Reproducibility and Biomarker-based Validity and Responsiveness of a Food Frequency Questionnaire to Estimate Protein Intake

Loek T. J. Pijls,1 Hendrik de Vries,1 Ab J. M. Donker,2 and Jacques Th. M. van Eijk1

This study, conducted in the Netherlands in 1994–1996, assessed the reproducibility, validity, and responsiveness of a food frequency questionnaire (FFQ) used with older subjects with type 2 diabetes. During a period of 6 months, 93 subjects had a stable protein intake according to their urinary urea excretion (UUE). An FFQ was completed before and after this period. Reproducibility was evaluated on the basis of the correlation (0.80) between duplicate FFQ estimates of protein intake. At the group level, the smallest detectable difference was 3 g. For 354 subjects, average protein intake according to the FFQ was similar to that from UUE estimates. Validity was evaluated on the basis of the crude and deattenuated correlations, which were 0.31 and 0.39, respectively. The crude correlation was highest for former smokers (r = 0.39) and for patients with a higher body mass index (r = 0.37), and it was very low for current smokers (r = 0.14, p > 0.20). The correlation did not vary with age or gender. Responsiveness was judged on the crude correlation between changes in intake according to the FFQ and UUE; the correlation was 0.22 (n = 134). The responsiveness ratio was 0.73. The authors concluded that FFQ estimates are well reproducible. At the group level, protein intake was estimated correctly, whereas individual protein intake and changes in intake were reflected to a moderate extent only. Am J Epidemiol 1999; 150:987–95.

Information on intake of nutrients is the basis for defining the associations between diet and health. If this information is inaccurate, spurious conclusions will result with regard to the effects of diet on health. Food frequency questionnaires (FFQs) potentially yield sufficiently accurate estimates of diet intake, are relatively inexpensive to administer, and make only modest demands on both subjects and researchers (1–4). To gain an understanding of the role of nutrition in good health at an older age, researchers need instruments that yield accurate intake estimates for older people (5).

The purpose of this study was to describe reproducibility and biomarker-based validity and responsiveness of a semiquantitative FFQ when used with older people with type 2 diabetes. It is well known that the numbers of type 2 diabetic patients are large and increasing, and in this group nutrition probably is of particular importance. FFQ estimates of protein intake were compared with estimates based on urinary urea excretion (UUE), the most valid method of estimating protein intake (6–8).

Low reproducibility definitely indicates that long-term intake or changes in intake have not been estimated validly (1, 9). Reproducibility of diet intake estimates is usually assessed when it is unknown whether actual change has occurred. Under such circumstances, it is not possible to determine, for instance, whether reproducibility has been underestimated as a result of actual changes having occurred. Thus, in our study reproducibility was assessed among subjects who had a stable protein intake according to UUE.

The validity of using the present FFQ compared with the dietary history method has already been demonstrated (10). However, since both methods depend on the patient’s memory and on use of the same food composition tables, good agreement does not necessarily indicate high validity. Therefore, the use of biomarkers is strongly recommended (1, 2, 11). Their sources of error, for example, biologic variation and variation in chemical analysis, are principally different from errors in questionnaire estimates (1, 3). The biomarker-based
validity of the FFQ was assessed by comparing its estimates of protein intake with UUE estimates.

If an instrument is used in intervention studies, its responsiveness should be described (12). Possibly because an FFQ generally is designed to estimate average intake during a minimum period of 1 month (13), its comparative responsiveness has been reported only once (12). To our knowledge the present study, which was conducted in 1994–1996, is the first to report on the biomarker-based responsiveness of an FFQ.

MATERIALS AND METHODS

Design

Reproducibility was evaluated on the basis of duplicate questionnaires that were completed by selected subjects who, according to their UUE, had a relatively stable protein intake. Biomarker-based validity of FFQ estimates of protein intake was assessed by cross-sectional comparison with UUE estimates. Responsiveness was evaluated by comparing, during a period of 6 months, changes in protein intake according to UUE with changes as estimated by the FFQs. In addition, "signal-to-noise" or responsiveness ratios were calculated (9). Diet-health associations can be confounded if the qualities of the instruments used depend on subject characteristics. Therefore, we assessed whether the validity of the FFQ varied across categories of age, gender, education, smoking habits, and degree of overweight. The study design was approved by the Medical Ethics Committee of the Vrije Universiteit Amsterdam, the Netherlands.

