.87	0.67	0.80	0.73	0.56	0.38	0.58
Cereal (g)	0.86	0.78	0.78	0.76	0.43	0.44	0.43
Red meat (g)	0.73	0.45	0.72	0.61	0.39	0.31	0.40
Poultry (g)	0.70	0.41	0.66	0.54	0.48	0.35	0.45
Fish and shrimp (g)	0.79	0.60	0.77	0.50	0.44	0.29	0.40
Eggs (g)	0.66	0.64	0.67	0.62	0.36	0.33	0.39
Dairy products (g)	0.72	0.55	0.70	0.51	0.29	0.34	0.31
Soya-based foods (g)	0.82	0.76	0.79	0.80	0.21	0.19	0.25
Vegetables (g)	0.86	0.69	0.86	0.62	0.44	0.37	0.43
Fruits (g)	0.71	0.56	0.69	0.45	0.47	0.40	0.51
Snacks/desserts (g)	0.80	0.71	0.81	0.77	0.32	0.36	0.33

ICC, intra-class correlation coefficients.

Table 4. Misclassification and weighted κ between the two FFQ and the second FFQ with mean 24-h dietary recall (m24-HDR)

	Per cent of agreement								Weighted κ	
	Same quartile		Adjacent quartile		One quartile apart		Opposite quartile			
	FFQ1 v. FFQ2	FFQ2 v. m24-HDR	FFQ1 v. FFQ2	FFQ2 v. m24-HDR	FFQ1 v. FFQ2	FFQ2 v. m24-HDR	FFQ1 v. FFQ2	FFQ2 v. m24-HDR	FFQ1 v. FFQ2	FFQ2 v. m24-HDR
Energy intake (kcal)	63.9	49.8	29.2	39.0	5.9	9.2	1.0	2.0	0.65	0.49
Protein (g)	70.7	51.7	21.9	36.6	6.4	8.3	1.0	3.4	0.70	0.49
Fat (g)	61.1	54.6	29.7	33.7	5.3	9.7	3.9	2.0	0.58	0.53
Carbohydrate (g)	73.2	45.4	21.5	42.5	4.9	9.1	0.4	3.0	0.74	0.44
Fibre (g)	77.5	45.4	19.0	36.6	3.0	12.2	0.5	5.8	0.79	0.37
Ca (mg)	67.8	50.2	23.4	33.7	7.3	10.7	1.5	5.4	0.66	0.43
P (mg)	68.8	45.4	24.9	42.9	5.4	9.7	1.0	2.0	0.69	0.45
Retinol (μ g)	58.5	49.8	32.6	38.5	4.9	10.2	4.0	1.5	0.57	0.49
Carotene (μ g)	68.2	48.8	24.9	31.7	5.9	12.6	1.0	6.9	0.68	0.38
Vitamin B ₁ (mg)	70.8	46.3	21.4	32.0	5.4	9.7	2.4	2.0	0.69	0.46
Vitamin B ₂ (mg)	61.5	47.8	30.3	35.2	7.7	14.0	0.5	3.0	0.62	0.42
Niacin (mg)	70.2	47.8	24.4	37.0	5.4	12.2	0.0	3.0	0.72	0.44
Vitamin C (mg)	79.0	50.2	17.5	37.1	3.0	10.7	0.5	2.0	0.80	0.48
Vitamin E (mg)	73.7	44.9	21.4	37.6	4.9	14.1	0.0	3.4	0.75	0.39
Cereal (g)	77.6	41.0	16.5	35.0	4.4	18.6	1.5	5.4	0.76	0.30
Red meat (g)	69.3	31.7	24.9	46.8	2.5	16.1	3.3	5.4	0.68	0.23
Poultry (g)	72.7	41.0	27.3	39.1	0.0	15.1	0.0	4.8	0.74	0.30
Fish and shrimp (g)	83.4	42.0	10.8	34.1	7.8	14.1	3.0	9.8	0.79	0.27
Eggs (g)	76.6	29.3	16.6	44.9	3.9	22.8	2.9	3.0	0.62	0.24
Dairy products (g)	73.2	33.2	18.5	39.1	3.4	20.5	4.9	7.2	0.68	0.18
Soya-based foods (g)	80.5	29.9	13.7	41.9	3.8	18.0	2.0	10.2	0.78	0.16
Vegetables (g)	79.5	39.1	16.6	41.9	2.9	13.7	1.0	5.3	0.80	0.32
Fruits (g)	60.5	42.4	31.2	35.1	3.4	13.2	4.9	9.3	0.57	0.35
Snacks/desserts (g)	67.3	38.7	25.8	40.5	4.9	14.6	2.0	6.2	0.67	0.29

