

ORIGINAL ARTICLE

Reproducibility and relative validity of a food frequency questionnaire developed among reproductive-aged women in Northeast China

Fanghua Liu^{1,2}, Qijun Wu^{1,2,3}, Tingting Gong^{2,4,5,*}¹Department of Epidemiology, School of Public Health, China Medical University, Shenyang 110122, Liaoning Province, China²Key Laboratory of Environmental Stress and Chronic Disease Control and Prevention (China Medical University), Ministry of Education, Shenyang 110122, Liaoning Province, China³NHC Key Laboratory of Advanced Reproductive Medicine and Fertility (China Medical University), National Health Commission, Shenyang 110122, Liaoning Province, China⁴Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang 110044, Liaoning Province, China⁵Department of Obstetrics and Gynecology, The Affiliated Reproductive Hospital of China Medical University, Shenyang 110013, Liaoning Province, China

ABSTRACT

Background: To evaluate the reproducibility and relative validity of a newly developed 118-item food frequency questionnaire (FFQ) for assessing habitual dietary intake among reproductive-aged women in Northeast China. **Methods:** A total of 207 eligible reproductive-aged women were enrolled. After exclusions for incomplete records or implausible energy intakes, 192 completed two FFQs (FFQ1 [first FFQ administration] and FFQ2 [second FFQ administration]) 9 months apart, 182 completed the FFQ and four planned 4-day weighed diet records (WDRs). Reproducibility was assessed using crude and energy-adjusted Spearman correlation coefficients (SCC), intraclass correlation coefficients (ICC), weighted kappa (κ_w), and quartile misclassification. Relative validity was assessed against WDRs using crude, energy-adjusted, and de-attenuated SCC, κ_w , and quartile misclassification. **Results:** For reproducibility, energy-adjusted SCC ranged from 0.32–0.60 (median 0.48) for food groups and 0.30–0.58 (median 0.44) for nutrients; κ_w ranged from 0.24–0.56 (median 0.43) and 0.26–0.53 (median 0.37), respectively. For validity, de-attenuated SCC ranged from 0.32–0.75 (median 0.57) for food groups and 0.29–0.58 (median 0.48) for nutrients. Same-or-adjacent quartile classification ranged from 76.15%–92.25% (median 85.72%) for reproducibility and 66.94%–85.48% (median 78.23%) for validity, with low extreme misclassification (median < 5% for both). **Conclusions:** The FFQ demonstrated acceptable reproducibility and relative validity for ranking the habitual intake of most foods and nutrients among reproductive-aged women in Northeast China, supporting its use in epidemiological studies of diet and reproductive health.

Keywords: food frequency questionnaire, validity, reproducibility

INTRODUCTION

Polyendocrine metabolic ovarian syndrome (PMOS), previously known as polycystic ovary syndrome, is a

common condition affecting reproductive-aged women worldwide.^[1] Evidence suggests that various dietary factors are linked to the development of PMOS, although controversy remains.^[2,3] To determine the rela-

*Corresponding Author:

Tingting Gong, Department of Obstetrics and Gynecology, The Affiliated Reproductive Hospital of China Medical University, Shenyang 110013, Liaoning Province, China. Email: gongtt@sj-hospital.org. <https://orcid.org/0000-0002-3813-8932>.

Received: 13 July 2026; Revised: 19 August 2026; Accepted: 10 September 2026

<https://doi.org/10.54844/prm.2026.1345>

tionship between diet and human diseases, it is important to accurately assess the food and nutrient intakes. Food frequency questionnaire (FFQ) is the most practical and cost-effective approach for evaluating food and nutrient intake, compared to methods such as the 24-hour dietary recall and weighed diet record (WDR).^[4] Consequently, it has been extensively employed in large-scale nutritional epidemiology studies across diverse populations.^[4] In addition to the FFQ for general adults, FFQs for different female populations have been established in China, mainly involving girls, female adolescents, pregnant women, postmenopausal women, and lactating mothers.^[5–9] However, we have not identified a reproducible and validated FFQ suitable for reproductive-aged women living in Northeast China. Thus, we developed a 118-item FFQ to assess habitual dietary intake.

Dietary practices differ considerably among populations with varying regional, ethnic, dietary, or cultural backgrounds. Therefore, for use in a specific population, the FFQ must be customized, reproducible, and properly validated.^[4,10,11] At present, there is no gold standard for validating dietary intake. The basic requirement for validation is that the errors of the reference method are independent of test method. The WDR is widely considered the optimal reference method for evaluating the validity of the FFQ.^[4]

This study examined the reproducibility and relative validity of the FFQ for assessing food and nutrient intakes. Reproducibility was evaluated using a test–retest design, whereas relative validity was assessed using WDRs as the reference method.

METHODS

Ethics statement

This study received approval from the Ethics Committee of Shengjing Hospital of China Medical University (Shenyang, China) (2024PS001F). Written informed consents were obtained from all participants prior to their enrollment.

Study population

The inclusion criteria were current resident in Northeast China for more than five years, > 3 years since menarche and aged not more than 40 years, without a history of any known infectious, neoplastic disease, or specific disorders that present similar signs and symptoms to PMOS, such as hyperprolactinemia, androgen-synthesizing neoplasms, and hypercortisolism, and free-living people without serious illnesses requiring special diets, and not on a weight-loss regimen. A total of 207 eligible participants were enrolled for the study.

Study design

The study was conducted from September 2023 through June 2024. Participants completed only two FFQs: FFQ1 at recruitment (baseline) and FFQ2 approximately nine months later. During this interval, each participant completed four 4-day WDR periods at approximately 3-month intervals. WDR1 was collected within 1 week after FFQ1, and WDR4 was completed before FFQ2. Anthropometric measurements were obtained on the day FFQ1 was completed. Figure 1 illustrates the study design.