Patients

Validity was assessed among 354 type 2 diabetic (14), primary care patients, who were selected by their general practitioners. Patients were younger than 79 years of age, were not in a period of recovery from severe illness, and did not have protein-losing enteropathy, venous leg ulcer, pressure ulcer, malignancy, or psychiatric or serious psychosocial problems. Their mean age was 64 years (standard deviation (SD), 9), 49 percent were men, and there were no vegetarians.

Responsiveness was studied among 134 of 354 patients. Their mean age was 64 years (SD, 8), and 59 percent were men. Patients met the following criteria: 1) microalbuminuria (30–300 mg/24 hours, mean of two 24-hour samples), relatively high albuminuria within the normoalbuminuric range (albuminuria >20 mg/24 hours in at least one sample or a urinary albumin concentration of >6.5 mg/liter in two samples), or diabetes duration of ≥5 years; 2) a baseline protein intake of ≥0.80 g/kg of body weight; 3) good compliance with dietitian appointments; 4) the ability, according to the dietitian, to understand diet information; 5) no reluctance to try to change their diet; and 6) provision of informed consent. A subgroup of 68 of the 134 patients was randomly allocated to receive dietary advice on protein restriction for a period of 18 months or longer.

Reproducibility of the FFQ estimates was assessed in a subgroup of 93 patients (mean age, 64 years (SD, 8); 57 percent men) with a stable protein intake. (Refer to the Data analysis section.)

Measurements

Protein intake was estimated from the FFQ and from UUE. The FFQ focused on the previous month, contained 75 items, and comprised 10 categories of frequency (10). By using the 1993 version of the Dutch food composition table (15), we converted food intake as recorded on the FFQ to intake of nutrients and energy. Subjects visited a dietitian, who explained the FFQ. It was filled in by subjects at home and was checked by the dietitian who, for example, determined whether the frequencies of consumption of different kinds of meat equaled the total frequency of meat consumption reported elsewhere on the FFQ. Any inconsistency was resolved together by the subject and the dietitian. To obtain the reference estimate of protein intake, the Maroni et al. formula (16) was applied to the UUE from duplicate 24-hour samples. The latter was estimated by multiplying 24-hour urine volume by urea concentration. This concentration was determined by using a BM/Hitachi 747/737 laboratory kit (Boeringer Mannheim, Mannheim, Germany). The laboratory device used was a Hitachi 747.

Data analysis

Calculations were made with the aid of SPSS software, version 5.02 (SPSS Inc., Chicago, Illinois). It was observed that the Spearman's correlations (data not shown) were of the same magnitude as the Pearson's correlations. Skewness apparently did not influence the Pearson's correlations; thus, no transformation was applied.

Reproducibility was assessed in a subgroup of 93 subjects in which protein intake, according to UUE, was relatively stable over 6–12 months. This period, compared with the first 6 months, was assumed to be the most suitable on which to base reproducibility estimates. Stable intake was defined as a change of less than 0.20 g/kg per day (17). Reproducibility was described by using Pearson's correlations between both
crude and energy-adjusted duplicate estimates of intake (1, 18, 19). We adjusted energy intake by using the residuals from regression analysis, with nutrient intake as the dependent variable and energy intake as the independent variable (1, 18). The smallest detectable difference in the group of subjects was estimated by using the following formula: \( \sqrt{(Z_\alpha + Z_\beta)^2 	imes \sigma^2} / n \) (9). In this formula, \( \alpha = 0.05 \) (Z\(_\alpha = 1.96\)), \( \beta = 0.20 \) (Z\(_\beta = 0.85\)), \( \sigma^2 \) = intrapersonal variance = variance in individual differences between two FFQ estimates / 2, and n = 93. The finding that a difference between two estimates varies according to level of intake is referred to as proportional bias (2, 20). Its presence was assessed by means of linear regression analysis; the difference between the estimates was entered as the dependent variable and their mean as the independent variable (2, 20).

