

## Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults

Els Siebelink\*, Anouk Geelen and Jeanne H. M. de Vries

Division of Human Nutrition, Wageningen University, PO Box 8129, 6700 EV Wageningen, The Netherlands

(Received 5 July 2010 – Revised 1 December 2010 – Accepted 23 December 2010 – First published online 22 February 2011)

### Abstract

It is generally assumed that a FFQ is not suitable to estimate the absolute levels of individual energy intake. However, in epidemiological studies, reported nutrients by FFQ are often corrected for this intake. The objective of the present study was to assess how accurately participants report their energy intakes by FFQ. We compared reported energy intake with actual energy intake needed to maintain stable body weights during eleven controlled dietary trials. FFQ were developed to capture at least 90% of energy intake. Participants, 342 women and 174 men, with a mean BMI of 22.8 (SD 3.1) kg/m<sup>2</sup> filled out the FFQ just before the trials. Energy intakes during the trials were calculated from provided foods and reported free-food items, representing 90 and 10% of energy intake, respectively. Mean reported energy intake was 97.5 (SD 12.7)% of actual energy intake during the trials; it was 98.9 (SD 15.2)% for women and 94.7 (SD 16.3)% for men ( $P=0.004$  for difference between sexes). Correlation coefficients between reported and actual energy intakes were 0.82 for all participants, 0.74 for women and 0.80 for men. Individual reported energy intake as a percentage of actual intake ranged from 56.3 to 159.6% in women and from 43.8 to 151.0% in men. In conclusion, the FFQ appeared to be accurate for estimating the mean level of energy intakes of these participants and for ranking them according to their intake. However, the large differences found on the individual level may affect the results of epidemiological studies in an unknown direction if nutrients are corrected for energy intakes reported by FFQ.

**Key words:** Nutritional assessment: FFQ: Energy intake: Evaluation studies

The use of FFQ to assess habitual intake of populations is the subject of discussion, because they may lack the ability to capture full heterogeneity in reported intake. One of the major reasons is that respondents are not well able to summarise their food intake over a longer period<sup>(1)</sup>. Because of this measurement error, FFQ may be a less powerful instrument to detect relationships between diet and disease in epidemiological studies than food records or 24h recalls<sup>(2–4)</sup>. On the other hand, FFQ appear to be still an essential instrument for epidemiological studies, because they are the cheapest and most feasible method to assess long-term consumption. As FFQ can be applied in larger populations than recalls or records, their use improves the power of a study<sup>(5)</sup>. It is also shown that for some purposes, reports by FFQ are better than those by other instruments, and therefore some researchers recommended not to abandon FFQ but to combine them with open-ended methods<sup>(6)</sup>.

The performance of FFQ may be improved by adjusting reported nutrients for energy intake<sup>(7)</sup>. Also, to assess the independent effect of nutrients on disease, energy adjustment is needed<sup>(8,9)</sup>. However, it is questionable whether this can be done without introducing error. Validation studies have shown that energy intake by FFQ is misreported<sup>(8,10–12)</sup>. But because

of limitations in the studies, it is not clear how large this error is. An important limitation is that the time frame of the reference method does not match that of the FFQ. Often methods with a shorter reference period than the FFQ, such as 24h recalls, food records or biomarkers, are used for comparison. In addition, the number of subjects in many validation studies is limited.

At our division, we apply dietary assessment methods before starting a controlled dietary trial to estimate the required energy level of participants to maintain stable body weights during the trial. Until 1995, we have used a 3d estimated or weighed food record<sup>(13)</sup>. Since 1996, the FFQ has been used because the method is less time-consuming. As the controlled experimental diets had known energy contents and the trials lasted at least 3 weeks, we have information of an almost 'gold standard' reference method for reported habitual energy intake by the FFQ for participants with a stable body weight. By combining the results of eleven trials, we were able to evaluate self-reported intake by FFQ in 516 subjects. For this evaluation, we aimed to assess how accurately participants report their energy intake by FFQ for assessing absolute individual levels of energy intake and ranking of individuals according to energy intake.

