Validity and reproducibility of a food frequency questionnaire for adults of São Paulo, Brazil

Validade e reprodutibilidade de um questionário de frequência alimentar para adultos de São Paulo, Brasil

Soraya Sant’Ana de Castro Selem, Aline Martins de Carvalho, Eliseu Verly-Junior, Jackeline Venâncio Carlos, Juliana Araujo Teixeira, Dirce Maria Lobo Marchioni, Regina Mara Fisberg

ABSTRACT: Objectives: To assess the validity and reproducibility of a food frequency questionnaire developed for estimating the food consumption of adults in São Paulo, Brazil, based population study. Methods: A sample of individuals aged above 20 years, of both genders, living in São Paulo, was used for the validation study (n = 77) and reproducibility study (n = 74) of the food frequency questionnaire. To verify the validity and reproducibility of energy and 19 nutrients were applied two food frequency questionnaires (60 items) and three 24-hour dietary recalls (24HR – reference method). The validity was verified by Spearman correlation coefficient (crude and de-attenuated) and weighted Kappa, and reproducibility by intraclass correlation coefficients and weighted kappa. Results: In analyzes of validity de-attenuated correlation coefficients ranged from 0.21 (carbohydrate) to 0.74 (energy), and weighted kappa exceeded 0.40 for 30% of the nutrients. Polyunsaturated fat and folate did not show significant correlation and weighted kappa. In reproducibility correlation coefficients ranged from 0.36 (polyunsaturated fat) to 0.69 (calcium), and weighted kappa exceeded 0.40 for 80% of the nutrients. Conclusion: The food frequency questionnaire analyzed has good validity and reproducibility for estimating the food consumption of adults in São Paulo compared to the reference method, so it is an appropriate instrument to be used in epidemiological studies on similar populations. Estimates of polyunsaturated fat and folate should be interpreted with caution.

Keywords: Diet surveys. Food-frequency questionnaire. Food consumption. Validation studies. Adult. São Paulo.
INTRODUCTION

In recent decades, the food frequency questionnaire (FFQ) has been the most used tool to assess food and nutrient intake in epidemiological studies, especially those whose goal is to relate dietary factors with health outcomes. This occurs due to the fact that this method is considered a survey that assesses usual diet presenting a low cost, easy application and processing, and generating information that allows sorting and categorizing of individuals according to exposure to certain dietary components\(^1\,^2\).

The development of an FFQ requires prior knowledge of food and nutrient intake in the population on which it will be applied. Usually, 24-hour dietary recalls (24HR) are applied in pilot studies, generating a database for establishment of food items that will make up the instrument, as well as the portion sizes\(^1\). Like any other instrument, the FFQ is not free of measurement errors. In this sense, validation and reproducibility studies are strategies to verify the accuracy and precision of their measurements\(^4\).

The FFQ developed by Fisberg et al.\(^3\) to estimate usual dietary intake of adults is the only one developed from a population-based study with a probability sample of the city of São Paulo, the most populous and multicultural city in Brazil. However, it is necessary to evaluate
the performance of this instrument for its application in this population. The objective of this study was to verify the accuracy and precision of the instrument.

**METHODS**

**THE INSTRUMENT**

The FFQ, designed to assess the habitual consumption of adults of both sexes during the year preceding its application, had as basis for its development dietary information obtained through a 24HR applied in a random sample of individuals aged 20 years or more (n = 1,477) residing in the city of São Paulo. This information was obtained in the Health Survey (ISA-Capital) study conducted in 2003. Foods that contributed up to 90% for the consumption of energy and nutrients were included and grouped, as well as foods that are sources of nutrients and that could discriminate the study subjects according to levels of consumption. The FFQ consists of 60 food items with frequencies of 0 – 10 times, and a unit of time: day, week, month and year. The portions were defined as small, medium, large and extra large, which correspond to the 25, 50, 75 and 95th percentiles of the consumption in grams of each food item. The median portion is the reference serving size, and is presented in household measures and in grams. Food items are organized into: soups and pasta, meat and fish, dairy products, legumes and eggs, rice and tubers, vegetables, sauces and spices, fruits, beverages, breads and biscuits, sweets and desserts. The FFQ also presents questions about recent changes in eating habits, consumption of dietary supplements, other important foods consumed that the instrument did not include, as well as consumption of visible fat from meat and skin from poultry.

