

Reproducibility, relative validity and calibration of a food frequency questionnaire for adults

Reprodutibilidade, validade relativa e calibração de um questionário de frequência alimentar elaborado para adultos

Reproducibilidad, validez relativa y calibración de un cuestionario de frecuencia alimentaria para adultos

Neuciani Ferreira da Silva ¹
 Rosely Sichieri ²
 Rosângela Alves Pereira ³
 Regina Maria Veras Gonçalves da Silva ⁴
 Márcia Gonçalves Ferreira ⁴

Abstract

The purpose of the present study was to evaluate the reproducibility and relative validity and calibrate the dietary intake assessment of a food frequency questionnaire (FFQ) using a random sample of 195 adults aged 20 to 50 years from the Central-West Region of Brazil. The reference method used by the study was two 24-hour recalls (24hR) that provided energy-adjusted deattenuated food intake data for comparison purposes. With respect to reproducibility, the average weighted kappa was 0.43 and exact agreement was 41.5%. With regard to relative validity, correlation coefficients ranged from 0.32 (thiamin) to 0.51 (carbohydrates), with a mean of 0.41. Deattenuation and adjustment for energy intake decreased most correlation coefficients in relation to crude values. The food frequency questionnaire showed good reliability and moderate validity for most nutrients based on classification into quartiles of energy and nutrient intake. The calibrated means of the FFQ were more similar to the means estimated from the 24hR and showed lower standard deviation.

Reproducibility of Results; Questionnaires; Adult

Resumo

O objetivo do estudo foi avaliar a reprodutibilidade e a validade relativa de um questionário de frequência alimentar (QFA) e estimar seus fatores de calibração. Uma amostra aleatória de 195 adultos, com idades entre 20 e 50 anos, foi estudada no Centro-oeste do Brasil. O método de referência utilizado foi o recordatório de 24 horas (R24h). Quanto à reprodutibilidade, o kappa médio ponderado foi de 0,43 e a concordância exata foi de 41,5%. Com relação à validade relativa, os coeficientes de correlação variaram de 0,32 (tiamina) a 0,51 (carboidratos), com média de 0,41. A deatenuação associada ao ajuste para a energia reduziu a maioria dos coeficientes de correlação em relação aos valores brutos obtidos para os nutrientes. O QFA avaliado mostrou boa reprodutibilidade e validade moderada para a maioria dos nutrientes, quando avaliada pela categorização em quartis. As médias do QFA calibrado ficaram mais próximas das obtidas pelo R24h, com desvios-padrão menores.

Reprodutibilidade dos Testes; Questionários; Adulto

¹ Instituto de Saúde Coletiva, Universidade Federal de Mato Grosso, Cuiabá, Brasil.

² Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil.

³ Instituto de Nutrição Josué de Castro, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

⁴ Faculdade de Nutrição, Universidade Federal de Mato Grosso, Cuiabá, Brasil.

Correspondence

N. F. Silva
 Departamento de Saúde Coletiva, Instituto de Saúde Coletiva, Universidade Federal de Mato Grosso.
 Rua 217, Quadra 44, nº 5, Setor 2, Cuiabá, MT
 78088-225, Brasil.
 neuciani@yahoo.com.br

Introduction

Measuring the exposure to dietary factors in epidemiological studies is a complex task because eating is influenced by the interaction of several factors, including an individual's biological characteristics, behavioral and affective issues, culture and socioeconomic status. These factors complicate not only the reporting of food consumption but also data interpretation. The growth in prevalence of diet-related and nutrition-related diseases, such as obesity and non-transmissible chronic diseases has stimulated the development of more appropriate dietary assessment tools^{1,2}.

Food frequency questionnaires (FFQ) have proven to be an essential tool for epidemiological studies, especially cohort and case-control studies, investigating the role of diet in the etiology, prevention and treatment of diseases, and for the evaluation of interventions aimed at promoting healthy dietary habits³. FFQs are widely used to estimate dietary intake in prospective studies because they allow researchers to estimate usual intake, are relatively inexpensive, and data analysis is practical and simple^{4,5}.

However, the elaboration of food lists used in the questionnaires is time consuming and the tool is also subject to measurement errors that often lead to biased estimates of the association between diet and disease development. Therefore, the evaluation of reproducibility, validation and calibration of FFQs is a critical element of effective dietary intake assessment⁶.

Evaluation of reproducibility is an estimate of the similarity between the results of the same FFQ conducted on two separate occasions, assuming no significant changes in dietary habits between the two time points⁷. Validity and calibration are assessed by comparing the results of the FFQ with the results of another dietary assessment method which serves as a reference method to quantify measurement error⁸. Validation assesses the accuracy of dietary intake estimates⁹, while calibration identifies correction factors to improve the accuracy of estimated food and nutrient intake obtained from FFQ data^{10,11}.

In Brazil, FFQs have been developed for the populations of the following cities: Rio de Janeiro (two questionnaires)^{12,13}, São Paulo (three questionnaires)^{14,15,16}, Goiânia, Goiás State (one questionnaire)¹⁷, Brasília (one questionnaire)¹⁸ and Porto Alegre, Rio Grande do Sul State (two questionnaires)^{19,20}. However, the majority of these FFQs have not been validated through population-based studies and deattenuated coefficients have not been calculated to correct the data for intrapersonal variability⁹.

This study uses internationally recommended statistical techniques to evaluate reproducibility and validity and calibrate the dietary intake assessment of a FFQ designed for adults from the Central-West Region of Brazil²¹. This tool will help in the study of diet-related diseases, especially obesity, metabolic syndrome and cardiovascular diseases, which have increased markedly in this region²².

Methods

Study design

Two 24-hour dietary recalls were conducted at an interval of 30 days. The 24-hour dietary recall (24hR) was chosen as the reference method since it is the most suitable method for the assessment of individual food intake in large samples²³. Reproducibility of the FFQ was evaluated by conducting the questionnaire on two separate occasions (FFQ1 and FFQ2), also at an interval of 30 days.