\*Corresponding author: E. Siebelink, fax +31 317 483342, email [els.siebelink@wur.nl](mailto:els.siebelink@wur.nl)

**Table 1.** Overview of the dietary trials  
(Mean values and standard deviations)

Study	Year	Reference	Duration (weeks)	Nutrient of interest	Period of the year	Number of participants		Reported as percentage of actual intake	
						Total	Included in the present analysis	Mean	SD
1	1996	Alles <i>et al.</i> <sup>(14)</sup>	6	Oligosaccharides	Sept–Oct	40	38	99.0	13.4
2	1996	Brouwer <i>et al.</i> <sup>(16)</sup>	4	Folic acid	Oct–Dec	77	72	97.0	10.5
3	1997	Castenmiller <i>et al.</i> <sup>(17)</sup>	3	β-Carotene	Feb–Mar	72	58	96.4	13.9
4	1999	Weggemans <i>et al.</i> <sup>(23)</sup>	8	Cholesterol	Jan–Mar	55	42	96.8	16.1
5	1999	de Roos <i>et al.</i> <sup>(20)</sup>	8	Fatty acids	Mar–May	32	28	92.9	17.9
6	1999	de Roos <i>et al.</i> <sup>(19)</sup>	8	Fatty acids and carbohydrates	Oct–Nov	35	23	95.9	16.5
7	2001	de Jong <i>et al.</i> <sup>(18)</sup>	4	Folic acid	Oct–Nov	73	67	95.7	18.7
8	2003	Van Loo-Bouwman <i>et al.</i> <sup>(21)</sup>	6	β-Carotene	May	24	22	99.2	12.5
9	2005	Winkels <i>et al.</i> <sup>(24)</sup>	4	Folic acid	June	75	70	100.2	14.2
10	2007	Bos <i>et al.</i> <sup>(15)</sup>	10	Fatty acids	Jan–Apr	57	41	100.7	16.4
11	2007	Wanders <i>et al.</i> <sup>(22)</sup>	9	Fatty acids	Sept–Nov	64	55	97.2	20.4
Total						604	516	97.5	12.7

## Methods

### Study design

We compared energy intakes reported by FFQ with those required to maintain body weights during controlled dietary trials. We obtained data of subjects participating in eleven controlled dietary trials (Table 1). The present study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human participants were approved by the Ethics Committee of Wageningen University. Written informed consent was obtained from all participants. Other details and results of the trials have been described elsewhere<sup>(14–24)</sup>.

We asked participants to fill in a FFQ 2–6 weeks before entering the trial in order to estimate their energy needs. At the start of the trial, participants were allocated to a diet with an energy content according to their FFQ report. During the trials, body weights of the participants were kept stable. Their body weights were measured twice a week, and if their body weight decreased or increased by more than 0.2 kg between two measurements, the participants received a diet with a higher or lower energy content, respectively.

### Participants

Participants were recruited from the Wageningen University and from the population of the city of Wageningen and its surroundings. In total, 604 participants were included in the eleven trials. For the present analysis, we left out participants who withdrew before the end of the trial (*n* 11) and whose body weights changed more than 2.0 kg from the end of week 2 until the end of the trial (*n* 34). Of those who participated in more than one trial, only the data from the first trial were used (*n* 43). Among the remaining 516 volunteers, 342 were women and 174 were men (Table 2). Most volunteers were students or staff members of the Wageningen University. On average, they had a normal body weight, and 67% were younger than 25 years.

### Assessment of energy intake by FFQ

During a screening visit, the participants filled out a FFQ. Trained dietitians checked whether it was filled out properly, and if necessary, additional information was obtained about unusual or missing reports.

The FFQ were developed by selecting foods from the latest food consumption data of the Dutch National Food

**Table 2.** Characteristics of the 516 participants of the eleven dietary trials included in the present analysis

(Mean values and standard deviations)

	All ( <i>n</i> 516)		Women ( <i>n</i> 342)		Men ( <i>n</i> 174)	
	Mean	SD	Mean	SD	Mean	SD
Age (years)	27.8	12.9	26.8	12.4	29.6	13.8
Height (cm)	174.3	9.2	170.1	6.5	182.7	8.0
Weight at the start of the trial (kg)	69.4	10.9	66.4	9.8	76.1	9.8
Weight at day 14 of the trial (kg)	69.3	10.8	65.8	9.6	76.1	9.7
Weight at the end of the trial (kg)	69.1	10.7	65.6	9.5	76.0	9.7
BMI at the start of the trial (kg/m <sup>2</sup> )	22.8	3.1	22.8	3.3	22.8	2.6

Consumption Surveys<sup>(25–27)</sup>, which contributed >0.5% to the intake of fat, fatty acids and cholesterol, and covered at least 90% of energy intake. Thereafter, foods were added to achieve face validity. The reference period of the questionnaire was 4 weeks, and portion size questions were included for spreads, cheese, milk in coffee, gravy, candy bars and beer.