**STUDY POPULATION AND DATA COLLECTION**

Between 2007 and 2008, a study was conducted in a representative subsample of the participants of ISA-Capital 2003 of both sexes, aged 20 or more, to estimate the prevalence of inadequate nutrient intake. The FFQ used was developed based on data from the 2003 ISA-Capital, and its accuracy and precision were verified in the present study. Two home visits, with an interval of one year between them, were performed for data collection, during which one FFQ and one 24HR — the reference method for validation — were conducted in each of the visits. The one-year interval between data collection was determined because the FFQ was developed to assess habitual intake during the year preceding its application. In addition, in the period between visits, another 24HR was collected by phone, and the average interval between 24HR collections was six months (Figure 1). The size of the subsample studied was reduced due to loss of follow-up, whose main reason was the failure to locate the individual. In the first home visit, 607 FFQs were collected, and 84 FFQs in
the second visit. Also, 20 individuals who presented under- and overestimation of food consumption were excluded from the analysis, according to the method proposed by Nielsen⁶. At the end, 77 individuals (34 men and 43 women) responded to the three 24HR and to the second FFQ, and 74 (32 men and 42 women) responded to the two FFQ.

For the application of the 24HR, a standard form and the Multiple Pass Method, developed by the United States Department of Agriculture (USDA) to assist in the collection of food consumption data, were used, minimizing errors in the measurement of diet⁷. The 24HR were applied so that all days of the week and seasons were represented. The standardization and quantification of foods and preparations were performed according to national recommendations, which allowed the conversion of household measurements in units of weight or volume⁸,⁹.

To quantify the energy and nutrients consumed, obtained by the FFQ and 24HR, the Nutrition Data System for Research software (version 2007, 2007, University of Minnesota, Minneapolis) was used.

**DATA ANALYSIS**

The following were selected for analysis: energy, protein, carbohydrates, fiber, total fat, saturated fat, monounsaturated fat, polyunsaturated fat, trans fats, cholesterol, vitamin A, beta-carotene, folate, vitamin C, vitamin E, calcium, phosphorus, iron, zinc and sodium. In all
analyzes regarding intake of nutrients (validation and reproducibility), intakes of nutrients values were used, adjusted for energy intake by the residual method\(^1\).

Validity was determined by comparing the values of energy intake and each nutrient by the second FFQ (FFQ2, applied the second home visit) and the average of the three 24HR. The Spearman’s correlation coefficient and weighted Kappa were used to verify compliance according to consumption quartiles.

To minimize the effect of intra-subject variability of intake measured by the 24HR, each correlation coefficient was multiplied by a deattenuation coefficient. This was calculated using the following formula:

\[
r_v = r_o \left(1 + \frac{\lambda}{n}\right)^{1/2}
\]

where \(\lambda\) is the ratio between intrapersonal variance and interpersonal intake variance; \(r_o\) is the observed correlation, \(n\) is the number of days of 24HR and \(r_v\) is the deattenuated correlation. The intra- and interpersonal variances were calculated for each nutrient using the method proposed by Iowa State University (ISU)\(^10\).

Reproducibility was checked by comparing the values of energy intake and each nutrient obtained by the first FFQ (FFQ1, applied on the first home visit) with those obtained by FFQ2, using the intraclass correlation and weighted Kappa to verify compliance according to consumption quartiles.

All statistical analyzes were performed using Stata software (version 10, 2007, StataCorp LP, College Station, Texas). A significance level of 5% was considered.

This study was approved by the Research Ethics Committee of the School of Public Health, University of São Paulo, and the participation was conditioned upon signing of a free and informed consent form. There are no conflicts of interest.