Sample

A total of 115 men and 115 women were randomly selected from a sub-sample of 686 adults aged between 20 and 50 years from a population-based survey conducted to estimate the prevalence of arterial hypertension²⁴. This sample size was chosen based on the findings of Cade et al.⁹ that established that a sample size of 100 individuals is sufficient for the analysis of FFQ validity.

Data collection

Data were collected between July and September 2007 through in-home interviews conducted by two trained nutritionists. Home visits were made on randomly selected weekdays and weekend days and if the individual was not located by the third visit the interview was conducted with the right-hand side next door neighbor of the original interviewee.

During the first visit, a socioeconomic questionnaire was filled out prior to completing the FFQ and 24hR. A second visit to conduct the second FFQ and 24hR was scheduled for a future date at least 30 days after the first visit. As recommended by Willett⁷, the FFQ was conducted before the reference method on both occasions.

The FFQ

The food list used in the FFQ was elaborated based on a dietary survey (24hR) of a sample of

104 individuals aged between 20 and 50 years²⁵. A total of 81 food items, comprising those items reported by at least 15% of sample members, were included in the list²⁵. The questionnaire included the following frequency of consumption options: more than three times a day; twice or three times a day; once a day; five or six times a week; twice to four times a week; once a week; once to three times a month; and never or almost never. Intake of 63 of the 81 food items was based on three portion size options. For the remaining items, including onions, peppers, butter and cream cheese, only information on intake frequency was requested. Reference portions were defined based on the most reported portion sizes in the 24hR. Standard portion sizes or units were adopted for certain foods, such as bread rolls or eggs. The time frame for the FFQ was the last six months preceding the interview²⁵.

The 24hR

Individuals were asked to describe the food consumed the day before the interview, including the type of food consumed, portion size, mode of preparation and time and location of consumption. Interview techniques were used to help interviewees recall food intake through associating consumption with the activities performed during the day. Pictures of food items and utensils²⁶ were also used to help individuals describe the amount consumed. Individuals were probed to remember foods that are often forgotten in 24hR, such as beverages, sauces, spreads, snacks and sugar²⁷.

Data treatment

A detailed review of the questionnaires was performed to screen out implausible energy intake reports (outliers with values below the second percentile or above the 98th percentile in both recalls, $n = 13$). Food consumption frequency reported in both FFQs was converted into daily equivalent frequency of consumption by attributing proportional values to the frequency of consumption options calculated with reference to a base value of 1.0 for the "once a day" option. For example, the value 2.5 was assigned to the option "twice to three times a day" based on the following daily equivalent frequency of consumption calculation: $[(2 + 3)/2] = 2.5$, while the value 0.79 was assigned to the "five to six times a week" option calculated as average frequency $[(5 + 6)/2]$ divided by seven (the number of days in a week).

Serving sizes were presented in their respective units of measurement based on standard amounts determined by previous studies²⁸

and measurements taken in the laboratory. Daily intake was calculated by multiplying the daily equivalent frequency of consumption by the respective amount per serving in grams or milliliters.

The NutWin software (Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, Brazil) was used to estimate daily energy and nutrient intake (carbohydrate; protein; total saturated and unsaturated fats; cholesterol; fiber; calcium; iron; vitamin C; thiamin; and folate) based on data from both the FFQs and 24hR. The NutWin software food composition table is based on data obtained from the United States Department of Agriculture (USDA). The nutritional composition of foods not on the NutWin list was taken from the Brazilian Food Composition Table²⁹. Nutrient intake was energy-adjusted using the residual method³⁰.

Mean usual energy and nutrient intake based on data from the 24hR was estimated using the PC-SIDE Program which uses the Iowa State University method^{31,32} to correct data for intraindividual variability (Department of Statistics, Iowa State University, Ames, USA). The means were deattenuated considering the ratio of within and between-subject variance.

Data analysis

Reproducibility of the FFQ was assessed using the weighted kappa statistic to evaluate the degree of agreement for classification into quartiles of energy and nutrient intake based on data from the two FFQs. Exact agreement (the proportion of individuals categorized into the same quartiles), adjacent agreement (the proportion of individuals categorized into adjacent quartiles) and disagreement (the proportion of individuals categorized into opposite and distant quartiles) were calculated. Furthermore, crude and energy-adjusted intraclass correlation coefficients (ICCs) were calculated to assess the linear relationship between energy and nutrient intake based on data from the FFQ1 and FFQ2 and degree of agreement. We also calculated the ICCs for portions and food items.

The relative validity of the FFQ was assessed by comparing the estimates obtained using data from the FFQ2 with the crude and deattenuated means of the two 24hR. The Student's t-test was used to ascertain differences between mean energy and nutrient intake based on data from the FFQ2 and crude mean energy and nutrient intake based on data from the two 24hR. Pearson's correlation coefficients were also calculated for energy and nutrient intake based on data from the FFQ2 and crude and deattenuated and energy-adjusted estimates from the 24hR. Finally, the study

also evaluated the degree of agreement between the two different methods based on the classification of individuals into quartiles of energy and nutrient intake using the weighted kappa. Exact agreement and disagreement were evaluated by comparing the proportion of individuals classified into same and opposite quartiles.

Calibration was performed using a linear regression model. Energy and nutrient intake values based on mean deattenuated energy and nutrient intake data from the 24hR were used as dependent variables and estimates based on data from the FFQ2 were used as independent variables. Values for each nutrient were calibrated using the following equation:

$$\text{Calibrated value} = \alpha + \lambda Q \quad (\text{equation 1})$$

where α is the regression constant, λ is the slope of the regression line, and Q is the estimated energy and nutrient intake from the FFQ.

The significance level was set at $p = 0.05$ for all analyses. The Kolmogorov-Smirnov test was used to test the distribution symmetry of continuous variables. Where necessary, the variables with asymmetric distributions were converted into logarithms to obtain symmetric distributions to fit the model.

This study was conducted in accordance with the Declaration of Helsinki guidelines and all relevant procedures were approved by the Ethics Committee of the Júlio Muller University Hospital (application number # 234/CEP-HU-JM/05). Participants signed an informed consent form after being informed of the goals of the study.