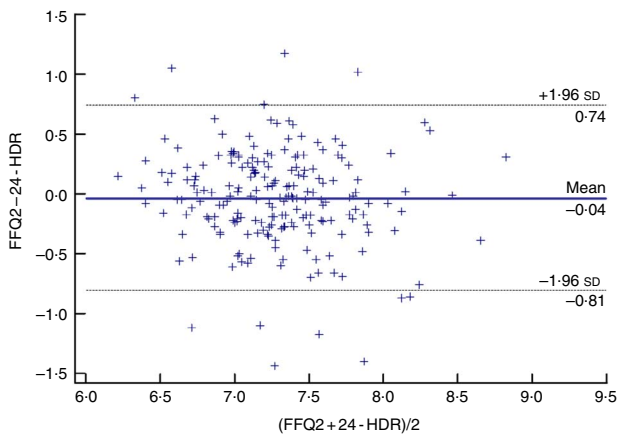


Fig. 2. The Bland–Altman plot for total energy intake. 24-HDR, 24-h dietary recalls.

When the intakes were categorised into quartiles, the ranges of agreement rates for the same or adjacent quartile classifications were 90.8–100%, when derived from the two FFQ, and 71.8–89.3%, when derived from FFQ2 and the m24-HDR. Extreme misclassification into opposite quartiles was <10.0% for all nutrients and food groups, with the exception of soya-based foods (10.2%). The weighted κ statistic showed moderate conformity, ranging from 0.57 to 0.80 for the two FFQ and from 0.16 to 0.53 for FFQ2 and the m24-HDR (Table 4). No difference was observed in the reproducibility and validity between men and women.

The Bland–Altman plots for total energy intake, protein, fat and carbohydrates are presented in Fig. 2–5. Anti-logging rendered mean difference and LOA of 96.1% (95% CI 44.5, 209.5), 99.0% (95% CI 32.0, 309.6), 94.2% (95% CI 39.5, 224.0) and 97.7% (95% CI 42.7, 222.5), respectively, for energy intake, protein, fat and carbohydrates. For almost all nutrients, <10% of the subjects were outside the LOA.

Discussion

In this study, we examined the reproducibility and validity of an FFQ used for the Nanjing cross-sectional nutrition and health survey. According to a previous review⁽¹⁰⁾, the number of food items listed in FFQ should range from 5 to 350. The FFQ used in this study was composed of eighty-seven food items, which covered about 90% of the commonly consumed foods in Nanjing. We evaluated the performance of the FFQ by comparing intake of nutrients and selected food groups obtained from this instrument with those derived from the 24-HDR.