In trials 1–8, a FFQ with a 104-item food list in a table format was used. This FFQ was a modification of the Vet Express<sup>(28)</sup>, a FFQ developed at Wageningen University in 1992 to assess energy, total fat, fatty acids and cholesterol. This FFQ was updated several times using data from a more recent food consumption survey<sup>(25,26)</sup> and newer versions of the Dutch Nutrient Databases<sup>(29,30)</sup>. Also, the table format was abandoned, and the questions in the FFQ were asked according to a nested approach<sup>(31)</sup>. This resulted in a new questionnaire with 125 items, which was used for trial 9. For other studies at our division, this 125-item FFQ was extended with questions to enable the assessment of dietary fibre and specific micronutrients<sup>(32)</sup>. The result was a 183-item questionnaire, which was used in trials 10 and 11.

### Energy intake during the trials

Based on our previous study with food records<sup>(13)</sup>, we assumed that the participants needed the first 2 weeks of the trials to adapt to the experimental diets. Therefore, we defined actual required energy intake as the mean energy intake calculated from provided experimental diets from day 14 until the end of the trial.

The energy content of the experimental diets ranged from 7 to 20 MJ. In trials 1–3, the experimental diets were supplied at twenty-seven energy levels, in increments of 0.5 MJ (120 kcal), and in the other trials at fourteen energy levels, in increments of 1 MJ (239 kcal).

The experimental diets were composed with nutrient contents according to the specific demands of each trial, while the other nutrients met the RDA of the Dutch Health Council<sup>(33)</sup>. Actual energy intakes during the trials were calculated from all consumed foods and beverages using the most recent Dutch food composition database<sup>(29,30,34,35)</sup>. For calculations of the experimental diets, we used the Food Calculation System (BAS nutrition software 2004, Arnhem, The Netherlands) in which the most recent Dutch food composition database was included.

In general, the diets consisted of conventional foods, but in some trials, specific test foods were provided such as margarines with a specific fat composition or milk fortified with folic acid. A total of eighty-nine participants (17%), who were lacto-vegetarians or disliked some types of meat, received meat replacers, resulting in a diet with a similar nutrient composition to that of their non-vegetarian counterparts.

During weekdays at lunch time, the participants consumed their hot meal at the division. All other foods were supplied daily as a package and consumed at home. On Fridays, the participants received a package with foods and beverages for the breakfast, lunch and hot meals of the weekend plus instructions for the preparation of these foods. We provided about 90% of the total daily energy. The remaining 10% of

energy had to be chosen by the participants from a so-called free-food item list. For each dietary trial, this list was adapted to the specific demands of the trial (e.g. items low in fat or in  $\beta$ -carotene in trials 3 and 8, respectively). Furthermore, we allowed participants unrestricted consumption of non-energy foods such as coffee and tea without milk and sugar, water, herbs and spices, lemon juice, vinegar and non-energy soft drinks. The daily choice of free-food items and any deviations from the guidelines were recorded by each subject in a diary. As participants visited the division each weekday, it was possible to advise them about their diets on these days if necessary. Participants were urged not to change their physical activities or smoking habits and asked to record any change in lifestyle in their diary.

### Checks of body weight and diets

Body weights were measured twice a week before participants used their hot meal, with participants wearing no shoes or heavy clothing and with empty pockets. If body weight was changed more than 0.2 kg, energy intake was adjusted to a higher or lower energy level. On average, body weight decreased by 0.25 (SD 0.22) kg between days 1 and 7, and 0.17 (SD 0.17) kg between days 8 and 14. Between day 14 and the end of the trial, the participants lost their body weight on average 0.20 (SD 0.72) kg or 6.2 (SD 39.0) g/d. During this period of 1–8 weeks, eighty-one (16%) participants lost or gained between 1.0 and 2.0 kg body weight; for the other participants (84%), weight loss or gain was less than 1.0 kg.

Duplicate portions of the provided experimental diets were collected every day for an imaginary participant with a daily energy intake of 11 MJ, stored at  $-20^{\circ}\text{C}$  and analysed for protein and fat after the trials. Carbohydrates were calculated by difference (carbohydrates (g) = 100 – fat (g) – protein (g) – ash (g) – water (g)). Energy content was calculated from the macronutrient composition of the duplicate portions using Atwater factors<sup>(36)</sup> and combined with the calculated energy content of the free-choice items using the most recent Dutch food composition database. The daily energy content of the provided experimental diets according to chemical analysis of duplicate portions and calculated composition of the free-choice items was on average 10.6 MJ (2536 kcal) and lower than the *a priori* calculated mean energy content of 11 MJ.

The participants' diaries were checked regularly, and anonymous questionnaires on compliance were filled out in trials 2 and 10. Both did not reveal deviations from the protocol that might have affected the results.