FFQ

The FFQ was constructed according to the method described by Willett,^[4] capturing average consumption and intake frequency over the preceding year. Food items were initially selected from the most frequently consumed items reported in the National Health and Dietary Survey in China^[12] and then supplemented with locally common foods identified through a pilot test to reflect regional dietary patterns.

In the end, the FFQ comprises 118 food items. Participants were asked to report their average intake of each food using a six-level scale (ranging from "almost never" to "more than twice per day"). For seasonal fruits and vegetables, participants were asked to indicate how often they consumed these items during the respective season.

WDR as reference

Four-day WDRs served as the reference method and were conducted at approximately 3-month intervals throughout the study. Each 4-day record covered one weekend day and three weekdays. Participants were asked to record the name and quantity of all foods and beverages consumed. Mixed dishes were converted into their original single food, and the weight of each food was calculated based on its proportion in the dish. Detailed instructions were provided on how to weigh and record the food amounts, and participants received analytical food scales with a precision of 0.1 g. After completing each record, participants were asked to check their diet records.

Data cleaning

Participants who did not complete both the FFQ1 and FFQ2 ($n = 9$) were excluded from the analyses. Participants with implausible energy intakes (< 500 or > 4500 kcal/day; $n = 6$) were also removed, following prior practice.^[13] Ultimately, 192 participants were included in the reproducibility study. For the validation study, seven participants failed to complete the WDR, and three had aberrant energy intake values, leaving 182 participants.

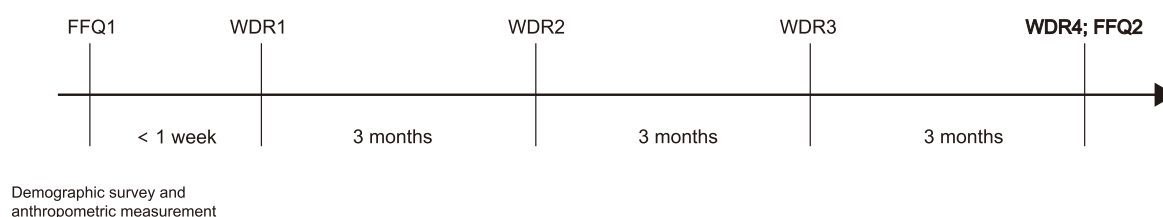


Figure 1. The design of the reproducibility and relative validity study among reproductive-aged women in Northeast China. FFQ1 was administered at baseline and FFQ2 approximately 9 months later. Four 4-day WDR periods were completed at approximately 3-month intervals, with WDR1 beginning within 1 week after FFQ1 and WDR4 completed before FFQ2. The study included two FFQ administrations and four 4-day WDR periods. FFQ1, first FFQ administration; FFQ2, second FFQ administration; WDR, weighed diet record.

Food and nutrient assessment

Daily intake of each food item was calculated by multiplying the average consumption frequency by the portion size. The portion size was determined as the median intake of each food item across all four 4-day WDR periods (16 recording days in total). Nutrient intake per food item was derived by multiplying daily food intake by the nutrient content per 100 grams. Major nutrient compositions were obtained from the Chinese Food Composition Tables.

Statistical analysis

Data were processed and analyzed using Microsoft Excel 16.0 and SAS 9.4. Categorical variables (*e.g.*, body mass index [BMI], educational level, annual income) were presented as frequencies (*n*) and percentages (%). Normally distributed continuous variables (*e.g.*, maternal age, height, weight, BMI) were expressed as mean and standard deviation (SD).

The residual method was employed to exclude the possibility of variation due to energy intake.^[14] Reproducibility between FFQ1 and FFQ2 was evaluated using crude and energy-adjusted Spearman correlation coefficients (SCC), intraclass correlation coefficients (ICC), weighted kappa (κ_w) statistics, and quartile-based misclassification analysis. The ICC was estimated using a two-way mixed-effects model with absolute agreement, based on single measurements rather than average measurements; participants were treated as the random effect and FFQ administration as the fixed effect. Validity of the FFQ relative to the WDR was assessed using crude and energy-adjusted SCC, weighted κ statistics, and quartile misclassification analysis. To account for within-person variation due to day-to-day fluctuations and seasonal effects, we de-attenuated the SCC using within- and between-person variance components from the WDR.^[15] The de-attenuation formula is:

$$r_t = r_0 \sqrt{1 + (S_w^2/S_b^2)/n}$$

where r_t is the true correlation, r_0 is the observed correlation, S_w^2/S_b^2 is the ratio of within- and between-person variances, and n is the number of WDR.

For quartile-based cross-classification, participants were ranked according to their intake and divided into four equal groups. Classifications were considered the same when the two methods placed a participant in the same quartile; adjacent when the quartile assignments differed by one quartile; and extreme (opposite) when a participant was classified into quartile 1 by one method and quartile 4 by the other, or vice versa. The percentage classified into the same or adjacent quartiles was calculated as the sum of the same and adjacent categories.

RESULTS

Characteristics of the participants

The characteristics of the participants are shown in Table 1. For the reproducibility study, the average age was 29.69 years and the average BMI was 26.49 kg/m². According to the BMI standard of Chinese adults,^[16] nine participants (4.68%) were underweight (BMI < 18.5 kg/m²), 44 (22.92%) were normal (18.5 kg/m² ≤ BMI < 24.0 kg/m²), 69 (35.94%) were overweight (24.0 kg/m² ≤ BMI < 28 kg/m²), and 70 (36.46%) were obese (BMI ≥ 28.0 kg/m²). Seven (3.65%) participants received assisted reproductive technology (ART). For the validation study, the mean age of the participants was 29.72 years and the mean BMI was 26.52 kg/m². Seven (3.85%) were underweight (BMI < 18.5 kg/m²), 44 (24.18%) were normal (18.5 kg/m² ≤ BMI < 24.0 kg/m²), 66 (36.26%) were overweight (24.0 kg/m² ≤ BMI < 28 kg/m²), and 65 (35.71%) were obese.