In the present study, the median intakes for almost all nutrients and food groups obtained from FFQ2 were higher than or equal to the values obtained from FFQ1. This might be due to the learning effect. That is to say participants might be more mindful of what they ate, and thus estimated the amount more appropriately after the previous surveys. Spearman’s correlation coefficients for reproducibility in our study ranged from 0.66 to 0.86 for food groups and from 0.66 to 0.88 for nutrients. The ICC between FFQ1 and FFQ2 were 0.66–0.86 for

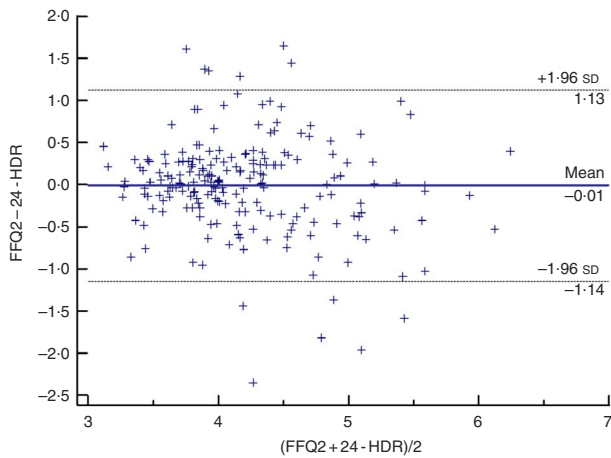


Fig. 3. The Bland–Altman plot for protein intake. 24-HDR, 24-h dietary recalls.

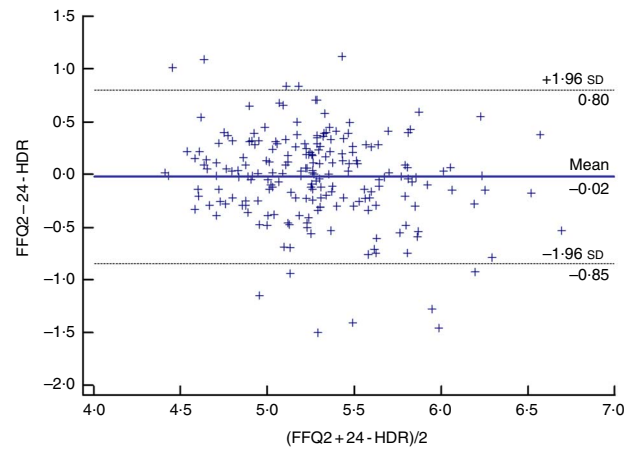


Fig. 5. The Bland–Altman plot for carbohydrate intake. 24-HDR, 24-h dietary recalls.

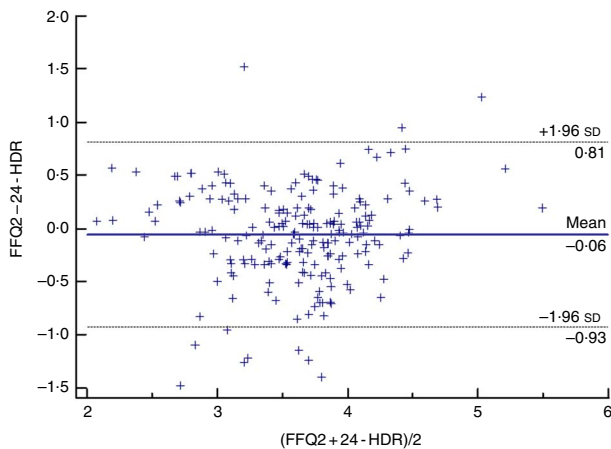


Fig. 4. The Bland–Altman plot for fat intake. 24-HDR, 24-h dietary recalls.

food groups and 0.65–0.88 for nutrients. Compared with other studies^(7–9,18–23), in which the correlation coefficients generally ranged from 0.20 to 0.80, the correlation coefficients in the present study were slightly higher. This may reflect the fact that the dietary habits of individuals in Nanjing are relatively stable. Energy adjustment did not improve the correlations for nutrients and food items. If the variability of nutrient consumption is related to energy intake, energy adjustment may increase correlation coefficients.