Results

The FFQs and 24hR were completed by 208 subjects (108 women and 100 men). The main reasons for loss were change of address and inability to attend the second interview. Thirteen individuals were excluded after data screening due to inconsistencies in dietary information. The final sample therefore consisted of 195 individuals (100 women and 95 men), corresponding to 85% of the original sample. The mean time interval between conducting the first and second FFQ was 36 days (standard deviation = six days). Mean age was 34 years (standard deviation = 10 years), with a predominance of individuals in the 20 to 29 year age range (42%). The majority of individuals (52%) had completed high school and 68% were living with a per capita household income of less than the Brazilian monthly minimum wage at the time of the study (R\$380 equivalent to US\$203.92). Compared to the two

24hR the FFQ1 and FFQ2 overestimated energy and nutrient intake (Table 1).

ICCs for food items ranged from 0.25 (sweet cookies 95%CI: 0.11;0.37) to 0.85 (beer 95%CI: 0.80;0.88) and were over 0.60 in 34.9% of the 63 items included in the questionnaire. The ICC for food portion sizes ranged between 0.40 and 0.60 in 57.1% of the items. ICCs for nutrients ranged from 0.49 (iron 95%CI: 0.38;0.59) to 0.65 (carbohydrates 95%CI: 0.56;0.72) (data not shown).

The weighted kappa of agreement of classification into quartiles of energy and nutrient intake between the FFQ1 and FFQ2 ranged from 0.33 (protein) to 0.60 (energy) ($p < 0.01$). Mean exact agreement was 41.5%, while mean disagreement was 4.5% (Table 2).

Pearson's correlation coefficients for energy and nutrient intake from the FFQ2 and crude mean energy and nutrient intake from the 24hR varied between 0.32 (folate) and 0.51 (carbohydrates). When correlated with the energy-adjusted and deattenuated 24hR mean energy and nutrient intake values ranged from 0.12 (unsaturated fat) to 0.41 (calcium) (Table 3).

Coefficients for the crude values for most nutrients dropped after adjustment for energy intake and increased in the case of calcium ($r = 0.40$ to $r = 0.41$), vitamin C ($r = 0.34$ to $r = 0.39$), thiamin ($r = 0.32$ to $r = 0.33$) and folate ($r = 0.35$ to $r = 0.36$). The strongest correlation after deattenuation and adjustment for energy found was for calcium, while the weakest correlation was for unsaturated fat ($r = 0.12$).

The weighted kappa of agreement of classification into quartiles of energy and nutrient intake between the FFQ2 and 24hR varied from 0.06 (unsaturated fat) to 0.53 (energy). Agreement was highest for the macronutrient fiber ($k = 0.38$). The results also revealed a reasonable degree of agreement for the micronutrients calcium and vitamin C ($k = 0.36$). Exact agreement was highest for energy (45.1%) and lowest for unsaturated fat (26.2%). Disagreement was highest for unsaturated fat (12.8%) (Table 4).

Table 5 presents the results of calibration, with values ranging from -0.75 (iron) to 0.46 (vitamin C). The results of calibration for energy, carbohydrates, protein, fat, unsaturated fat, cholesterol, fiber, thiamin and folate were statistically significant. As expected, the mean calibrated values based on data from the FFQ were similar to the energy-adjusted mean values from the 24hR, and were associated with a considerable reduction in standard deviation.

Table 1

Mean energy intake, adjusted mean * and 95% confidence interval (95%CI) of nutrients included in the food frequency questionnaire (FFQ 2) and means of the two 24-hour recalls (24hR). Adults from the Central-west Region of Brazil, 2007 (N = 195).

Variable	FFQ1		FFQ2		24hR	
	Mean	95%CI	Mean	95%CI	Mean	95%CI
Energy (Kcal)	3,008.0	2,837;3,178.2	2,642.0	2,498.8;2,785.3	1,623.6	1,537.3;1,709.9
Carbohydrate (g)	395.6	387.2;404.0	345.9	339.2;352.6	189.6	184.4;194.8
Protein (g)	108.4	105.3;111.6	92.9	90.3;95.5	66.6	64.3;68.9
Lipid (g)	88.1	86.0;90.3	75.7	73.9;77.5	50.2	48.5;51.9
Saturated fat (g)	16.1	15.4;16.8	25.9	24.7;27.1	15.7	15.1;16.5
Unsaturated fat (g)	29.9	28.7;31.2	35.3	34.3;36.3	27.8	26.7;28.7
Cholesterol (mg)	242.6	224.4;260.9	320.8	303.3;338.3	209.4	194.6;224.2
Fiber (g)	22.5	20.6;24.3	23.6	22.1;25.1	10.0	9.5;10.7
Calcium (mg)	597.8	563.2;632.4	678.2	634.6;721.8	295.1	272.6;317.6
Iron (mg)	14.7	14.2;15.1	14.0	13.5;14.5	9.0	8.7;9.3
Vitamin C (mg)	211.3	188.6;234.0	172.7	157.9;187.5	87.4	64.4;110.4
Thiamin (mg)	2.5	2.2;2.7	2.6	2.3;2.9	1.1	1.0;1.2
Folate (mcg)	326.4	308.0;344.8	211.8	202.3;221.3	110.3	102.3;118.5

* Energy adjusted using the residual method.

Table 2

Weighted kappa of energy adjusted intake and agreement (%) between the two food frequency questionnaires (FFQ1/FFQ2). Adults from the Central-west Region of Brazil, 2007 (N = 195).