Reproducibility

Table 2 shows reproducibility of the FFQ for the food groups intake. The intakes of only three food groups obtained from FFQ1 were statistically significantly different from the intakes obtained from FFQ2. SCCs

Table 1: Characteristics of participants in the study

Characteristics	Reproducibility study (n = 192)	Validation study (n = 182)
Age at recruitment (years)	29.69 ± 3.99	29.72 ± 4.01
BMI (kg/m ²)	26.49 ± 4.88	26.52 ± 4.87
BMI		
Underweight (n, %)	9 (4.68)	7 (3.85)
Normal weight (n, %)	44 (22.92)	44 (24.18)
Overweight (n, %)	69 (35.94)	66 (36.26)
Obesity (n, %)	70 (36.46)	65 (35.71)
Education level		
Middle school or below	67 (34.90)	63 (34.62)
College or higher	125 (65.10)	119 (65.38)
Annual income (RMB)		
< 100,000	169 (88.02)	160 (87.91)
≥ 100,000	23 (11.98)	22 (12.09)
Current smokers	10 (5.21)	8 (4.40)
Current drinkers	15 (7.81)	14 (7.69)
Received ART	7 (3.65)	7 (3.85)

Data are presented as mean ± standard deviation for continuous variables and n (%) for categorical variables. ART, assisted reproductive technology; BMI, body mass index.

Table 2: Reproducibility of FFQ for main food items and food groups

Main food items and food groups (g)	Intake value			Correlation coefficient				κ _w
	FFQ1	FFQ2	P	Crude SCC	Adjusted SCC [#]	Crude ICC	Adjusted ICC [#]	
Cereals and cereal products	415.01 (251.67-527.41)	412.50 (245.41-520.26)	0.72	0.64*	0.51*	0.65*	0.52*	0.43
Snacks/desserts	13.74 (6.37-22.78)	15.21 (7.16-25.47)	0.12	0.55*	0.44*	0.58*	0.46*	0.43
Milk and dairy	85.74 (40.37-159.48)	92.46 (45.24-175.63)	0.04	0.62*	0.60*	0.64*	0.51*	0.40
Red meat	17.16 (7.31-42.84)	18.50 (7.98-45.22)	0.35	0.41*	0.33*	0.51*	0.35*	0.27
Poultry	4.02 (4.02-17.16)	4.02 (4.02-17.16)	0.99	0.65*	0.51*	0.80*	0.63*	0.56
Processed meat	36.33 (21.18-66.03)	34.84 (20.24-63.51)	0.41	0.58*	0.46*	0.60*	0.47*	0.43
Fish and seafood	9.38 (3.35-17.42)	10.24 (3.80-18.92)	0.26	0.58*	0.49*	0.60*	0.50*	0.45
Eggs	17.65 (6.70-35.36)	16.51 (6.26-33.43)	0.48	0.57*	0.32*	0.54*	0.42*	0.36
Pickled foods	8.31 (4.46-15.30)	8.80 (4.65-16.21)	0.56	0.63*	0.50*	0.66*	0.52*	0.52
Tubers	20.02 (12.73-53.25)	21.44 (14.58-58.66)	0.21	0.52*	0.41*	0.50*	0.40*	0.34
Fresh vegetables	104.98 (64.24-180.57)	88.50 (52.40-155.24)	< 0.01	0.56*	0.45*	0.53*	0.42*	0.48
Legumes and legume products	41.00 (21.61-80.26)	40.28 (21.10-78.52)	0.68	0.54*	0.43*	0.58*	0.46*	0.43
Fresh fruits	113.67 (53.05-224.88)	90.20 (41.58-180.69)	< 0.01	0.44*	0.35*	0.31*	0.24*	0.25
Nuts	2.98 (1.01-4.40)	3.15 (1.10-4.65)	0.62	0.51*	0.37*	0.42*	0.33*	0.24
Tea	33.50 (0-133.50)	33.50 (0-133.50)	0.99	0.75*	0.58*	0.72*	0.56*	0.51
Sugar drink	38.86 (16.75-105.86)	41.20 (18.50-112.46)	0.38	0.72*	0.58*	0.75*	0.64*	0.49
Coffee	0 (0-16.75)	0 (0-16.75)	0.99	0.74*	0.59*	0.85*	0.62*	0.50
Alcohol	0 (0-16.75)	0 (0-16.75)	0.99	0.68*	0.54*	0.78*	0.61*	0.42

Intake value are presented as Median (P₂₅, P₇₅). FFQ1, first FFQ administration; FFQ2, second FFQ administration; ICC, intraclass correlation coefficient; SCC, spearman correlation coefficient; κ_w, weighted kappa. P values were derived from the Wilcoxon signed-rank test; #, Energy-adjusted intakes by the residual method; *, P < 0.05.

between FFQ1 and FFQ2 for the main food items and group intake were above 0.50 except in red meat (SCC = 0.41) and fresh fruits (SCC = 0.44). The crude and energy-adjusted SCCs between FFQ1 and FFQ2 ranged from 0.41 in red meat to 0.75 in tea (median = 0.58) and from 0.32 in eggs to 0.60 in milk and dairy (median = 0.48), respectively. The ICCs for the main food items and group intake were above 0.50 except for fresh fruits (ICC = 0.31) and nuts (ICC = 0.42). The crude and energy-adjusted ICC ranged from 0.31 in fresh fruits to 0.85 in coffee (median = 0.60) and from 0.24 in fresh fruits to 0.64 in sugar drink (median = 0.49), respectively. Weighted κ values ranged from 0.24 in nuts to 0.56 in poultry (median = 0.43).