To determine validity, we first assessed systematic bias, that is, the mean of the differences between the FFQ and UUE estimates. Proportional bias was determined as described for reproducibility. The Pearson’s correlation coefficient between the two estimates was calculated and deattenuated by adjustment for within-person variation in both the UUE and FFQ estimates (1, 21, 22). The deattenuation implied that the crude \( r \) was multiplied by

\[
\sqrt{(1 + \frac{SD_{UUE\, within}}{SD_{UUE\, between}})} \times \sqrt{(1 + \frac{SD_{FFQ\, within}}{SD_{FFQ\, between}})}
\]

(1, 21, 22). SD\(_{between}\) was estimated from the SD\(_{6\, months}\) in the subgroup with a stable protein intake between 6 and 12 months (n = 93). SD\(_{within}\) was estimated as the SD\(_{change\, 12-6\, months}\) in the same subgroup, divided by 2.

Responsiveness was evaluated similarly. Calculated were the difference and correlation between estimates of change over the period between baseline and 6 months. Correlation coefficients were deattenuated by applying the formula given above. SD\(_{between}\) was estimated from the SD\(_{change\, 12-6\, months}\) in the above-mentioned subgroup with a stable protein intake between 6 and 12 months (n = 93). SD\(_{within}\) was estimated as the SD\(_{change\, 18-12\, months\,-\, change\, 12-6\, months}\) in the same subgroup, divided by 2. In addition, the responsiveness ratio was computed (9). It was defined as the ratio between the 1) decrease in protein intake\(_{FFQ}\) in the subgroup of subjects who, according to UUE, had reduced their protein intake by more than 0.20 g/kg per day (17) between baseline and 6 months and 2) standard deviation of the change in protein intake\(_{FFQ}\) in the subgroup with a stable protein intake\(_{UUE}\).

RESULTS

Reproducibility

The crude correlation was 0.80 between duplicate estimates of protein intake (table 1). It ranged from 0.47 for plant protein to 0.88 for alcohol. Energy-adjusted correlations were somewhat weaker, except for plant protein (r = 0.64). No proportional bias was observed for total protein intake or for most other nutrients listed in table 1 (p > 0.05). However, for plant protein, linear regression analysis showed that there was no difference at an intake of 20 g, and beta (per 10 g) was -3.8 g (95 percent confidence interval (CI): -6.2, -1.3). This indicates that with every 10 g that the mean of duplicate estimates exceeded 20 g, the estimate based on the second FFQ was about 4 g lower than that based on the first FFQ. Carbohydrate estimates from the two FFQs were more or less similar at about 151 g, and beta (per 10 g) was significant but small: 1.7 g (95 percent CI: 0.4, 3.1). When 93 patients were considered, the smallest detectable difference for total protein intake was 2.7 g; for one person, the smallest detectable difference would have been 26 g.

Validity

The mean FFQ estimate of protein intake (88 g; SD, 24) was of the same magnitude as the UUE estimate (84 g; SD, 21), and the difference between them, as shown in table 2, was far from statistically significant (p > 0.20). Proportional bias was present. Regression analysis showed no difference at an intake of 70 g, and beta (per 10 g) was 2.0 g (95 percent CI: 0.4, 3.5). Thus, intakes of less than 70 g were underestimated and higher intakes were overestimated by the FFQ. Per every 10 g that the mean of the two estimates exceeded 70 g, the FFQ estimate was 2.0 g higher than that obtained by using UUE. However, the difference in relative terms did not depend on the level of intake (p > 0.20).