Various time intervals between FFQ1 and FFQ2, from 15 d to several years, have been reported in previous studies^(6,24). Reproducibility tests are based on the assumption that diet does not change between two questionnaires; thus, reproducibility may ideally be obtained by two administered questionnaire surveys with a short interval^(25–28). However, subjects are more likely to remember and repeat their responses in that case. In this study, we administered FFQ1 and FFQ2 with an almost 1-year interval, which can reduce the above-mentioned error. As categorised dietary intake rather than the absolute amount of intake has been more commonly used in epidemiological studies of diet and chronic diseases, we performed misclassification analyses, showing that the percentages of participants correctly classified

into the same or adjacent categories and the weighted κ values in our study were higher than those reported by other validation studies^(8,9). In the Suihua female adolescent study, the agreement rates for classifying nutrient and food group intakes into the same or adjacent categories were 70.8–92.9%, and the weighted κ values were 0.35–0.60⁽⁹⁾. In another study, the rates ranged from 73.0 to 86.0% for the agreement and from 0.20 to 0.50 for the weighted κ values⁽⁸⁾.

The results of estimated relative validity depend on several factors such as choice of reference method, degree of homogeneity of the intake values within the population, recall period and the number of days of record collection⁽²⁹⁾. Selecting the appropriate reference method by which to assess the test measurement is a major part of the validation process. In a review on the validation of FFQ, the authors showed that 75% of studies validated the FFQ against another dietary method. When co-operation or literacy of study subjects is limited, 24-h recalls may be more appropriate⁽¹⁰⁾. Moreover, it is important that the measurement errors of the FFQ and reference method should be independent. In the present study, relative validity has been tested by comparing FFQ2 with the average of the four three consecutive 24-HDR, one for each season. The multiple recalls were able to minimise the effect of daily and seasonal dietary intake variation on the dietary assessment. Compared with the 24-HDR, the FFQ tends to underestimate intake of most nutrients. The mean differences shown in Bland–Altman plots were all negative. We also found that the FFQ underestimated the intakes of red meat, poultry, egg and dairy products. These results are very similar to the Shanghai men's study⁽⁷⁾. Some of the measurement error may reflect a bias of study participants seeking social approval⁽³⁰⁾.

The correlation coefficients of our study were consistent with the results reported in other Chinese population studies, which ranged from 0.15 to 0.72^(6–9). The validity correlation decreased for most of the nutrients and food groups after adjustments for energy. This might be due to the between-person variation in the intakes of nutrients and food groups in subjects. There was very little change in the correlations of nutrient and food intakes assessed by the FFQ and averaged 24-HDR after we further adjusted for the within-person variation from the multiple 24-HDR. If the frequency of consumption is low and the within-person variability

is too high, the correlation coefficients can be attenuated⁽³¹⁾. Our results indicate that the dietary intake of individuals in Nanjing did not vary significantly during the four seasons; this may be because Nanjing is situated in the Yangtze River Delta, with easy access and a well-developed infrastructure for the movement of goods, so that the same foods can be purchased from the market throughout the year. The misclassification analyses showed that more than 70% of the subjects were classified into the same or adjacent quartile for food group and nutrient intakes by both methods, which compares well with other studies^(6–9,32,33). The weighted κ value for most nutrients and food groups reached the ‘acceptable’ threshold.

Natural-log (ln) transformations were performed by Bland–Altman analysis in order to narrow the LOA. A wide LOA indicates that the potential for large differences between methods and agreement is considered poor. The LOA of our study were wider than those reported in two previous studies^(7,8); however, the mean differences of nutrient intakes were approximately 0. This may indicate that the FFQ is more suitable for ranking intakes than estimating absolute intakes, and the misclassification in nutrient intake is less likely to cause systematic biases.

There are several limitations to our study. First, due to the lack of a perfect standard for measuring dietary intake to assess the validity of a dietary instrument, we chose dietary recalls as the reference method. This method was advantageous in its ability to collect actual intake on specific days, but also has some weak points such as under-reporting in dietary recalls. We attempted to minimise under-reporting by checking dietary recalls by following-up incomplete or ambiguous information directly with respondents. Second, the analysis of reproducibility and validity was confined to those adults aged 31–80 years. It is unclear whether our findings can be generalised to children, adolescents and the younger adult populations. Finally, the data would be more representative if three consecutive 24-HDR were collected monthly, instead of each season.