Energy/adjusted nutrients	Weighted kappa *	Exact agreement (%)	Exact and adjacent agreement (%)	Disagreement (%)
Energy (Kcal)	0.60	47.7	87.7	2.1
Carbohydrate (g)	0.42	38.5	77.9	3.6
Protein (g)	0.33	35.4	75.9	5.6
Lipid (g)	0.36	36.4	77.9	6.2
Saturated Fat (g)	0.34	37.9	75.9	5.6
Unsaturated Fat (g)	0.40	45.1	76.9	5.1
Cholesterol (mg)	0.47	43.6	80.5	4.1
Fiber (g)	0.38	37.9	76.9	4.6
Calcium (mg)	0.53	44.6	84.6	3.1
Iron (mg)	0.35	39.0	76.9	5.6
Vitamin C (mg)	0.51	47.7	83.6	3.6
Thiamin (mg)	0.41	41.0	82.6	6.7
Folate (mcg)	0.53	47.7	83.1	3.1
Mean	0.43	41.5	79.7	4.5

* $p < 0.001$.

Table 3

Pearson's correlation coefficients for energy and nutrient intake from the food frequency questionnaire (FFQ 2) and means of the two 24-hour recall (24hR). Adults from the Central-west Region of Brazil, 2007 (N = 195).

Variables *	Raw **	Deattenuated *** adjusted #
Energy (Kcal)	0.49	-
Carbohydrate (g)	0.51	0.29 *
Protein (g)	0.43	0.27 *
Lipid (g)	0.40	0.19 *
Saturated Fat (g)	0.35	0.16 *
Unsaturated Fat (g)	0.43	0.12
Cholesterol (mg)	0.39	0.29 *
Fiber (g)	0.46	0.37 *
Calcium (mg)	0.40	0.41 *
Iron (mg)	0.42	0.24 *
Vitamin C (mg)	0.34	0.39 *
Thiamin (mg)	0.32	0.33 *
Folate (mcg)	0.35	0.36 *
Mean	0.41	0.29

* Values converted into logarithms;

** $p < 0.05$;

*** Corrected using the variance ratio (intra-individual and inter-individual variability);

Energy adjusted using the residual method.

Discussion

The FFQ assessed by this study showed good reproducibility and moderate relative validity. The results of calibration to correct estimates indicate that this FFQ is an appropriate tool for assessing food intake in adults from the Central-west Region of Brazil.

The analysis of reproducibility reveals that the FFQ is generally reliable considering that more than 75% of the participants were classified into same or adjacent quartiles of energy and nutrient intake. ICCs observed by this study were similar to those found by Salvo & Gimeno¹⁶ and Kesse-Guyot³³, but slightly lower than values reported by Zanolla et al.¹⁹ and Marques-Vidal et al.³⁴.

The average time interval between conducting the two FFQs (30 days) is considered adequate because it is not likely to influence eating habits and individuals do not have enough time to become familiar with the questionnaire⁹.

The evaluation of relative validity was based on correlation coefficients and agreement for classification into quartiles of energy and nutri-

ent intake. Correlation coefficients are widely used by studies evaluating dietary assessment methods. Cade et al.⁹ reported that 83% of studies assessing the validity of FFQs used this statistical measure in their analyses. In a review of FFQ validation studies, Thompson & Byers³⁵ observed that correlation coefficients ranged from 0.4 to 0.7. The mean correlation coefficient for crude estimates of energy and nutrient intake presented by the present study is within an acceptable level (mean $r = 0.41$), and is similar to mean values found by Jackson et al.³⁶ ($r = 0.44$), Kusama et al.³⁷ ($r = 0.46$), Marques-Vidal et al.³⁴ ($r = 0.42$), Block et al.³⁸ and Deschamps et al.³⁹ ($r = 0.52$). FFQ validation studies carried out in Brazil also showed similar values: Fornés et al.¹⁷ and Ribeiro et al.¹⁸ found a mean Pearson correlation coefficient of 0.50, and Henn et al.²⁰ observed a mean value of 0.49.

For most nutrients, correlation coefficients decreased after energy-adjustment and deattenuation, remaining below 0.30. However, the correlation coefficients for calcium, vitamin C, thiamin and folate remained constant or increased after adjustment. Similarly low deattenuated correlation coefficients were reported by Fornés et al.¹⁷ in an assessment of a FFQ for construction workers in the Central-west region of Brazil.

The correlation coefficients observed for fiber, calcium, vitamin C, thiamin and folate were higher than those reported by Chen et al.⁴⁰, who evaluated the validity of a FFQ developed for adults in Bangladesh, observing correlation coefficients below 0.3 for these and other nutrients, such as iron and vitamin B6.

According to Freedman et al.⁴¹, the low-magnitude of correlation may be a result of several factors, including: biased reporting, where people with high food intake tend to underestimate consumption; reference method bias; variations in food intake during the study period; and difficulties in recalling food intake and correctly estimating food portions.

Decreases in correlation coefficients after deattenuation and adjustment for energy intake were also observed in other studies in Brazil^{12,17,19,20}. Adjustment for total energy intake is a common procedure⁷. In the present study this process led to the elimination of large variations for most nutrients, except calcium, showing that this may not be a necessary step in analysis in validation studies where a strong correlation is observed with total food intake.

Our review of research conducted in Brazil, including studies carried out by Salvo & Gimeno¹⁶, Fornés et al.¹⁷, and Lima et al.⁴², shows that adjustment for energy intake led to lower correlation coefficients for most items,

Table 4

Weighted kappa of energy adjusted intake and agreement (%) between the food frequency questionnaire (FFQ 2) and the means of the two 24-hour recalls (24hR) based on classification into quartiles of energy and nutrient intake. Adults from the Central-west Region of Brazil, 2007 (N = 195).

Variables *	Weighted kappa	Exact agreement (%)	Exact and adjacent agreement (%)	Disagreement (%)
Energy (Kcal)	0.53 **	45.1	85.1	3.1
Carbohydrate (g)	0.19 **	34.9	69.7	9.2
Protein (g)	0.30 **	30.3	73.3	5.1
Lipid (g)	0.13	27.2	66.7	8.7
Saturated Fat (g)	0.13	33.8	68.2	10.8
Unsaturated Fat (g)	0.06	26.2	67.7	12.8
Cholesterol (mg)	0.33 **	36.4	72.3	4.1
Fiber (g)	0.38 **	34.9	75.4	3.1
Calcium (mg)	0.36 **	35.4	75.9	4.6
Iron (mg)	0.30 **	38.5	74.4	7.2
Vitamin C (mg)	0.36 **	39.0	73.8	4.1
Thiamin (mg)	0.13	32.3	68.7	10.8
Folate (mcg)	0.27 **	31.3	74.9	7.7
Mean	0.27	34.3	72.8	7.0

* Energy adjusted logarithms;

** $p < 0.05$.

Table 5

Calibration parameters (α and λ), 95% confidence interval (95%CI) and standard deviation (SD) for energy and nutrient intake from the food frequency questionnaire (FFQ 2), two 24-hour recalls (24hR) and calibrated FFQ. Adults from the Central-west Region of Brazil, 2007 (N = 195).