Protein intake was overestimated particularly among current smokers and among subjects with a low body mass index. The crude correlation between the two estimates of protein intake was 0.31. If deattenuated for the variation in both the FFQ and UUE estimates, the correlation was 0.39 (table 2); if deattenuated for such variation in the UUE estimates only (23), it was 0.35. The crude correlation was higher for former smokers (r = 0.39) and for patients with a higher body mass index (r = 0.37), and it was strikingly low and not even statistically significant for current smokers (r = 0.14, p > 0.20). The correlation tended to be higher for subjects with a lower level of education but did not vary with age or gender.
TABLE 1. Reproducibility data: mean intake of protein and other macronutrients and energy according to a food frequency questionnaire (FFQ), Pearson's correlation between duplicate FFQ estimates for subjects with stable protein intake* according to urinary urea excretion, and smallest detectable differences for 93 patients with type 2 diabetes, the Netherlands, 1994–1996

<table>
<thead>
<tr>
<th>Nutrient or energy</th>
<th>Unit</th>
<th>FFO 1</th>
<th>FFO 2</th>
<th>Pearson's r</th>
<th>SDD§</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Crude</td>
<td>Energy adjusted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total protein</td>
<td>g/day</td>
<td>87 (21)</td>
<td>0.80</td>
<td>0.76</td>
<td>2.7</td>
</tr>
<tr>
<td>Animal protein</td>
<td>g/day</td>
<td>58 (18)</td>
<td>0.80</td>
<td>0.77</td>
<td>2.3</td>
</tr>
<tr>
<td>Plant protein</td>
<td>g/day</td>
<td>26 (9)</td>
<td>0.47</td>
<td>0.64</td>
<td>1.6</td>
</tr>
<tr>
<td>Total fat</td>
<td>g/day</td>
<td>65 (19)</td>
<td>0.71</td>
<td>0.58</td>
<td>2.9</td>
</tr>
<tr>
<td>Saturated fat</td>
<td>g/day</td>
<td>23 (7)</td>
<td>0.75</td>
<td>0.70</td>
<td>1.1</td>
</tr>
<tr>
<td>Total carbohydrate</td>
<td>g/day</td>
<td>154 (44)</td>
<td>0.82</td>
<td>0.76</td>
<td>6.0</td>
</tr>
<tr>
<td>Dietary fiber</td>
<td>g/day</td>
<td>14 (4)</td>
<td>0.85</td>
<td>0.82</td>
<td>0.5</td>
</tr>
<tr>
<td>Alcohol</td>
<td>g/day</td>
<td>6.8 (11)</td>
<td>0.88</td>
<td>0.87</td>
<td>1.0</td>
</tr>
<tr>
<td>Energy</td>
<td>MJ/day</td>
<td>6.6 (2)</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Change <0.20 g/kg per day (17).
† Mean intake (standard deviation) according to the first of two consecutive FFQs.
‡ Energy adjustment according to the residual method; the residuals from regression analysis were used, with nutrient intake as the dependent variable and energy intake as the independent variable (1, 18); p < 0.001.
§ SDD, smallest detectable difference = √((2α + Zβ)² × σ²)/n
α = 0.05 (Zα = 1.96), β = 0.20 (Zβ = 0.85), and n = 93
σ² = intridual variance estimated by: σ² of differences between two FFQ measurements / 2

Responsiveness

According to the FFQ and to UUE, the mean change in protein intake during a period of 6 months was -11 g (SD, 24) and 0 g (SD, 17), respectively (table 3). The difference in these two estimates varied with their mean; that is, proportional bias occurred. When this mean indicated a decrease, the FFQ overestimated this

TABLE 2. Biomarker-based validity of protein intake estimates: discrepancy and correlation between estimates from a food frequency questionnaire (FFQ) and urinary urea excretion (UUE), patients with type 2 diabetes, the Netherlands, 1994–1996