**RESULTS**

The average consumption of energy and 10 nutrients estimated by FFQ2 was similar to that estimated by 24HR. In comparison with FFQ1, FFQ2 underestimated the consumption of 15 nutrients (Table 1).

In validity analyzes, the correlation coefficients ranged from 0.25 (carbohydrates) to 0.63 (energy). After deattenuation, there was a higher correlation for 12 nutrients, and the coefficients ranged from 0.21 (carbohydrates) to 0.74 (energy). The weighted Kappa exceeded 0.40 for 30% of the nutrients analyzed, and 45% of the nutrients showed values between 0.30 and 0.40. Polyunsaturated fat and folate showed no significant correlation and weighted Kappa compared to the reference method (Table 1).

In reproducibility, the correlation coefficients ranged from 0.36 (polyunsaturated fat) to 0.69 (calcium), and the weighted Kappa exceeded 0.40 for 80% of the nutrients (Table 1).
Table 1. Estimation of daily intake, correlation coefficients and concordance analysis to validation and reproducibility analyses.

<table>
<thead>
<tr>
<th></th>
<th>Estimation of daily intake</th>
<th>Difference</th>
<th>Correlation Coefficients</th>
<th>Concordance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FFQ1 (n = 74)</td>
<td>FFQ2 (n = 77)</td>
<td>24HR (n = 77)</td>
<td>FFQ1 x FFQ2 (n = 74)</td>
</tr>
<tr>
<td></td>
<td>Mean (adjusted by energy)</td>
<td>p-value</td>
<td>Crude</td>
<td>Deattenuated</td>
</tr>
<tr>
<td>Energy (kcal)</td>
<td>1692.58</td>
<td>1566.96</td>
<td>1834.07</td>
<td>0.15</td>
</tr>
<tr>
<td>Protein (g)</td>
<td>66.32</td>
<td>64.41</td>
<td>68.34</td>
<td>0.01</td>
</tr>
<tr>
<td>Carbohydrate (g)</td>
<td>238.39</td>
<td>213.07</td>
<td>200.2</td>
<td>0.00</td>
</tr>
<tr>
<td>Fiber (g)</td>
<td>18.87</td>
<td>16.77</td>
<td>14.62</td>
<td>0.00</td>
</tr>
<tr>
<td>Total Fat (g)</td>
<td>59.48</td>
<td>54.82</td>
<td>59.61</td>
<td>0.01</td>
</tr>
<tr>
<td>Saturated fat (g)</td>
<td>18.52</td>
<td>17.62</td>
<td>18.18</td>
<td>0.11</td>
</tr>
<tr>
<td>Monounsaturated fat (g)</td>
<td>21.12</td>
<td>19.65</td>
<td>20.9</td>
<td>0.03</td>
</tr>
<tr>
<td>Polyunsaturated fat (g)</td>
<td>13.58</td>
<td>12.22</td>
<td>13.57</td>
<td>0.00</td>
</tr>
<tr>
<td>Trans Fat (g)</td>
<td>2.99</td>
<td>2.61</td>
<td>2.27</td>
<td>0.15</td>
</tr>
<tr>
<td>Fat (g)</td>
<td>160.42</td>
<td>150.63</td>
<td>173.71</td>
<td>0.17</td>
</tr>
<tr>
<td>Vitamin A (IU)</td>
<td>864.55</td>
<td>736.1</td>
<td>651.66</td>
<td>0.00</td>
</tr>
<tr>
<td>Beta-Carotene (mcg)</td>
<td>3634.96</td>
<td>3134.33</td>
<td>1991.2</td>
<td>0.02</td>
</tr>
<tr>
<td>Folate (mg)</td>
<td>568.66</td>
<td>518.39</td>
<td>505.67</td>
<td>0.00</td>
</tr>
<tr>
<td>Vitamin C (mg)</td>
<td>134.83</td>
<td>118.48</td>
<td>81.74</td>
<td>0.01</td>
</tr>
<tr>
<td>Vitamin E (mg)</td>
<td>5.49</td>
<td>4.95</td>
<td>4.93</td>
<td>0.00</td>
</tr>
<tr>
<td>Calcium (mg)</td>
<td>612.2</td>
<td>598.08</td>
<td>519.2</td>
<td>0.05</td>
</tr>
<tr>
<td>Phosphorus (mg)</td>
<td>939.51</td>
<td>909.23</td>
<td>936.69</td>
<td>0.00</td>
</tr>
<tr>
<td>Iron (mg)</td>
<td>12.6</td>
<td>11.8</td>
<td>12.68</td>
<td>0.00</td>
</tr>
<tr>
<td>Zinc (mg)</td>
<td>9.38</td>
<td>8.6</td>
<td>9.31</td>
<td>0.00</td>
</tr>
<tr>
<td>Sodium (mg)</td>
<td>2938.33</td>
<td>2695.47</td>
<td>2692.73</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*p < 0.05, FFQ 1: food frequency questionnaire - application on the first home visit, FFQ 2: food frequency questionnaire - application on the second home visit, 24HR: 24-hour dietary recall.
DISCUSSION