Variables *	α (95%CI)	λ ** (95%CI)	FFQ2 (SD)	24hR (SD)	Calibrated FFQ
Energy (Kcal) ***	940.2 (657.5;1,222.9)	0.37 (0.27;0.47) #	2,715.1 (1,026.2)	1,967.6 (847.6)	1,944.8 (379.7)
Carbohydrate (g)	114.6 (66.6;162.5)	0.29 (0.16;0.42) #	361.6 (52.5)	217.2 (51.2)	219.5 (15.2)
Protein (g)	46.3 (30.9;61.7)	0.38 (0.22;0.53) #	97.1 (19.8)	83.3 (22.2)	83.2 (7.5)
Lipid (g)	46.5 (34.5;58.4)	0.22 (0.07;0.37) #	78.4 (13.9)	64.5 (15.4)	63.8 (3.1)
Saturated Fat (g)	16.5 (12.5;20.4)	0.16 (-0.11;0.42)	14.3 (4.5)	18.8 (8.7)	18.8 (0.7)
Unsaturated Fat (g)	23.9 (19.2;28.5)	0.20 (0.03;0.36) #	27.0 (7.9)	29.2 (9.5)	29.3 (1.6)
Cholesterol (mg)	132.9 (94.6;171.2)	0.37 (0.21;0.53) #	214.8 (97.7)	210.4 (118.2)	212.4 (36.2)
Fiber (g)	11.0 (8.7;13.3)	0.21 (0.11;0.31) #	20.5 (10.5)	14.9 (7.6)	15.3 (2.2)
Calcium (mg)	366.9 (277.3;456.5)	0.08 (-0.08;0.25)	519.9 (200.5)	410.1 (219.1)	408.5 (16.0)
Iron (mg)	40.4 (-38.7;119.6)	-0.75 (-6.53;5.04)	13.4 (3.2)	29.2 (124.3)	30.4 (2.4)
Vitamin C (mg)	70.8 (-79.4;221.1)	0.46 (-0.22;1.15)	190.9 (123.0)	156.4 (535.7)	158.6 (56.6)
Thiamin (mg)	0.7 (0.5;1.0)	0.27 (0.20;0.34) #	2.3 (2.2)	1.4 (1.2)	1.3 (0.6)
Folate (mcg)	65.8 (42.5;89.1)	0.10 (0.03;0.17) #	297.8 (111.4)	96.3 (60.9)	95.6 (11.1)

* Energy-adjusted nutrient intake from the two food recalls (dependent variable) and food frequency questionnaire (independent variable);

** Calibration parameter;

*** Mean crude energy intake from the two food recalls (dependent variable) and food frequency questionnaire (independent variable);

$p < 0.05$.

while deattenuation minimally altered crude correlation coefficients. Based on the findings of other studies^{16,38,43,44}, a greater increase in correlation coefficients following deattenuation could have been expected for most nutrients.

The use of only two 24hR may be insufficient for estimating usual energy and nutrient intake, thus limiting the relevance of the correlation coefficients. Higher correlation coefficients have been observed in studies that used a greater number of reference method replications. For example, Kroke et al.⁴⁵, in a study using twelve 24hR, observed coefficients ranging from 0.54 (fiber) to 0.84 (energy), while Johansson et al.⁴⁶, in a study using 10 replications of the 24hR, found values between 0.44 (protein and cholesterol) and 0.63 (carbohydrate). A study of a sample of adults in Brazil carried out by Crispin et al.⁴⁷ using four 24hR found low correlations for only two nutrients (calcium = 0.13 and lipids = 0.39), even in individuals with a low level of education. The values for the other nutrients assessed by the study were above 0.40 and ranged from 0.40 (protein) to 0.98 (retinol) among individuals with a higher level of education.

According to Cade et al.⁹, an increase in the time interval between conducting dietary recalls is associated with improvements in the validity of the FFQ. Furthermore, Molag et al.⁴⁸ observed that correlation coefficients of studies that used a time interval of between eight and 14 days were significantly higher than in studies that used a time interval of between one to seven days. Pereira et al.⁴⁹, in an analysis of dietary intake in adults, observed a high degree of within and between-subject variance, with values of up to 15.6 for cholesterol in men. These authors showed that the number of replications needed to estimate usual energy and nutrient intake depends on the nutrient being analyzed and on the expected correlation coefficient: for example, to achieve an expected correlation coefficient of 0.7 requires an average of eight replications within a range of one (for carbohydrates) to 15 (for cholesterol).

The classification of individuals into quantile categories (tertiles, quartiles, or quintiles) reveals the capacity of both methods to allocate individuals according to level of nutrient intake and is useful for estimating disease risk or relative protection measures¹⁹. Masson et al.⁵⁰ recommends that no more than 10% of individuals should be classified into opposite quartiles and at least 50% should be classified into the same quartile in epidemiological dietary intake studies. In the present study, average exact agreement for energy and nutrient intake was 34%, and disagreement was below 10% for most of the items analyzed,

except saturated and unsaturated fats and thiamin. It should be noted that the sum of exact and adjacent agreement was more than 70% for most nutrients, indicating a good degree of concordance and emphasizing that energy adjustment is not always necessary for assessing validation studies. The rate of exact agreement was identical to that found by Henn et al.²⁰. In the same study the rate of disagreement was slightly lower (6.1%) than in the present study. Similarly, in a study by Zanolla et al.¹⁹, exact agreement ranged from 35% (vitamins A and C) to 40% (energy), and average disagreement was 4%, and in a study carried out by Pakseresht & Sharma⁵¹ exact agreement ranged from 30% (protein) to 45% (energy) and average disagreement was 4.5%.