<table>
<thead>
<tr>
<th>Subject characteristics</th>
<th>No.</th>
<th>FFQ-UUE difference</th>
<th></th>
<th></th>
<th>Pearson's r</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Absolute (g/day)</td>
<td>Relative (%)*</td>
<td></td>
<td>Crude</td>
<td>p value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean† p value‡</td>
<td>Mean† p value‡</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>354</td>
<td>3 (26)</td>
<td>9 (38)</td>
<td>0.31§</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>169</td>
<td>2.6 (29)</td>
<td>6.9 (36)</td>
<td>0.28</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>185</td>
<td>3.7 (24)</td>
<td>11.0 (40)</td>
<td>0.21</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤65</td>
<td>164</td>
<td>1.5 (2.9)</td>
<td>7.0 (41)</td>
<td>0.31</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>&gt;65</td>
<td>190</td>
<td>4.7 (24)</td>
<td>10.9 (35)</td>
<td>0.28</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>184</td>
<td>4.0 (25)</td>
<td>9.6 (35)</td>
<td>0.36</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>125</td>
<td>1.2 (29)</td>
<td>7.0 (40)</td>
<td>0.24</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>106</td>
<td>4.5 (24)</td>
<td>10.9 (38)</td>
<td>0.26</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>151</td>
<td>-1.0 (25)</td>
<td>2.4 (33)</td>
<td>0.39</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>57</td>
<td>8.8 (32)</td>
<td>18.0 (48)</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤27</td>
<td>166</td>
<td>9.1 (26)</td>
<td>&lt;0.001</td>
<td>16.5 (37)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>&gt;27</td>
<td>187</td>
<td>-2.1 (28)</td>
<td>2.3 (37)</td>
<td>0.37</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

* 100 x (FFQ-UUE)/UUE.
† Standard deviation in parentheses.
‡ Versus other subgroup; listed only if p ≤ 0.20.
§ Deattenuated (1) Pearson's correlation, 0.39.
¶ Former versus current smokers.
TABLE 3. Two approaches* to determining the biomarker-based responsiveness of a food frequency questionnaire (FFQ) for patients with type 2 diabetes, the Netherlands, 1994–1996

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age (years)</th>
<th>Protein intake</th>
<th>Responsiveness ratio††</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;65</td>
<td>&gt;65</td>
<td>Substantial decrease¶</td>
</tr>
<tr>
<td></td>
<td>≤27</td>
<td>&gt;27</td>
<td>Stable§</td>
</tr>
<tr>
<td></td>
<td>Substantial increase**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* 1) The discrepancy and correlation between FFQ and urinary urea excretion (UUE) estimates of change in protein intake over 6 months and 2) the responsiveness ratio for change in protein intake according to the FFQ.
†† Mean (standard deviation in parentheses).
¶ Versus other subgroup (does not apply to Pearson's $r$); listed only if $p < 0.20$.
§ Deattenuated (1) Pearson's correlation, 0.96.
6 months – intake0 months < −0.20 g/kg per day according to UUE.
5 −0.20 g/kg per day ≤ (intake6 months – intake0 months) ≤ 0.20 g/kg per day according to UUE.
** Intake6 months – intake0 months > 0.20 g/kg per day according to UUE.
†† Ratio of mean change according to the FFQ for patients with a decreased protein intake according to UUE, to the standard deviation of change according to the FFQ in a group with unchanged intake according to UUE.

Am J Epidemiol Vol. 150, No. 9, 1999
on the relation between diet and health status has been obtained from studies with selected groups of participants who were rather aware of their diet and health.

Reproducibility

To our knowledge, the present study is the first to assess reproducibility of an FFQ in a group of subjects who, according to a reference method such as UUE, have a relatively stable diet intake. The FFQ was as reproducible as other self- or interviewer-administered, 44- to 104-item FFQs with an interval of 4–6-months (22, 24, 25) and a diet history (26). In a Spanish population, 4-day food records and FFQs with 3-month and more than 12-month intervals, respectively, were less reproducible (27). Note, however, that food records estimate only a few days of diet intake, and it would appear that lower reproducibility possibly reflects residual day-to-day variation. An FFQ usually estimates diet intake over several weeks, so that day-to-day variation plays a minor role. Furthermore, this and other studies (22, 28) show that the reproducibility of the estimates of intake of a certain nutrient, for example, alcohol, can be very high if the number of foods containing that nutrient is rather small ($r = 0.88$).