As shown in Table 3, we found that six out of 29 intakes of energy and nutrients obtained from FFQ1 were statistically significantly different from the intakes obtained from FFQ2. We found that SCC for energy and all nutrients were above 0.40, except for vitamin E (SCC = 0.39). The crude and energy-adjusted SCC between FFQ1 and FFQ2 for reproducibility ranged from 0.39 in vitamin E to 0.71 in zinc (median = 0.56) and from 0.30 in vitamin A to 0.58 in zinc (median = 0.44), respectively. The crude and energy-adjusted ICC ranged from 0.21 in polyunsaturated fatty acid (PUFA) to 0.68 in thiamine (median = 0.55) and from 0.15 in carotene to 0.53 in potassium (median = 0.42), respectively. The weighted κ values ranged from 0.26 for calcium to 0.53 for niacin (median = 0.37), indicating fair or moderate agreement.

The degree of misclassification associated with the categorized intakes that were assessed using the FFQs was examined as the proportion of participants who were classified into same, adjacent, and extreme quartile (Supplementary Table 1). The percentage of participants classified into the same or adjacent quartiles and the extreme quartiles by FFQ1 and FFQ2 ranged from 79.85% to 92.25% (median = 85.72%) and from 0 to 11.63% (median = 1.55%), respectively. After energy adjustment, the same or adjacent quartile percentages ranged from 75.97% to 93.80% (median = 85.27%), and extreme quartile percentages ranged from 0 to 8.53% (median = 2.71%).

When the nutrient intake was categorized into quartiles (Supplementary Table 2), the agreement rates for the same or adjacent quartile classifications ranged from 76.15% monounsaturated fatty acids (MUFA) to 91.47% (vitamin A). After energy adjustment, the range was from 73.64% (protein) to 90.84% (insoluble fiber). The crude and energy-adjusted percentages classified into opposite quartiles ranged from 0 (retinol, vitamin A, vitamin E, saturated fatty acids (SFA), iron, manganese, selenium, insoluble fiber, and energy) to 3.85% (PUFA)

and from 0 (copper) to 7.75% (protein and riboflavin), respectively.

Validity

As shown in Table 4, the crude and energy-adjusted SCC between the FFQ and WDR ranged from 0.28 in fresh fruit to 0.72 in tea (median = 0.55) and from 0.21 in fresh fruits to 0.58 in cereals and cereal products (median = 0.45), respectively. De-attenuation slightly improved SCC for most variables. The de-attenuation values ranged from 0.32 to 0.75 (median = 0.57). Agreement between the FFQ and WDR was fair or moderate ($\kappa_w > 0.20$) for the intake of the main food groups, except for fresh fruit ($\kappa_w = 0.17$).

As shown in Supplementary Table 3, the percentage of participants classified into the same or adjacent quartile ranged from 66.94% (fresh fruits) to 85.48% (tea), while the percentage classified into extreme quartiles ranged from 1.61% (tea) to 9.68% (fresh fruits). After energy adjustment, the ranges of agreement rates for the same or adjacent and extreme quartile classifications were 66.13%-83.06% (median = 75.41%) and 2.42%-9.68% (median = 5.65%), respectively, when the main food items and groups were derived from the FFQ and WDR. Weighted κ statistics ranged from 0.18 (nuts) to 0.42 (tea).

The results of the validity of the FFQ for energy and nutrient intake are presented in Table 5. We found that 15 out of the 29 selected nutrients were significantly different between the FFQ and WDR. The crude SCC for nutrients and energy between the FFQ and the WDR ranged from 0.26 for vitamin C to 0.57 for iron (median = 0.48); the energy-adjusted SCC ranged from 0.26 for cholesterol to 0.55 for phosphorus (median = 0.38); and the de-attenuated coefficients ranged from 0.29 for soluble fiber to 0.58 for iron (median = 0.48). Weighted κ values ranged from 0.14 to 0.37 (median = 0.28), and all values were above 0.20 except for fiber ($\kappa_w = 0.15$), insoluble fiber ($\kappa_w = 0.14$), cholesterol ($\kappa_w = 0.17$), and zinc ($\kappa_w = 0.18$).

For the classification into quartiles (Supplementary Table 4), the proportion of participants who were classified into the same or adjacent quartiles by the FFQ and WDR ranged from 68.55% for insoluble fiber to 83.06% for iron (median = 78.23%). Extreme quartile of participants was < 10.00% for all energy and nutrients.

DISCUSSION

This study evaluated the reproducibility and relative validity of a new 118-item FFQ developed for reproductive-aged women in Northeast China, using repeated FFQ administrations and WDR. Overall, the FFQ