Masson et al.⁵⁰ suggested a minimum kappa value for reasonable agreement of 0.40, whereas Crewson⁵² noted that kappa values between 0.21 and 0.40 indicate reasonable agreement. The kappa values observed by this study were reliable for energy (0.53) and reasonable for protein (0.30), cholesterol (0.33), fiber (0.38), calcium (0.36), iron (0.30), vitamin C (0.36) and folate (0.26). Reasonable values were also reported by Zanolla et al.¹⁹ and Pakseresht & Sharma⁵² for most nutrients.

The moderate validity of the FFQ may partially explained by the reference method (24hR) used by this study. It is widely recognized that, in comparison to the double-labeled water method that assesses energy expenditure, the 24hR underestimates energy intake⁵³. Underreporting in the 24hR may also explain why the FFQ means of energy and nutrient intake were significantly higher than the 24hR means. A review of the literature found that the typical level of underreporting among adults is approximately 20% and that higher levels are common among obese individuals⁵⁴. According to Buzzard⁵⁵, underreporting of consumption is highest among obese individuals, women, teenagers and the elderly. Furthermore, underreporting seems to be selective and foods with more calories, such as cakes, soft drinks, cheese, sandwiches, butter, sauces and sugar, are generally poorly reported^{43,55}. Behavior related to underreporting of food intake is complex and includes perceptual, emotional, cognitive, social, moral and physical factors. Several aspects of underreporting remain largely unexplored, significantly affecting food intake assessment⁴³.

With respect to calibration using linear regression, intercept and slope (λ) values of approximately zero and 1.0, respectively, indicate absence of bias in the questionnaire, i.e., mean FFQ intake is similar to the mean intake estimated using the reference method⁵⁶. As observed in other calibration studies^{12,57}, slope values under 1.0 were observed in the present study. In this

respect, it should be noted that, even under optimal study conditions, food consumption assessment is subject to several sources of error that affect data quality^{7,58}, including misreporting, within-subject variability, use of a limited number of foods items or excessive aggregation of items during grouping of food items, which may induce subjects to under or over-report intake. Furthermore, the attenuation coefficient slope, independence between the errors of both methods of dietary intake assessment and a lack of systematic errors in the reference method may be attributed to the violation of theoretical assumptions essential to the calibration process⁷.

It was observed that standard deviations were lower for calibrated data than for crude data, corroborating findings of other studies^{11,56,57}. Calibration allows researchers to correct errors that occur in the classification of individuals since extreme values are affected by linear correction due to the linearity between FFQ and the reference method values⁵⁹.

For both the FFQ and 24hR, effective data collection depends on the ability of individuals to recall and describe their dietary intake. Certain sources of error are common to participants, resulting in positive correlations between measurement errors. Thus, variance in intake levels may be somewhat overestimated when recalls are used as a reference method^{57,60}.

This study represents a valuable contribution to nutritional epidemiology research in this region. Further studies are necessary to help refine the present FFQ and revise the food list and reference portions. Continuous evaluation of dietary assessment methods is essential considering current changes in food intake and the importance of FFQ for monitoring food habits and detecting groups at risk due to dietary intake and nutritional deficiencies.

Resumen

El objetivo del estudio fue evaluar la reproducibilidad y la validez relativa de un cuestionario de frecuencia alimentaria (CFA) y estimar sus factores de calibración. Fue estudiada una muestra aleatoria de 195 adultos del centro-oeste de Brasil, con edades entre 20 y 50 años. El método de referencia utilizado fue el recordatorio de 24 horas (R24h). En cuanto a la reproducibilidad, el kappa medio ponderado fue 0,43 y la concordancia exacta fue de un 41,5%. En relación con la validez relativa, los coeficientes de correlación variaron de 0,32 (tiamina), a 0,51 (carbohidratos), con una media de 0,41. Se observó una disminución de los coeficientes de correlación obtenidos, en cuanto a los nutrientes sin atenuación y ajustados por el consumo de energía, cuando se comparan con los coeficientes obtenidos para los valores brutos. El CFA evaluado demostró buena reproducibilidad y una validez moderada para la mayoría de los nutrientes cuando fue evaluado por la categorización por cuartiles. Las medias del CFA calibradas quedaron más cerca a las medias obtenidas por los R24h, con menores desviaciones estándar.

Reproducibilidad de Resultados; Cuestionarios; Adulto

Contributors

N. F. Silva participated in data collection, analysis and interpretation and in the drafting of this manuscript. R. A. Pereira, R. M. V. G. Silva and R. Sichieri participated in project conception and design, data analysis and interpretation and in revising this manuscript. M. G. Ferreira was the principal investigator, was responsible for project conception and design, research team training, coordination of data collection, treatment, analysis and interpretation, and revision of the final text.

Acknowledgments

We are grateful to CNPq (grant number # 4021022005/3) for funding this project. The authors are also grateful to Marina Campos Araujo for her support with data analysis.