### Statistical analysis

The reported average energy intakes and their 95% CI were computed from the FFQ reports. To show systematic differences, we plotted differences between the reported energy intake and the required energy intake against the average of the two methods in a so-called Bland–Altman plot. The difference in reported intakes expressed as percentage of actual

intake between men and women was tested by unpaired Student's *t* test. Because of a non-normal distribution of the biases, the difference in bias between men and women was tested by the Mann–Whitney test. To assess associations between reported and actual energy intakes, we used Pearson's correlation coefficients and applied Fisher's *Z* transformations to calculate 95% CI of the correlation coefficients. We classified the participants in those reporting <90%, between 90 and 110% regarded as accurate intakes and >110% of actual energy intake during the trials and tested the differences in BMI between these groups using one-way ANOVA and the *post hoc* Tukey test. Regression analysis was used to determine the relationships of sex, age, BMI, type of list (table format *v.* nested approach) and season to the difference between reported and actual intakes. All statistical tests were performed in SPSS for Windows version 15.0 (SPSS, Inc., Chicago, IL, USA).

## Results

The reported intake as a percentage of actual intake required to maintain a stable body weight ranged from 92.9 to 100.7% in the eleven trials (Table 1). Mean reported energy intake was significantly lower than actual energy intake for all participants and for men, but not for women (difference between sexes:  $P=0.004$ ; Table 3). As a consequence, the FFQ underestimated the difference in energy intake between sexes compared with the reference method. The difference between the reported energy intakes of men and women was 3.2 (95% CI 2.9, 3.5) MJ and that between their actual energy intakes was 3.8 (95% CI 3.0, 4.6) MJ.

The reported energy intake was highly correlated with the actual intake: Pearson's correlation coefficients were 0.82 (95% CI 0.80, 0.85) for all participants, 0.74 (95% CI 0.69, 0.78) for women and 0.80 (95% CI 0.74, 0.85) for men.

Some participants misreported energy intakes more than others. The individual reported energy intakes as a percentage of actual energy intakes showed a large variation and ranged from 56.3 to 159.6% for women and from 43.8 to 151.0% for men.

On the individual level, the Bland–Altman plot (Fig. 1) showed both over- and under-reporting of energy intake. The under and upper 95% limits of agreement varied from  $-3.3$  to  $3.8$  MJ. The plot also showed a general trend of under-reporting at lower intakes and over-reporting at higher intakes for both men and women.

Significant differences between reported and actual intakes were found for participants aged  $\leq 30$  years, with a BMI  $> 25$  kg/m<sup>2</sup>, and for those who reported in autumn or by a FFQ with a table format, but no statistically significant differences were found between strata. Subsequent analyses showed that reported energy intake was inversely associated with BMI for women ( $r = -0.13$ ;  $P=0.013$ ) and men ( $r = -0.26$ ;  $P=0.001$ ), and with age for men ( $r = -0.21$ ;  $P=0.006$ ) but not for women ( $r = -0.0002$ ;  $P=0.971$ ). In addition, mean BMI of those reporting  $> 110\%$  of actual energy intake was statistically significantly lower than in those reporting accurately (90–110%) or below 90% of actual intake (Table 4).

The correlation coefficients of the reported energy intake using the FFQ in a table format (trials 1–8:  $n$  350,  $r$  0.84) when compared with actual energy intake were similar to that using the two versions of the FFQ with questions

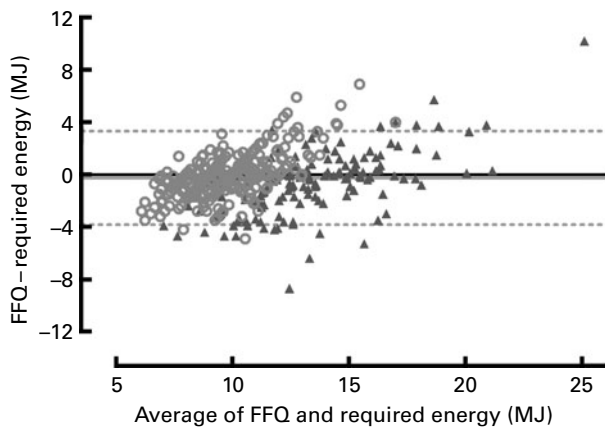
**Table 3.** Reported energy intake *v.* actual energy intake required to maintain body weight by characteristics of the 516 participants of the eleven dietary trials

(Mean values, standard deviations, 95% confidence intervals, number of subjects and percentages)