Table 3: Reproducibility of FFQ for energy and nutrients intake

Nutrients	Intake value		P	Correlation coefficient				κ_w
	FFQ1	FFQ2		Crude SCC	Adjusted SCC [#]	Crude ICC	Adjusted ICC [#]	
Energy (kcal)	1086.05 (872.40-1600.33)	1115.49 (890.27-1640.36)	0.45	0.58*	-	0.55*	-	0.43
Protein (g)	44.01 (34.47-63.63)	43.15 (33.51-62.86)	0.28	0.55*	0.44*	0.54*	0.42*	0.27
Fat (g)	28.44 (19.76-38.78)	29.80 (21.23-40.57)	0.09	0.49*	0.39*	0.52*	0.40*	0.35
Carbohydrate (g)	186.98 (142.14-257.75)	184.22 (140.50-252.46)	0.52	0.54*	0.43*	0.59*	0.46*	0.36
Fiber (g)	10.73 (6.41-16.27)	9.20 (5.57-13.83)	< 0.01	0.62*	0.50*	0.56*	0.43*	0.49
Soluble fiber (g)	3.64 (2.02-5.76)	3.75 (2.12-5.90)	0.72	0.48*	0.38*	0.45*	0.34*	0.33
Insoluble fiber (g)	6.95 (4.23-10.40)	6.37 (3.85-9.51)	0.02	0.57*	0.45*	0.53*	0.41*	0.48
Cholesterol (mg)	75.67 (50.87-113.00)	77.28 (51.51-115.80)	0.61	0.52*	0.41*	0.51*	0.36*	0.37
SFA (g)	6.52 (4.10-9.09)	6.68 (4.15-9.25)	0.48	0.50*	0.39*	0.54*	0.41*	0.40
MUFA (g)	6.77 (4.29-9.15)	6.53 (4.10-8.99)	0.34	0.45*	0.35*	0.41*	0.31*	0.31
PUFA (g)	3.33 (2.09-5.06)	3.42 (2.15-5.15)	0.55	0.43*	0.34*	0.21*	0.28*	0.28
Vitamin A (μg)	532.53 (243.20-919.50)	485.61 (220.55-860.48)	0.02	0.40*	0.30*	0.35*	0.27*	0.31
Carotene (μg)	1731.84 (1094.33-2426.70)	1610.21 (950.47-2200.50)	< 0.01	0.42*	0.33*	0.38*	0.15*	0.37
Retinol (μg)	71.91 (37.07-123.91)	75.44 (39.56-130.24)	0.14	0.58*	0.47*	0.56*	0.43*	0.36
Thiamine (mg)	0.48 (0.36-0.76)	0.47 (0.35-0.75)	0.68	0.51*	0.40*	0.68*	0.37*	0.29
Riboflavin (mg)	0.79 (0.58-1.37)	0.82 (0.60-1.42)	0.35	0.60*	0.48*	0.57*	0.44*	0.36
Niacin (mg)	12.40 (8.85-16.43)	12.55 (8.95-16.60)	0.71	0.64*	0.51*	0.62*	0.48*	0.53
Folate (μg)	116.43 (73.40-184.63)	125.80 (80.52-195.48)	0.04	0.53*	0.42*	0.50*	0.38*	0.36
Vitamin C (mg)	91.03 (52.85-167.24)	75.27 (40.56-135.82)	<0.01	0.44*	0.35*	0.42*	0.32*	0.31
Vitamin E (mg)	7.68 (4.81-13.22)	7.85 (5.03-13.60)	0.58	0.39*	0.37*	0.46*	0.35*	0.31
Calcium (mg)	370.40 (257.99-649.43)	368.55 (255.44-645.28)	0.88	0.56*	0.44*	0.58*	0.45*	0.26
Phosphorus (mg)	733.29 (563.38-1028.53)	750.41 (580.26-1055.63)	0.22	0.61*	0.53*	0.59*	0.46*	0.42
Potassium (mg)	1546.23 (1051.98-2483.80)	1520.60 (1030.58-2450.41)	0.41	0.63*	0.51*	0.60*	0.53*	0.41
Magnesium (mg)	172.51 (128.63-243.10)	175.82 (131.53-248.65)	0.36	0.65*	0.52*	0.64*	0.50*	0.48
Iron (mg)	16.16 (11.66-24.25)	15.85 (11.43-23.97)	0.42	0.68*	0.55*	0.66*	0.51*	0.42
Zinc (mg)	6.36 (4.68-8.40)	6.48 (4.75-8.50)	0.59	0.71*	0.58*	0.63*	0.49*	0.42
Selenium (μg)	18.47 (13.10-28.47)	18.24 (12.89-28.18)	0.65	0.67*	0.52*	0.65*	0.50*	0.41
Copper (mg)	1.31 (0.92-2.19)	1.38 (0.98-2.30)	0.18	0.69*	0.56*	0.61*	0.47*	0.43
Manganese (mg)	3.06 (2.15-4.11)	3.18 (2.24-4.28)	0.76	0.67*	0.54*	0.35*	0.19*	0.40

Intake value are presented as Median (P_{25} , P_{75}). FFQ1, first FFQ administration; FFQ2, second FFQ administration; ICC, intraclass correlation coefficient; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acid; SCC, spearman correlation coefficient; SFA, saturated fatty acids; κ_w , weighted kappa. P values were derived from the Wilcoxon signed-rank test; #, Energy-adjusted intakes by the residual method; *, $P < 0.05$.

showed acceptable reproducibility and relative validity for the intake of most food groups and nutrients. These findings support its use in epidemiological studies of diet and reproductive health in this population, although estimates for several specific foods and nutrients should be interpreted cautiously.

The reproducibility results indicated that the FFQ provided reasonably stable rankings of habitual dietary intake over a 9-month interval. The median crude SCCs were 0.58 for food groups and 0.56 for nutrients, while the corresponding median ICC were 0.60 and 0.55. Approximately 85% of participants were classified into

the same or adjacent intake quartile by the two FFQ administrations. These results are comparable to those reported in the Northeast Cohort Study of China, in which the median reproducibility SCC was 0.53 for nutrients and 0.59 for food groups after energy adjustment.^[17] Together, these findings indicate that the present FFQ has acceptable reproducibility for ranking dietary intake. Energy adjustment reduced the correlation coefficients for many foods and nutrients. This pattern was also observed in the Northeast Cohort Study of China^[17] and may reflect increased correlated measurement error after adjustment for total energy intake. Lower reproducibility was observed for fresh fruit, nuts,

Table 4: Relative validity of the FFQ for main food items and food groups (median values and percentiles)