References

1. Fisberg RM, Martini LA, Slater B. Métodos de inquéritos alimentares. In: Fisberg RM, Slater B, Marchioni DML, Martini LA, organizadores. Inquéritos alimentares: métodos e bases científicos. São Paulo: Edições Manole; 2005. p. 1-31.
2. Pereira RA, Sichieri R. Métodos de avaliação do consumo alimentar. In: Kac G, Sichieri R, Gigante DP, organizadores. Epidemiologia nutricional. Rio de Janeiro: Editora Fiocruz/Editora Atheneu; 2007. p. 181-200.
3. Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of energy and macronutrient intake of a food frequency questionnaire developed for the German Part of the EPIC Project. *Int J Epidemiol* 1997; 26(1 Suppl):S71-81.
4. McPherson RS, Hoelscher DM, Alexander M, Scanlon KS, Serdula MK. Dietary assessment methods among school-aged children: validity and reliability. *Prev Med* 2000; 31:S11-33.
5. Furlan-Viebig R, Pastor-Valero M. Desenvolvimento de um questionário de frequência alimentar para o estudo da dieta e doenças não transmissíveis. *Rev Saúde Pública* 2004; 38:581-4.
6. Subar AF. Developing dietary assessment tools. *J Am Diet Assoc* 2004; 104:769-70.
7. Willett W. Nutritional epidemiology. New York: Oxford University Press; 1998.
8. Carithers TC, Talegkar SA, Rowser ML, Henry OR, Dubbert PM, Bogle ML, et al. Validity and calibration of food frequency questionnaires used with African American adults in the Jackson Heart Study. *J Am Diet Assoc* 2009; 109:1184-93.
9. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilization of food-frequency questionnaires: a review. *Public Health Nutr* 2002; 5:567-87.
10. Thompson FE, Midthune D, Subar AF, Kipnis V, Kahle LL, Schatzkin A. Development and evaluation of a short instrument to estimate usual dietary intake of percentage energy from fat. *J Am Diet Assoc* 2007; 107:760-7.
11. Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, et al. Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutr* 2002; 5:915-23.
12. Araujo MC, Yokoo EM, Pereira RA. Validation and calibration of a semi-quantitative food frequency questionnaire designed for adolescents. *J Am Diet Assoc* 2010; 110:1170-7.
13. Sichieri R, Everhart JE. Validity of a Brazilian food frequency questionnaire against dietary recalls and estimated energy intake. *Nutr Res* 1998; 18:1649-59.
14. Ishihara J, Iwasaki M, Kunieda CM, Hamada GS, Tsugane S. Food frequency questionnaire is a valid tool in the nutritional assessment of Brazilian women of diverse ethnicity. *Asia Pac J Clin Nutr* 2009; 18:76-80.
15. Pereira GA, Genaro PS, Santos LC, Sarkis KS, Pinheiro MM, Szjenfeld VL, et al. Validation of a food frequency questionnaire for women with osteoporosis. *J Nutr Health Aging* 2009; 13:403-7.
16. Salvo VL, Gimeno SG. Reproducibility and validity of a food frequency questionnaire. *Rev Saúde Pública* 2002; 36:505-12.
17. Fornés NS, Stringhini ML, Elias BM. Reproducibility and validity of a food-frequency questionnaire for use among low-income Brazilian workers. *Public Health Nutr* 2003; 6:821-7.
18. Ribeiro AB, Sávio KEO, Rodrigues MLCF, Costa THM, Schmitz BAS. Validação de um questionário de frequência de consumo alimentar para população adulta. *Rev Nutr* 2006; 19:553-62.
19. Zanolla AF, Olinto MT, Henn RL, Wahrlich V, Anjos LA. Assessment of reproducibility and validity of a food frequency questionnaire in a sample of adults living in Porto Alegre, Rio Grande do Sul State, Brazil. *Cad Saúde Pública* 2009; 25:840-8.
20. Henn RL, Fuchs SC, Moreira LB, Fuchs FD. Development and validation of a food frequency questionnaire (FFQ-Porto Alegre) for adolescent, adult and elderly populations from Southern Brazil. *Cad Saúde Pública* 2010; 26:2068-79.
21. Millen B. Dietary behavior assessment: historical and recent innovations. *J Am Diet Assoc* 2010; 110:1166-9.
22. Schmidt MI, Duncan BB, Silva GA, Menezes AM, Monteiro CA, Barreto SM, et al. Chronic non-communicable diseases in Brazil: burden and current challenges. *Lancet* 2011; 377:1949-61.
23. Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J Am Diet Assoc* 2006; 106:1640-50.
24. Barbosa LS, Scala LCN, Ferreira MG. Associação entre marcadores antropométricos de adiposidade corporal e hipertensão arterial na população adulta de Cuiabá, Mato Grosso. *Rev Bras Epidemiol* 2009; 12:237-47.
25. Ferreira MG, Silva NF, Schmidt FD, Silva RMVG, Sichieri R, Guimarães LV, et al. Development of a Food Frequency Questionnaire for adults in a population-based sample in Cuiabá, Mid-Western Region of Brazil. *Rev Bras Epidemiol* 2010; 13:1-11.
26. Zabotto CB, Vianna RPT, Gil MF. Registro fotográfico para inquéritos dietéticos: utensílios e porções. Campinas: Núcleo de Estudos e Pesquisas em Alimentação, Universidade Estadual de Campinas; 1996.
27. Kac G, Sichieri R, Gigante DP, organizadores. Epidemiologia nutricional. Rio de Janeiro: Editora Fiocruz/Editora Atheneu; 2007.
28. Pinheiro ABV, Lacerda EMA, Benzecry EH, Gomes MCS, Costa VM. Tabela para avaliação de consumo alimentar em medidas caseiras. Rio de Janeiro: Editora Atheneu; 2005.
29. Núcleo de Estudos e Pesquisas em Alimentação, Universidade Estadual de Campinas. TACO: Tabela Brasileira de Composição de Alimentos. 2ª Ed. Campinas: Núcleo de Estudos e Pesquisas em Alimentação, Universidade Estadual de Campinas; 2006.
30. Willett W, Stampfer MJ. Total energy intake: implications for epidemiological analyses. *Am J Epidemiol* 1986; 12:17-27.