	<i>n</i>	%	Reported intake (MJ)		Actual intake (MJ)		Difference between reported and actual intake (MJ)			Reported as percentage of actual intake	
			Mean	SD	Mean	SD	Mean	95% CI	<i>P</i>	Mean	SD
All	516	100	10.7	3.2	10.9	2.5	-0.25	-0.41, -0.09	0.002	97.5	12.7
Sex											
Women	342	66.3	9.6	2.2	9.6	2.2	-0.07	-0.22, 0.96	0.421	98.9	15.2
Men	174	33.7	12.8	3.7	13.4	2.4	-0.61*	-0.95, -0.27	0.001	94.7	16.3
Age (years)†											
$\leq 30$	401	77.9	10.7	3.2	11.0	2.6	-0.30	-0.48, -0.12	0.001	97.1	15.1
$> 30$	113	21.9	10.7	3.1	10.8	2.4	-0.06	-0.40, 0.29	0.749	99.2	17.5
BMI (kg/m <sup>2</sup> )											
$\leq 25$	416	80.6	10.9	3.3	11.1	2.6	-0.18	-0.35, 0.00	0.055	98.2	15.5
$> 25$	100	19.4	9.8	2.6	10.4	2.1	-0.56	-0.90, -0.22	0.001	94.6	16.2
Season											
Spring	120	23.2	10.5	2.9	10.7	2.4	-0.19	-0.50, 0.13	0.248	98.3	15.0
Summer	93	18.0	11.0	3.4	11.2	2.8	-0.22	-0.63, 0.19	0.295	98.0	17.8
Autumn	162	31.4	10.7	2.7	11.0	2.7	-0.36	-0.65, -0.07	0.015	96.3	15.1
Winter	141	27.3	10.7	3.0	10.9	2.2	-0.20	-0.48, 0.09	0.182	97.8	15.4
FFQ											
Table format	350	67.8	10.8	3.3	11.1	2.6	-0.33	-0.52, -0.14	0.001	96.6	14.9
Nested	166	32.2	10.5	2.8	10.6	2.3	-0.08	-0.36, 0.20	0.577	99.3	17.0

\* Mean value was significantly different from the bias in women ( $P=0.004$ ).

† There are two missing values for age.



**Fig. 1.** Difference between the reported energy intake by FFQ and the actual energy intake required to maintain a stable body weight plotted against the mean of reported energy intake by FFQ and actual energy intake with 95% limits of agreement. ▲, Men (*n* 174); ○, women (*n* 342).

according to a nested approach (trial 9: *n* 70, *r* 0.82; trials 10–11: *n* 96, *r* 0.73).

Also, of the variables introduced into the regression model, sex, age and BMI contributed statistically significantly to the model, whereas the type of list and season (spring, summer, autumn or winter) in which the FFQ were filled out did not.

The regression analysis provided the following regression equation:

$$y = 1.6 + 0.59 \text{sex} + 0.02 \text{age} - 0.12 \text{BMI},$$

where *y* is the difference between the reported and actual energy intakes in MJ, sex is 0 for men and 1 for women, age is reported in years and BMI in kg/m<sup>2</sup>.

## Discussion

Energy intake reported by FFQ showed, on average, very good agreement with actual energy intake during controlled feeding trials, while body weights were kept stable. Also, FFQ were very well able to rank the participants according to energy intakes, but on the individual level, we found large differences. The present study provided a unique design to evaluate the reported energy intake of a large sample of participants using an almost 'gold standard' reference. Studies using references of this quality, including

doubly labelled water and indirect calorimetry, are often too expensive or too difficult to apply in a large sample. However, it may be questioned whether our reference method is truly a 'gold standard'.

The energy intakes assessed by the reference method might have been affected by changes in body weight or physical activity during the controlled trials and errors in the food composition table. To avoid the effects of changes in body weight, we used only the data of the participants after day 14 of the trials when they were used to the test diets and expected to have stable body weights. The average decrease in body weight from day 14 to the end of the study was only 6.2 (SD 39.0) g/d. This means that on average, the actual energy requirement of the participants was 0.2 MJ/d higher than the energy intake calculated from the provided diets, assuming that 1 kg of body weight equals 30 124.8 kJ (7200 kcal)<sup>(37)</sup>. In addition, we used the same food composition tables for calculating the energy intake from the FFQ as for calculating the actual energy intake during the trials. Therefore, errors in both methods originating from the food composition table are not independent. Assuming that systematic errors introduced by the food composition table in FFQ and reference method were the same, these errors would not have changed our conclusions for the mean estimated differences between the methods on the group and individual levels. Also, random errors in the food composition tables would not have affected these differences. However, because of these correlated errors, ranking of individuals according to their energy intake estimated from the FFQ may have agreed better with the reference method than with true intake, resulting in higher correlation coefficients. Yet, comparison of the energy content of the diets by chemical analysis for the energy level of 11 MJ only showed a small overestimation of, on average, 0.4 MJ for calculated energy intake during the trials. Although chemical analysis was only conducted for one average energy level of all test diets, we may expect the same difference for the other energy levels as diets were devised in a very standardised way. Therefore, we think that our reference method can be regarded as an almost 'gold standard' of energy intake.