Main food items and food groups (g)	Intake value		P	SCC			κ_w
	FFQ	WDR		Crude	Adjusted [#]	De-attenuated	
Cereals and cereal products	412.37 (247.14-526.82)	388.54 (241.26-498.17)	0.07	0.61*	0.58*	0.63*	0.38
Snacks/desserts	14.26 (6.34-23.82)	11.18 (4.72-18.45)	0.03	0.55*	0.43*	0.57*	0.32
Milk and dairy	90.14 (41.82-167.53)	71.62 (31.45-142.36)	< 0.01	0.59*	0.51*	0.60*	0.41
Red meat	17.82 (7.43-43.16)	14.67 (5.32-35.84)	0.03	0.47*	0.36*	0.48*	0.23
Poultry	4.15 (4.02-17.43)	5.34 (1.12-20.26)	0.14	0.53*	0.44*	0.53*	0.35
Processed meat	35.61 (20.37-65.14)	29.82 (15.46-57.18)	0.01	0.51*	0.39*	0.52*	0.34
Fish and seafood	10.04 (3.52-18.76)	8.31 (2.24-15.82)	0.08	0.57*	0.46*	0.58*	0.37
Eggs	17.26 (6.42-33.85)	14.15 (4.37-29.64)	0.06	0.49*	0.37*	0.50*	0.29
Pickled foods	8.52 (4.48-15.63)	10.34 (5.14-18.72)	0.11	0.63*	0.52*	0.63*	0.44
Tubers	21.14 (12.86-56.24)	18.27 (9.14-47.35)	0.18	0.44*	0.34*	0.46*	0.26
Fresh vegetables	96.84 (56.42-165.17)	81.35 (45.28-146.72)	< 0.01	0.41*	0.32*	0.43*	0.24
Legumes and legume products	41.25 (21.08-80.46)	34.72 (16.82-69.53)	0.04	0.55*	0.45*	0.56*	0.36
Fresh fruits	101.46 (45.37-203.18)	75.82 (30.46-157.24)	< 0.01	0.28*	0.21*	0.32*	0.17
Nuts	3.12 (1.08-4.62)	2.37 (0.64-3.92)	0.24	0.35*	0.27*	0.35*	0.21
Tea	33.64 (5.12-134.26)	27.85 (2.34-116.48)	0.48	0.72*	0.54*	0.75*	0.43
Sugar drink	40.82 (17.15-110.37)	31.54 (9.82-91.64)	0.02	0.67*	0.53*	0.68*	0.40
Coffee	0 (0-33.56)	4.18 (0.45-15.72)	0.03	0.69*	0.55*	0.70*	0.42
Alcohol	0 (0-8.38)	3.39 (0.37-18.64)	0.03	0.65*	0.50*	0.66*	0.39

Intake value are presented as Median (P25, P75). FFQ, food frequency questionnaire; WDR, weighed diet record; κ_w , weighted kappa. P values were derived from the Wilcoxon signed-rank test; #, Energy-adjusted intakes by the residual method; *, $P < 0.05$.

eggs, and vitamin E. Seasonal variation, infrequent consumption, and difficulty in estimating usual portion sizes may partly explain these findings.

Relative validity against the WDRs was also acceptable for most dietary variables. The median crude SCCs were 0.55 for food groups and 0.48 for nutrients, and the corresponding deattenuated medians were 0.57 and 0.48. The proportions of participants classified into the same or adjacent quartile ranged from 67% to 85% for food groups and from 69% to 83% for nutrients, whereas extreme misclassification was below 10%. These findings suggest that the FFQ is generally capable of distinguishing participants with relatively low or high dietary intake, which is an important requirement in studies examining diet-disease associations.

Significant differences in absolute intake were found between the FFQ and WDR for 15 of the 29 nutrients examined. Such differences are expected because the two methods assess dietary intake over different time frames and have different sources of measurement error. The FFQ estimates habitual intake over an extended period and depends on participants' memory and their ability to estimate usual portion sizes, whereas the WDR

captures foods consumed on specific recording days and may be affected by day-to-day variation or changes in eating behavior during recording. The relatively weak correlations for fresh fruit, nuts, fiber, cholesterol, vitamin C, and zinc may reflect seasonal or infrequent consumption, substantial within-person variation, and difficulties in estimating portion sizes or food composition. These estimates should therefore be interpreted with caution.

The estimates are consistent with validation studies conducted in other Chinese populations. In the Taizhou adult FFQ, reproducibility and validity coefficients were generally lower (0.25-0.61 and 0.20-0.58, respectively), possibly because the reference method comprised four 24-hour recalls rather than weighed records and the questionnaire covered a broader age range.^[11] A rural Chinese FFQ with a one-month retest interval yielded higher reproducibility (0.58-0.92), probably because shorter intervals capture less true dietary change.^[18] Our results also fall within the range reported for Chinese female adolescents (validity 0.41-0.65) and lactating mothers (validity 0.23-0.72),^[5,9] supporting the generalizability of FFQ-based ranking in diverse Chinese female populations.