31. Nusser SM, Carriquiry AL, Dodd KW, Fuller WA. A semi parametric transformation approach to estimating usual intake distributions. *J Am Stat Assoc* 1996; 91:1440-9.
32. Nusser SM, Fuller WA, Guenther PM. Estimating usual dietary intake distributions: adjusting for measurement error and no normality in 24-hour food intake data. In: Lyberg L, Biemer P, Collins M, De Leeuw E, Dippo C, Schwarz N, et al., editors. *Survey measurement and process quality*. New York: Wiley and Sons; 1997. p. 689-709.
33. Kesse-Guyot E, Castetbon K, Touvier M, Hercberg S, Galan P. Relative validity and reproducibility of a food frequency questionnaire designed for French adults. *Ann Nutr Metab* 2010; 57:153-62.
34. Marques-Vidal P, Ross A, Wynn E, Rezzi S, Paccaud F, Decarli B. Reproducibility and relative validity of a food-frequency questionnaire for French-speaking Swiss adults. *Food Nutr Res* 2011; 55:103402.
35. Thompson FE, Byers T. Dietary assessment resource manual. *J Nutr* 1994; 124(11 Suppl):2245S-317S.
36. Jackson M, Walker S, Cade J, Forrester T, Cruickshank JK, Wilks R. Reproducibility and validity of a quantitative food-frequency questionnaire among Jamaicans of African origin. *Public Health Nutr* 2001; 4:971-80.
37. Kusama K, Le DSNT, Hanh TTM, Takahashi K, Hung NTK, Yoshiike N, et al. Reproducibility and validity of a food frequency questionnaire among Vietnamese in Ho Chi Minh City. *J Am Coll Nutr* 2005; 24:466-73.
38. Block G, Wakimoto P, Jensen C, Mandel S, Green RR. Validation of a food frequency questionnaire for Hispanics. *Prev Chronic Dis* 2006; 3:A77.
39. Deschamps V, Lauzon-Guillain B, Lafay L, Borys JM, Charles MA, Romon M. Reproducibility and relative validity of a food-frequency questionnaire among French adults and adolescents. *Eur J Clin Nutr* 2009; 63:282-91.
40. Chen Y, Ahsan H, Parvez F, Howe GR. Validity of a food-frequency questionnaire for a large prospective cohort study in Bangladesh. *Br J Nutr* 2004; 92:851-9.
41. Freedman LS, Carroll RJ, Wax Y. Estimating the relation between dietary intake obtained from a food frequency questionnaire and true average intake. *Am J Epidemiol* 1991; 134:310-20.
42. Lima FEL, Slater B, Latorre MRDO, Fisberg RM. Validade de um questionário de frequência alimentar desenvolvido para população feminina no nordeste do Brasil. *Rev Bras Epidemiol* 2007; 10:483-90.
43. Scagliusi FB, Lancha Júnior AH. Subnotificação da ingestão energética na avaliação do consumo alimentar. *Rev Nutr* 2003; 16:471-81.
44. Boucher B, Catterchio M, Kreiger N, Nadalin V, Block T, Block G. Validity and reliability of the Block98 food-frequency questionnaire in a sample of Canadian women. *Public Health Nutr* 2006; 9:84-93.
45. Kroke A, Klipstein-Grobusch K, Voss S, Möseneder J, Thielecke F, Noack R. Validation of self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, end repeated 24-h dietary recall methods. *Am J Clin Nutr* 1999; 70:439-47.
46. Johansson I, Hallmans G, Wikman A, Biessy C, Riboli E, Kaaks R. Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease Cohort. *Public Health Nutr* 2002; 5:487-96.
47. Crispim SP, Ribeiro RCL, Silva MMS, Rosado LEFP, Rosado GP. The influence of education in the validation process of a food frequency questionnaire for adults in Viçosa, Minas Gerais, Brazil. *Eur J Clin Nutr* 2006; 60:1311-6.
48. Molag ML, Vries JHM, Ocké MC, Dagnelie PC, Brandt PAVD, Jansen MCJE, et al. Design characteristics of food frequency questionnaires in relation to their validity. *Am J Epidemiol* 2007; 166:1468-78.
49. Pereira RA, Araujo MC, Lopes TS, Yokoo EM. How many 24-hour recalls or food records are required to estimate usual energy and nutrient intake? *Cad Saúde Pública* 2010; 26:2101-11.
50. Masson LF, McNeill G, Tomany JO, Simpson JA, Peace HS, Wei L, et al. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutr* 2003; 6:313-21.
51. Pakseresht M, Sharma S. Validation of a quantitative food frequency questionnaire for Inuit population in Nunavut, Canada. *J Hum Nutr Diet* 2010; 23 (Suppl 1):67-74.
52. Crewson PE. Reader agreement studies. *Am J Roentgenol* 2005; 184:1391-7.
53. Slater B, Philippi ST, Fisberg RM, Latorre MRDO. Validation of a semi-quantitative adolescent food frequency questionnaire applied at a public school in São Paulo, Brazil. *Eur J Clin Nutr* 2003; 57:629-35.
54. Bellisle F. O método da água duplamente marcada e pesquisas de ingestão alimentar: um confronto. *Rev Nutr* 2001; 14:125-33.
55. Buzzard M. 24-hour dietary recall and food record methods. In: Willett W, editor. *Nutritional epidemiology*. New York: Oxford University Press; 1998. p. 50-73.
56. Slater B, Marchioni DML, Voci SM. Aplicação de regressão linear para correção de dados dietéticos. *Rev Saúde Pública* 2007; 41:190-6.
57. Voci SM, Slater B, Silva MV, Marchioni DML, Latorre MRDO. Estudo de calibração do Questionário de Frequência Alimentar para Adolescentes (QFAA). *Ciênc Saúde Coletiva* 2011; 16:2335-43.

58. Marchioni DML, Slater B, Fisberg RM. Minimizando erros na medida da ingestão dietética. In: Fisberg RM, Slater B, Marchioni DML, Martini LA, organizadores. Inquéritos alimentares: métodos e bases científicos. São Paulo: Edições Manole; 2005. p. 159-66.
59. Kynast-Wolf G, Becker N, Kroke A, Brandstetter BR, Wahrendorf J, Boeing H. Linear regression calibration: theoretical framework and empirical results in EPIC, Germany. *Ann Nutr Metab* 2002; 46:2-8.
60. Kaaks R, Riboli E, van Staveren W. Calibration of dietary intake measurements in prospective cohort studies. *Am J Epidemiol* 1995; 142:548-56.

Submitted on 17/Aug/2012

Final version resubmitted on 28/Mar/2013

Approved on 10/Apr/2013