Another limitation may have been a change in energy requirement between the period of filling out the FFQ and the trial because of a change in lifestyle. The FFQ were filled out about 2–6 weeks before the start of the trial.

**Table 4.** Sex distribution and BMI of the 516 participants classified according to the accuracy of their reports (Mean values, standard deviations and percentages)

Reported intake as percentage of actual energy requirements	<90				90–110				>110			
	Mean	SD	<i>n</i>	%	Mean	SD	<i>n</i>	%	Mean	SD	<i>n</i>	%
Number of participants			136	26			295	57			85	17
Women			81	24			198	58			63	18
Men			55	32			97	56			22	13
All	23.4 <sup>a</sup>	2.9			22.8 <sup>a</sup>	3.2			21.9 <sup>b</sup>	2.8		
Women	23.4 <sup>a</sup>	2.9			22.8 <sup>a</sup>	3.5			22.2 <sup>a</sup>	3.0		
Men	23.2 <sup>a</sup>	2.8			22.8 <sup>a</sup>	2.5			21.2 <sup>b</sup>	1.9		

<sup>a,b</sup> Mean values with unlike superscript letters were significantly different (*P* < 0.05).

However, we do not expect large differences, because of the short term between FFQ reports and the start of the trials and because participants were asked not to change their smoking habits and physical activity. According to the food diaries, the participants complied well with these guidelines.

We evaluated the three types of FFQ and information from different seasons in one analysis. We think this is justified because of a similar development of the FFQ and the fact that they all accounted for at least 90% of the energy intake. Also, we did not find differences in performance between different types of FFQ or FFQ applied in different seasons. However, FFQ including questions with a nested approach<sup>(31)</sup> may be easier to fill out, and longer FFQ may perform better than shorter FFQ<sup>(38)</sup>. In the present study, we could not confirm this. An explanation may be that the reports by the longer FFQ in other studies were reported by older participants in whom under-reporting was more common<sup>(39)</sup>. According to the literature<sup>(8,39)</sup>, accurate reporting of energy intake is influenced by several factors including sex, age, educational level, BMI, psychosocial factors and lifestyle. Inclusion of sex, age and BMI into our model confirmed these associations, even in our rather homogeneous population.

The FFQ in the present study were developed to cover an energy intake of at least 90% of actual energy intake. Theoretically, this implies that reports by our FFQ may underestimate energy intake by maximally 10%. Thus, if we would for this reason compare the FFQ reports with 90% of actual energy intake during the trials, the conclusion of our evaluation would have been that the FFQ, on average, slightly overestimate energy intake.

It may have been expected that our participants would yield good reports of their food consumption, because they were young and mostly highly educated<sup>(40)</sup>. They were motivated to enter a controlled dietary trial, to fill out a FFQ and to be aware that the purpose of the FFQ was to estimate their required energy intake during the trial. It is not unthinkable that some over-reported their consumption because they were afraid to receive too little food during the trial. Thus, our FFQ may not provide similar good results in other studies or populations as in the present study.

Reports of energy intakes by the FFQ during our study showed better results both on the group mean level and to rank participants according to their intake than other studies. In other studies, underestimation on the group level ranged from 10 to 36%. Andersen *et al.*<sup>(10)</sup> reported an underestimation of 11% in a group of seventeen women with a comparable age with that of our population. Subar *et al.*<sup>(7)</sup> found an underestimation of 36% for women (*n* 206) and 34% for men (*n* 245) in the age range of 40–69 years. Kroke *et al.*<sup>(11)</sup> found an underestimation of 19% for a group of twenty-eight women and men aged between 35 and 67 years. Our participants were on average much younger, but if we compare the results of a similar age subgroup of our population, the difference between reported intake and the reference level is also much smaller than in other studies. We found a mean overestimation of 1.9% for women (*n* 57) and 4.6% underestimation for men (*n* 37) in the participants of  $\geq 40$  years (results not shown). In a subgroup

of our population (*n* 17) aged 65–86 years with an average BMI of 24.5 kg/m<sup>2</sup>, the underestimation was somewhat larger with 13% (1.5 MJ) but still in the lower range compared with other studies<sup>(41)</sup>. When evaluating ranking of participants, we found correlation coefficients of 0.74 for women and 0.80 for men, whereas Subar *et al.*<sup>(7)</sup> reported correlation coefficients of 0.10 (*n* 206) for women and 0.19 (*n* 245) for men. Even in the older men (*n* 17), the correlation coefficient between the reported energy intake and the actual energy intake of 0.67 can still be considered as reasonably good compared with other studies<sup>(41)</sup>.