In summary, our study evaluated the reproducibility and validity of an eighty-seven-item FFQ developed specifically for investigating the relationship between dietary factors and chronic diseases in the Nanjing community-based cross-sectional study. The results of this study indicated that the FFQ can reasonably categorise usual intake of major nutrients and food groups among the study population, although it may not quantify the absolute intake of some nutrients or foods.

Acknowledgements

The authors are grateful to all the dedicated fieldworkers who took part in the survey, and all the participants who facilitated the survey implementation at each community.

The present study was supported by Nanjing Municipal Medical Science and Technique Development Foundation, China (2012-YKK12166).

Q. Y. and F. X. contributed to the study design, data analysis and manuscript writing; Q. Y., X. H., Z. W., H. Y., X. C., H. Z., C. W., Y. L. and L. S. were responsible for data collection; X. H., Z. W., H. Y., X. C., H. Z., C. W., Y. L. and L. S. were responsible for manuscript revision.

The authors report no conflicts of interest.

References

1. Cho SS, Qi L, Fahey GC Jr, *et al.* (2013) Consumption of cereal fiber, mixtures of whole grains and bran, and whole grains and risk reduction in type 2 diabetes, obesity, and cardiovascular disease. *Am J Clin Nutr* **98**, 594–619.
2. Rice BH, Quann EE & Miller GD (2013) Meeting and exceeding dairy recommendations: effects of dairy consumption on nutrient intakes and risk of chronic disease. *Nutr Rev* **71**, 209–223.
3. Willett WC, Sampson L, Stampfer MJ, *et al.* (1985) Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* **122**, 51–65.
4. Pietinen P, Hartman AM, Haapa E, *et al.* (1988) Reproducibility and validity of dietary assessment instruments. II. A qualitative food frequency questionnaire. *Am J Epidemiol* **128**, 667–676.
5. Sharma S, Cade J, Jackson M, *et al.* (1996) Development of food frequency questionnaires in three population samples of African origin from Cameroon, Jamaica and Caribbean migrants to the UK. *Eur J Clin Nutr* **50**, 479–486.
6. Shu XO, Yang G, Jin F, *et al.* (2004) Validity and reproducibility of the food frequency questionnaire used in the Shanghai Women’s Health Study. *Eur J Clin Nutr* **58**, 17–23.
7. Villegas R, Yang G, Liu D, *et al.* (2007) Validity and reproducibility of the food-frequency questionnaire used in the Shanghai men’s health study. *Br J Nutr* **97**, 993–1000.
8. Zhuang M, Yuan Z, Lin L, *et al.* (2012) Reproducibility and relative validity of a food frequency questionnaire developed for adults in Taizhou, China. *PLOS ONE* **7**, e48341.
9. Xia W, Sun C, Zhang L, *et al.* (2011) Reproducibility and relative validity of a food frequency questionnaire developed for female adolescents in Suihua, North China. *PLoS ONE* **6**, e19656.
10. Cade J, Thompson R, Burley V, *et al.* (2002) Development, validation and utilisation of food-frequency questionnaires – a review. *Public Health Nutr* **5**, 567–587.
11. Watson JF, Collins CE, Sibbritt DW, *et al.* (2009) Reproducibility and comparative validity of a food frequency questionnaire for Australian children and adolescents. *Int J Behav Nutr Phys Act* **6**, 62.
12. Shim JS, Oh K & Kim HC (2014) Dietary assessment methods in epidemiologic studies. *Epidemiol Health* **36**, e2014009.
13. Qin Y, Melse-Boonstra A, Yuan B, *et al.* (2012) Zinc biofortification of rice in China: a simulation of zinc intake with different dietary patterns. *Nutrients* **4**, 517–528.
14. Yang Y (2005) *Chinese Food Composition Table 2004*. Beijing: Peking University Medical Press.
15. Huybrechts I, De Backer G, De Bacquer D, *et al.* (2009) Relative validity and reproducibility of a food-frequency questionnaire for estimating food intakes among Flemish preschoolers. *Int J Environ Res Public Health* **6**, 382–399.
16. Willett WC, Howe GR & Kushi LH (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 1220S–1228S discussion 1229S–1231S.
17. Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.
18. Ibiebele TI, Parekh S, Mallitt KA, *et al.* (2009) Reproducibility of food and nutrient intake estimates using a semi-quantitative FFQ in Australian adults. *Public Health Nutr* **12**, 2359–2365.
19. Marques-Vidal P, Ross A, Wynn E, *et al.* (2011) Reproducibility and relative validity of a food-frequency questionnaire for French-speaking Swiss adults. *Food Nutr Res* **55**, 5905–5913.
20. Kelemen LE, Anand SS, Vuksan V, *et al.* (2003) Development and evaluation of cultural food frequency questionnaires for South Asians, Chinese, and Europeans in North America. *J Am Diet Assoc* **103**, 1178–1184.