Essential in the assessment of reproducibility is a washout period between applications. When completing the second questionnaire, subjects should no longer remember what they filled in on the first one. In contrast with previous studies, the present study design made it possible to exclude subjects with substantial diet changes as determined by an objective criterion. If no such reference method is available, the washout period should, to minimize real changes, not be too long (1, 13). In previous studies, there was at least 1 (23, 27–30) or even 3 years (31) between applications. This also could have contributed to the findings, compared with those in the present study, of lower estimates of reproducibility. The same applies to the fact that, in the present study, reproducibility was assessed on a second and third FFQ filled out by the subjects.

Validity

An instrument’s ability to rank persons according to their actual intake is generally viewed as its most essential characteristic (1, 2). This ranking ability can be evaluated by using Spearman’s correlation. It has also been advocated that to set diet guidelines and to evaluate the magnitude of purported health problems, instruments should estimate absolute intakes (32). The Pearson’s correlation reflects the degree of agreement between the absolute levels of estimates. In the present study, no clear discrepancies were observed between Pearson’s and Spearman’s correlations (data not shown).

Crude and deattenuated Pearson’s correlations between estimates of protein intake according to the FFQ and UUE were 0.31 and 0.39, respectively. In contrast with an FFQ and UUE, FFQs and food records depend on the same food composition tables and, at least to some extent, on subjects’ memory. Thus, correlations between the latter two methods may overestimate true validity. FFQ estimates of protein intake and the mean of two 7-day food records demonstrated deattenuated correlations of 0.32 (first food record) and 0.44 (second food record) (23), which are of the same magnitude as the correlation between the FFQ and UUE in the present study. It can be questioned whether a correlation of about 0.4 is good enough (32). Such correlation implies that the FFQ estimates explain only $(0.4)^2 \times 100 = 16$ percent of the variation in UUE estimates.

The precoded character of FFQs limits their size but may also have disadvantages in comparison with food records and 24-hour recalls. Foods are grouped if they contain similar amounts of certain nutrients. This may limit the validity of FFQ estimates of other nutrients whose content varies largely within the food groups listed (32). In the present FFQ, this probably applied merely to micronutrients rather than to macronutrients. Precoding also results in a loss of information on preparation methods and on type or brand of food, for example (32). However, the present FFQ did include detailed information on the types and brands of fat used in frying and with bread and, for example, on the types of meat and dairy products used. It also contained an open-ended question so that subjects could record foods not listed.

To date, validity of estimates of protein intake based on an FFQ, compared with the reference method based on UUE, has been reported for four FFQs (22, 33, 34). A Swedish FFQ yielded a weaker crude correlation ($r = 0.19$), even when subjects reporting incompleteness of 24-hour urine samples were excluded ($r = 0.27$) (34). This finding is noteworthy since that FFQ contained 224 items and UUE was measured over a period of 4 days, whereas the FFQ used in the present study included only 75 items and UUE was measured over 2 days. In two American, 130-item FFQs that also used duplicate samples, correlations with UUE were weak ($r = 0.15$ and $r = 0.24$) (33). According to the Swedish (34) and our FFQ, protein intakes were on average 25 and 9 percent higher, respectively, than those based on UUE. Yet another Dutch FFQ, compared with UUE, estimated protein intake to be 9 g/day lower (22).

As far as recall methods are concerned, diet history is the generally accepted reference method (1, 2).
Compared with the agreement found between our FFQ and UUE, the diet history's association with UUE estimates appeared to be only slightly stronger (26). Correlations between UUE and 3-day (35) or 4-day (34) food records are reported to be similar ($r = 0.31$) or higher ($r = 0.47$), respectively. In addition to UUE, there are other biomarkers for diet intake, for example, fatty acids in adipose tissue and α-tocopherol and β-carotene in serum. Compared with our findings, the correlations of these biomarkers with diet history estimates of fatty acids and vitamins E and A have been found to be similar or weaker (22, 36, 37) or stronger (38).