Although the agreement for ranking individuals was good, the differences found on the individual level ranging from –50 to +50% between reported and actual energy intakes were large but similar to those of other studies<sup>(10,11)</sup>. As we define reported energy intakes on the individual level within  $\pm 10\%$  of actual energy intakes as acceptable, only 57% of the participants reported within that range. Men had higher requirements and wider ranges in misreporting than women. The bias appeared to be intake-related, with under-reporting at lower intakes and over-reporting at higher intakes for both men and women.

Another explanation for the better performance of the FFQ compared with other FFQ for assessing mean group intake and ranking individuals to their intake may be their shorter reference period. It was 1 month, whereas many other FFQ use 1 year. In general, people find it hard to report their intake over a long period, taking all seasonal variation into account. Although for energy intake a month is expected to be sufficient<sup>(42)</sup>, this may be different for nutrients or foods with a larger variation. On the other hand, we used a relatively long reference period for our reference method (1–8 weeks), whereas other studies had a maximum reference duration of 14 d using the doubly labelled water method<sup>(10–12)</sup>.

It is surprising that our FFQ perform better on the group level and for ranking of individuals than other FFQ, but that their performance on the individual level is as inaccurate. A better performance on the group level may be explained by the fact that in our selected population, under- and overestimation occurred to the same extent resulting in, on average, a small difference. The better ranking that we found may be, but only for a small part, explained by correlated errors due to the use of the same food composition table to calculate energy intake for both the FFQ and the reference method.

Thus, reported intake by our FFQ, on average, equals the actual energy intake to maintain body weight and ranks the participants reasonably well according to their energy intake but is not accurate at the individual level. For adjustment of energy intake when studying nutrient–disease relationships<sup>(9)</sup>, it is required that FFQ accurately determine absolute energy intakes. Data from the Observing Protein and Energy Nutrition study<sup>(1)</sup> showed a failure of FFQ to provide a sufficiently accurate report of absolute energy intakes to enable the detection of their moderate associations with disease. Yet, it was also shown that because of correlated errors in reporting protein and energy, energy-adjusted protein was less affected by measurement error than absolute protein intake. The results of our validation could suggest that adjustment of nutrients

to energy intake reported by FFQ may result in the introduction of substantial error in epidemiological studies. However, in case of correlated errors between the nutrient of interest and energy, error might be reduced by adjusting for energy.

We conclude that despite the large differences in accuracy between individuals, the FFQ used in the present study can be useful to pick up dietary changes in trials if the population is large enough, because systematic errors are only small on the group level. In addition, our FFQ can be applied in epidemiological studies to rank individuals accurately according to their energy intake, but if nutrient intakes are adjusted for energy as reported by FFQ, this may affect the results of these studies in an unknown direction.

### Acknowledgements

E. S. and J. H. M. d. V. designed the study; E. S. collected and analysed the data, and wrote the manuscript; A. G. supervised the data analysis, and together with J. H. M. d. V. revised the earlier versions of the manuscript. None of the authors had a personal or financial conflict of interest. The present study received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