21. Malekshah AF, Kimiagar M, Saadatian-Elahi M, *et al.* (2006) Validity and reliability of a new food frequency questionnaire compared to 24 h recalls and biochemical measurements: pilot phase of Golestan cohort study of esophageal cancer. *Eur J Clin Nutr* **60**, 971–977.
22. Ocke MC, Bueno-de-Mesquita HB, Goddijn HE, *et al.* (1997) The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol* **26**, Suppl. 1, S37–S48.
23. Pietinen P, Hartman AM, Haapa E, *et al.* (1988) Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. *Am J Epidemiol* **128**, 655–666.
24. Vereecken CA & Maes L (2003) A Belgian study on the reliability and relative validity of the Health Behaviour in School-Aged Children food-frequency questionnaire. *Public Health Nutr* **6**, 581–588.
25. Hakim IA, Hartz V, Harris RB, *et al.* (2001) Reproducibility and relative validity of a questionnaire to assess intake of black tea polyphenols in epidemiological studies. *Cancer Epidemiol Biomarkers Prev* **10**, 667–678.
26. Bogers RP, Van Assema P, Kester AD, *et al.* (2004) Reproducibility, validity, and responsiveness to change of a short questionnaire for measuring fruit and vegetable intake. *Am J Epidemiol* **159**, 900–909.
27. Bolca S, Huybrechts I, Verschraegen M, *et al.* (2009) Validity and reproducibility of a self-administered semi-quantitative food-frequency questionnaire for estimating usual daily fat, fibre, alcohol, caffeine and theobromine intakes among Belgian post-menopausal women. *Int J Environ Res Public Health* **6**, 121–150.
28. Gulliford MC, Mahabir D & Rocke B (2004) Reliability and validity of a short form household food security scale in a Caribbean community. *BMC Public Health* **4**, 22.
29. Block G, Woods M, Potosky A, *et al.* (1990) Validation of a self-administered diet history questionnaire using multiple diet records. *J Clin Epidemiol* **43**, 1327–1335.
30. Hebert JR, Clemow L, Pbert L, *et al.* (1995) Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *Int J Epidemiol* **24**, 389–398.
31. Salvini S, Hunter DJ, Sampson L, *et al.* (1989) Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int J Epidemiol* **18**, 858–867.
32. Haftenberger M, Heuer T, Heidemann C, *et al.* (2010) Relative validation of a food frequency questionnaire for national health and nutrition monitoring. *Nutr J* **9**, 36.
33. Deschamps V, Lauzon-Guillain B, de, Lafay L, *et al.* (2009) Reproducibility and relative validity of a food-frequency questionnaire among French adults and adolescents. *Eur J Clin Nutr* **63**, 282–291.