Closer agreement between protein intake estimated by the UUE reference method and by an FFQ is probably not feasible. Even if urine is collected during a period of several days, residual day-to-day variation will result in an underestimation of true validity. The same applies to imprecision resulting from incompleteness of 24-hour urine samples. Biochemical markers are available that estimate sample completeness according to which adjustments can be made (8). However, the question is whether the use of such markers is feasible in large-scale studies. Most of the variation in FFQ estimates of nutrient intake is explained by frequency of consumption. Yet the incremental information obtained from further extending the number of categories of frequency is limited. Extension of the number of categories of portion size and inclusion of novel foods do not yield much incremental information either (11, 39).

In one other study, subjects who had a fairly high level of education appeared to be able to keep diet records relatively well (35). In the present study, however, we found that FFQ validity with regard to estimates of protein intake tended to be higher among subjects with a lower level of education.

Energy intake is known to be underestimated by both diet records and recall methods (40, 41). As in two-thirds of other studies (42), comparison of our FFQ estimates of energy intake with estimated basal metabolic rate (43–46) suggested substantial underreporting. An average ratio of 1.05 (SD, 0.28) was observed, which is much lower than the expected ratio of 1.55 for a sedentary population (43–46). Thus, since the FFQ overestimated protein intake by less than 10 percent but underestimated energy intake, it must also have underestimated intake of fats and carbohydrates. However, this may have been more typical of the present study population than for the FFQ itself. For decades, diabetes patients have been told to restrict their intake of carbohydrates and, more recently, fat. They apparently underreported their consumption of foods rich in carbohydrates and fat but not their consumption of the most important sources of protein. For similar reasons, the present (data not shown) and other population-based studies (22, 47) show that energy intake is underestimated in particular by overweight subjects.

### Responsiveness

We know of no other study that has assessed the responsiveness of an FFQ in comparison with the reference method based on UUE. Considering the crude correlation between changes according to the FFQ and to UUE, the FFQ was only moderately responsive. Expectedly, the latter correlation was lower than the correlation between single estimates; imprecision in single estimates adds up when calculating changes. Liu et al. (21) and Willett (1) described a method to deattenuate correlations based on cross-sectional comparisons. Since comparison of two estimates of change over the same time period can also be regarded as cross-sectional, the correlation between the change estimates was deattenuated accordingly. However, since this application is new and unexplored, the resulting very high correlation of 0.96 should be interpreted with great caution.

The signal-to-noise ratio, which should exceed 1.0 (9), was only 0.73. As was the case in a previous study involving patients who received education on protein restriction (35), a discrepancy was found between change according to the FFQ and to UUE. It illustrates the subjectivity of the FFQ: subjects reported, to some extent, a desirable instead of their actual diet. However, the larger decrease according to the FFQ compared with UUE may also have resulted from the well-known fact that the estimates of food records and questionnaires generally tend to be lower when applied the second versus the first time. The design of the FFQ used in the present study permitted ample variability in frequency of consumption and in portion sizes; an increase would expectedly not have resulted in higher responsiveness.

Two previous intervention studies concerned with fat restriction addressed responsiveness. No other reference method was applied, but it was found that FFQ estimates of a change in fat intake were fairly similar to an estimate from a 4-day food record (12, 29). However, it cannot be determined whether the FFQs were responsive to real changes.

### Conclusions

Among older type 2 diabetic subjects, FFQ estimates of the intake of protein and other nutrients have good reproducibility. At the group level, an FFQ provides a reasonable estimate of protein intake and changes in intake. To a very moderate extent only, it
reflects actual individual protein intake, and the same applies to its responsiveness to changes in protein intake. These measurement properties should receive careful consideration when the FFQ is applied in diet-health studies.

ACKNOWLEDGMENTS

Financial support for this study was provided by the Dutch Organization for Scientific Research (The Hague, the Netherlands; grant 904-58-051) and the Health Research and Development Council (The Hague, the Netherlands; grant 28-2298).

Becel’s Nutritional Calculation Program (Van den Bergh Foods, Rotterdam, the Netherlands) was used to process food questionnaire data. Dietary counseling of patients was provided by dietitians Martine de Clercq, José Drijvers, Anneke Droop, Elles Grobbe, and Irene Havekes-Kleyweg.

REFERENCES

Validity of a Food Frequency Questionnaire