Table 5: Relative validity of the FFQ for energy and nutrients intake (median values and percentiles)

Nutrients	Intake value		<i>P</i>	SCC			κ_w
	FFQ	WDR		Crude	Adjusted [#]	De-attenuated	
Energy (kcal)	1105.42 (885.40-1625.50)	1082.50 (850.43-1580.65)	0.52	0.53*	-	0.53*	0.34
Protein (g)	43.82 (33.85-63.05)	45.20 (35.10-65.82)	< 0.01	0.41*	0.39*	0.42*	0.28
Fat (g)	29.37 (20.55-39.80)	31.40 (22.51-42.18)	< 0.001	0.48*	0.34*	0.48*	0.34
Carbohydrate (g)	184.85 (140.20-254.63)	178.64 (135.27-245.40)	< 0.01	0.49*	0.37*	0.49*	0.28
Fiber (g)	9.80 (5.80-14.99)	8.51 (5.23-12.80)	< 0.01	0.30*	0.42*	0.30*	0.15
Soluble fiber (g)	3.75 (2.12-5.97)	3.28 (1.81-5.15)	< 0.01	0.29*	0.37*	0.29*	0.25
Insoluble fiber (g)	6.55 (3.94-9.89)	6.10 (3.65-9.23)	0.26	0.27*	0.41*	0.31*	0.14
Cholesterol (mg)	76.92 (51.40-115.23)	82.49 (55.67-125.85)	< 0.001	0.35*	0.26*	0.36*	0.17
SFA (g)	6.64 (4.11-9.26)	6.95 (4.30-9.83)	0.82	0.45*	0.41*	0.45*	0.20
MUFA (g)	6.58 (4.15-9.00)	6.80 (4.25-9.45)	0.08	0.48*	0.31*	0.50*	0.28
PUFA (g)	3.42 (2.19-5.26)	3.55 (2.36-5.63)	0.25	0.49*	0.34*	0.49*	0.33
Vitamin A (μg)	512.45 (235.46-895.40)	495.24 (225.46-850.61)	0.94	0.37*	0.34*	0.37*	0.24
Carotene (μg)	1685.28 (1030.57-2350.49)	1540.61 (980.52-2180.46)	0.15	0.36*	0.37*	0.36*	0.24
Retinol (μg)	74.15 (38.50-128.43)	65.85 (32.44-110.58)	< 0.001	0.45*	0.31*	0.45*	0.24
Thiamine (mg)	0.47 (0.35-0.75)	0.49 (0.36-0.78)	< 0.01	0.48*	0.33*	0.48*	0.32
Riboflavin (mg)	0.82 (0.60-1.42)	0.85 (0.62-1.48)	< 0.01	0.51*	0.42*	0.51*	0.32
Niacin (mg)	12.52 (8.95-16.60)	13.49 (9.57-17.83)	0.03	0.55*	0.53*	0.55*	0.32
Folate (μg)	122.31 (77.55-191.24)	118.52 (72.43-185.68)	0.55	0.41*	0.38*	0.41*	0.26
Vitamin C (mg)	82.45 (45.87-150.41)	75.20 (40.51-135.87)	< 0.01	0.26*	0.37*	0.37*	0.26
Vitamin E (mg)	7.82 (4.95-13.50)	7.63 (4.88-13.11)	0.35	0.51*	0.53*	0.51*	0.23
Calcium (mg)	371.25 (258.41-650.28)	385.46 (265.26-670.48)	0.32	0.34*	0.29*	0.34*	0.33
Phosphorus (mg)	745.53 (575.20-1045.68)	768.21 (595.44-1080.50)	0.04	0.49*	0.55*	0.50*	0.33
Potassium (mg)	1538.61 (1045.59-2475.22)	1510.47 (1010.51-2420.61)	0.45	0.44*	0.38*	0.44*	0.32
Magnesium (mg)	175.47 (131.26-247.55)	170.50 (125.47-238.25)	0.38	0.48*	0.42*	0.48*	0.32
Iron (mg)	16.12 (11.65-24.15)	15.80 (11.46-23.57)	0.95	0.57*	0.39*	0.58*	0.37
Zinc (mg)	6.45 (4.75-8.58)	6.82 (5.13-9.29)	< 0.01	0.52*	0.48*	0.52*	0.18
Selenium (μg)	18.22 (12.87-28.16)	22.49 (15.26-35.67)	< 0.001	0.32*	0.35*	0.33*	0.34
Copper (mg)	1.37 (0.96-2.23)	1.42 (1.05-2.35)	0.38	0.54*	0.54*	0.54*	0.30
Manganese (mg)	3.10 (2.20-4.18)	3.23 (2.30-4.45)	< 0.001	0.50*	0.48*	0.51*	0.34

Intake value are presented as Median (P25, P75). FFQ, food frequency questionnaire; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acid; SCC, spearman correlation coefficients; SFA, saturated fatty acids; WDR, weighed diet record; κ_w , weighted kappa. *P* values were derived from the Wilcoxon signed-rank test; #, Energy-adjusted intakes by the residual method; *, *P* < 0.05.

Comparisons with international FFQs further support this interpretation. The PERSIAN Cohort FFQ showed food-group validity correlations of 0.23-0.79 and reproducibility correlations of 0.42-0.72,^[19] while the Tehran Lipid and Glucose Study reported mean validity coefficients of 0.39-0.53 and reliability coefficients of 0.59-0.60.^[20] Corrected correlations between FFQ and dietary records in Flemish preschoolers ranged from 0.32 to 0.75.^[21] Our ranges are similar to these benchmarks, even though the instruments, item numbers, reference periods, and study populations differed.

An important strength of this study was the use of repeated FFQs and WDRs. The evaluation also included both food groups and nutrients, as well as correlation and intake-classification analyses. In addition, the questionnaire was developed for reproductive-aged women in Northeast China and therefore considered the dietary

characteristics of the target population.

Several limitations should be acknowledged. First, participants were recruited from a single research center. The findings may therefore not be fully generalizable to all reproductive-aged women. Second, the 9-month interval between FFQ administrations may reflect both measurement error and true dietary change. Third, we did not include dietary supplement intake in the nutrient intake calculation. This omission may lead to an underestimation of nutrient intake, particularly among frequent supplement users.