This study suggests that the FFQ developed to estimate usual dietary intake of adults in the city of São Paulo has good validity and reproducibility for most nutrients analyzed. The underestimation of the second FFQ compared to the first has been observed in other studies\(^\text{11,12}\). This may occur because the participants learn to respond to the questionnaire more accurately. And it is possible to assume that the overestimation of consumption of four nutrients (carbohydrates, fiber, beta-carotene and vitamin C) by FFQ2 in comparison with 24HR happens by overestimating the consumption of sources of these nutrients, such as fruits, a common bias among individuals seeking for social approval of their diet\(^\text{13}\). It is noteworthy that the overestimation of nutrients can hide the poor consumption of the population\(^\text{11}\).

The FFQ did not show good validity in estimating polyunsaturated fat and folate, a result observed in other studies\(^\text{4,14,15}\). Thus, it is considered that the estimates of polyunsaturated fat and folate should be interpreted with caution. Nevertheless, the performance of the analyzed FFQ can be considered satisfactory, and the results obtained are similar to other published studies, considering the diversity of these in relation to the study population, the reference method, the time elapsed between the surveys and other characteristics of the instrument\(^\text{14-16}\).

The study has some strengths, such as the fact that the FFQ was developed based on a probability sample of São Paulo, the largest city in Brazil, with immigrants from all over the country and a diet that is influenced by different cultures. In addition, there was an effort to seek the participants of the original study to evaluate the performance of the FFQ. The method for calculation of intra- and interpersonal variances used in the deattenuation is considered one of the best\(^\text{10}\). Finally, accuracy and precision were evaluated for energy and 19 nutrients, higher than in most published studies, which often refer only to the number of macronutrients.

There are also limitations to consider. Although statistical procedures aimed at correcting the correlation coefficients were applied, providing a better estimate, the reference method used has sources of errors in common with the FFQ, because it depends on the respondent’s memory and ability to recall and report the foods and portions consumed. The use of recovery biomarkers as the reference method in validation analysis could provide more reliable results for the study, but these tests are expensive, often invasive, besides being available for few nutrients. Thus, the use of 24HR is still a viable and widely used alternative\(^\text{3}\). The sample size may compromise the statistical power of the analysis, but reduced sample sizes are common in this type of study due to loss of follow-up\(^\text{11,12,17,18}\).

CONCLUSION

The FFQ considered has good validity and reproducibility to estimate the intake of most nutrients in adults in the city of São Paulo compared to the reference method, so it is an appropriate instrument for epidemiological studies with similar populations. Estimates of polyunsaturated fat and folate, however, should be interpreted with caution.
REFERENCES

17. Mendes LL, Campos SF, Malta DC, Bernal RTI, Sá NNB, Velásquez-Meléndez G. Validade e reproducibilidade de marcadores do consumo de alimentos e bebidas de um inquérito telefônico realizado na cidade de Belo Horizonte (MG), Brasil. Rev Bras Epidemiol 2011; 14(1 Suppl): 80-9.

Received on: 01/09/2014
Final version presented on: 02/19/2014
Accepted on: 04/29/2014