CONCLUSION

The 118-item FFQ demonstrated acceptable reproducibility and relative validity for ranking the habitual intake

of most foods and nutrients among reproductive-aged women in Northeast China. However, estimates for fresh fruit, nuts, fiber, cholesterol, vitamin C, zinc, and other variables with relatively weak correlations should be interpreted cautiously. Further validation in larger populations is warranted.

DECLARATIONS

Supplementary information

Supplementary materials are only available at the official site of the journal (<https://www.hksmp.com/journals/prm>).

Acknowledgments

None.

Author contributions

Liu FH, Gong TT: Conceptualization, Design, Manuscript drafting, Critical revision. Liu FH, Wu QJ, Gong TT: Data collection, Data cleaning, Discrepancy checking, Data analysis. All authors interpreted the data, read the manuscript, and approved the final version.

Source of funding

None.

Ethical approval

This study was approved by the Ethics Committee of Shengjing Hospital of China Medical University (Shenyang, China) (2024PS001F).

Informed consent

Written informed consents were obtained from all participants before enrollment.

Use of large language models, AI and machine learning tools

None declared.

Data availability statement

Processed datasets used in the analysis are available from the corresponding author upon reasonable request.

Conflict of interest

Wu QJ serves as a member of the Editorial Board of *Placenta and Reproductive Medicine*. To avoid any potential conflict of interest, Wu QJ had no involvement in the editorial handling or decision-making process for this manuscript. The other authors declare no conflicts of interest.

REFERENCES

1. Lizneva D, Suturina L, Walker W, Brakta S, Gavrilova-Jordan L, Azziz R. Criteria, prevalence, and phenotypes of polycystic ovary syndrome.

- Fertil Steril*. 2016;106(1):6-15.
2. Darand M, Sadeghi N, Salimi Z, *et al*. Is the MIND diet useful for polycystic ovary syndrome? A case-control study. *BMC Womens Health*. 2024;24(1):282.
3. Vaziri Esfarjani F, Dorfeshan P, Mansouri A, *et al*. Association between dietary inflammatory index, empirical dietary inflammatory patterns, dietary and lifestyle inflammation scores, and polycystic ovary syndrome: a case-control study. *BMC Endocr Disord*. 2025;25(1):202.
4. Willett W C. *Nutritional Epidemiology*. Oxford University Press; 2013: 71.
5. Xia W, Sun C, Zhang L, *et al*. Reproducibility and relative validity of a food frequency questionnaire developed for female adolescents in Suihua, North China. *PLoS One*. 2011;6(5):e19656.
6. Li M, Halldorsson TI, Bjerregaard AA, *et al*. Relative validity and reproducibility of a food frequency questionnaire used in pregnant women from a rural area of China. *Acta Obstet Gynecol Scand*. 2014;93(11):1141-1149.
7. Shu XO, Yang G, Jin F, *et al*. Validity and reproducibility of the food frequency questionnaire used in the Shanghai Women's Health Study. *Eur J Clin Nutr*. 2004;58(1):17-23.
8. Xu L, J Dibley M, D'Este C. Reliability and validity of a food-frequency questionnaire for Chinese postmenopausal women. *Public Health Nutr*. 2004;7(1):91-98.
9. Ding Y, Li F, Hu P, *et al*. Reproducibility and relative validity of a semi-quantitative food frequency questionnaire for the Chinese lactating mothers. *Nutr J*. 2021;20(1):20.
10. Fraser GE. A search for truth in dietary epidemiology. *Am J Clin Nutr*. 2003;78(3 Suppl):521S-525S.
11. Zhuang M, Yuan Z, Lin L, *et al*. Reproducibility and relative validity of a food frequency questionnaire developed for adults in Taizhou, China. *PLoS One*. 2012;7(11):e48341.
12. Li LM, Rao KQ, Kong LZ, *et al*. [A description on the Chinese national nutrition and health survey in 2002]. *Zhonghua Lin Xing Bing Xue Za Zhi*. 2005;26(7):478-484.
13. Wang Q, Sun Y, Xu Q, *et al*. Higher dietary inflammation potential and certain dietary patterns are associated with polycystic ovary syndrome risk in China: A case-control study. *Nutr Res*. 2022;100:1-18.
14. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr*. 1997;65(4 Suppl):1220S-1228S; discussion 1229S-1231S. [PMID 9094926 DOI: 10.1093/ajcn/65.4.1220S]
15. Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am J Epidemiol*. 1988;127(2):377-386.
16. Wang YF, Sun MX, Xue H, *et al*. [Understanding the China Blue Paper on Obesity Prevention and Control and policy implications and recommendations for obesity prevention and control in China]. *Chinese Journal of Preventive Medicine*. 2019;53(9):875-884.
17. Cui Q, Xia Y, Liu Y, *et al*. Validity and reproducibility of a FFQ for assessing dietary intake among residents of northeast China: northeast cohort study of China. *Br J Nutr*. 2023;129(7):1252-1265.
18. Liu X, Li H, Zhao Y, *et al*. Reproducibility and relative validity of a food frequency questionnaire for a diet-related study in a rural Chinese population. *Nutr J*. 2022;21(1):3.
19. Eghtesad S, Hekmatdoost A, Faramarzi E, *et al*. Validity and reproducibility of a food frequency questionnaire assessing food group intake in the PERSIAN Cohort Study. *Front Nutr*. 2023;10:1059870.
20. Mirmiran P, Esfahani FH, Mehrabi Y, Hedayati M, Azizi F. Reliability and relative validity of an FFQ for nutrients in the Tehran lipid and glucose study. *Public Health Nutr*. 2010;13(5):654-662.
21. Huybrechts I, De Backer G, De Bacquer D, Maes L, De Henauw S. Relative validity and reproducibility of a food-frequency questionnaire for estimating food intakes among Flemish preschoolers. *IJERPH*. 2009;6(1):382-399.