### References

- Kipnis V, Subar AF, Midthune D, *et al.* (2003) Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* **158**, 14–21.
- Bingham SA, Luben R, Welch A, *et al.* (2003) Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet* **362**, 212–214.
- Freedman LS, Potischman N, Kipnis V, *et al.* (2006) A comparison of two dietary instruments for evaluating the fat-breast cancer relationship. *Int J Epidemiol* **35**, 1011–1021.
- Schatzkin A, Kipnis V, Carroll RJ, *et al.* (2003) A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. *Int J Epidemiol* **32**, 1054–1062.
- Thompson FE, Kipnis V, Midthune D, *et al.* (2008) Performance of a food-frequency questionnaire in the US NIH-AARP (National Institutes of Health-American Association of Retired Persons) Diet and Health Study. *Public Health Nutr* **11**, 183–195.
- Lissner L, Troiano RP, Midthune D, *et al.* (2007) OPEN about obesity: recovery biomarkers, dietary reporting errors and BMI. *Int J Obes* **31**, 956–961.
- Subar AF, Thompson FE, Kipnis V, *et al.* (2001) Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *Am J Epidemiol* **154**, 1089–1099.
- Livingstone MB & Black AE (2003) Markers of the validity of reported energy intake. *J Nutr* **133**, Suppl. 3, 895S–920S.
- Willett WC, Howe GR & Kushi LH (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 1220S–1228S; discussion 1229S–1231S.
- Andersen LF, Tomten H, Haggarty P, *et al.* (2003) Validation of energy intake estimated from a food frequency questionnaire: a doubly labelled water study. *Eur J Clin Nutr* **57**, 279–284.
- Kroke A, Klipstein-Grobusch K, Voss S, *et al.* (1999) Validation of a 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, and repeated 24-h dietary recall methods. *Am J Clin Nutr* **70**, 439–447.
- Subar AF, Kipnis V, Troiano RP, *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* **158**, 1–13.
- de Vries JH, Zock PL, Mensink RP, *et al.* (1994) Underestimation of energy intake by 3-d records compared with energy intake to maintain body weight in 269 nonobese adults. *Am J Clin Nutr* **60**, 855–860.
- Alles MS, Hartemink R, Meyboom S, *et al.* (1999) Effect of transgalactooligosaccharides on the composition of the human intestinal microflora and on putative risk markers for colon cancer. *Am J Clin Nutr* **69**, 980–991.
- Bos MB, de Vries JH, Feskens EJ, *et al.* (2010) Effect of a high monounsaturated fatty acids diet and a Mediterranean diet on serum lipids and insulin sensitivity in adults with mild abdominal obesity. *Nutr Metab Cardiovasc Dis* **20**, 591–598.
- Brouwer IA, van Dusseldorp M, West CE, *et al.* (1999) Dietary folate from vegetables and citrus fruit decreases plasma homocysteine concentrations in humans in a dietary controlled trial. *J Nutr* **129**, 1135–1139.
- Castenmiller JJ, van de Poll CJ, West CE, *et al.* (2000) Bioavailability of folate from processed spinach in humans. Effect of food matrix and interaction with carotenoids. *Ann Nutr Metab* **44**, 163–169.
- de Jong RJ, Verwei M, West CE, *et al.* (2005) Bioavailability of folic acid from fortified pasteurised and UHT-treated milk in humans. *Eur J Clin Nutr* **59**, 906–913.
- de Roos NM, Bots ML, Siebelink E, *et al.* (2001) Flow-mediated vasodilation is not impaired when HDL-cholesterol is lowered by substituting carbohydrates for monounsaturated fat. *Br J Nutr* **86**, 181–188.
- de Roos NM, Siebelink E, Bots ML, *et al.* (2002) *Trans* monounsaturated fatty acids and saturated fatty acids have similar effects on postprandial flow-mediated vasodilation. *Eur J Clin Nutr* **56**, 674–679.
- Van Loo-Bouwman CA, West CE, van Breemen RB, *et al.* (2009) Vitamin A equivalency of beta-carotene in healthy adults: limitation of the extrinsic dual-isotope dilution technique to measure matrix effect. *Br J Nutr* **101**, 1837–1845.
- Wanders AJ, Brouwer IA, Siebelink E, *et al.* (2010) Effect of a high intake of conjugated linoleic acid on lipoprotein levels in healthy human subjects. *PLoS One* **5**, e9000.
- Weggemans RM, Zock PL, Meyboom S, *et al.* (2000) Apolipoprotein A4-1/2 polymorphism and response of serum lipids to dietary cholesterol in humans. *J Lipid Res* **41**, 1623–1628.
- Winkels RM, Brouwer IA, Siebelink E, *et al.* (2007) Bioavailability of food folates is 80% of that of folic acid. *Am J Clin Nutr* **85**, 465–473.
- Netherlands Nutrition Centre (1998) *Zo eet Nederland 1998 (in Dutch)*. (Dutch National Food Consumption Survey 1998). The Hague: Netherlands Nutrition Centre.
- Netherlands Nutrition Centre (1993) *Zo eet Nederland 1992 (in Dutch)*. (Dutch National Food Consumption Survey 1992). The Hague: Netherlands Nutrition Centre.
- Hulshof KFAM & Van Staveren WA (1991) The Dutch National Food Consumption Survey: design, methods and first results. *Food Policy* **16**, 257–260.
- Feunekes GI, Van Staveren WA, De Vries JH, *et al.* (1993) Relative and biomarker-based validity of a food-frequency

- questionnaire estimating intake of fats and cholesterol. *Am J Clin Nutr* **58**, 489–496.
29. Nevo (1993) *Dutch Food Composition Database 1993*. The Hague: Stichting Nevo.
  30. Nevo (2001) *Dutch Food Composition Database 2001*. The Hague: Stichting Nevo.
  31. Subar AF, Thompson FE, Smith AF, *et al.* (1995) Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. *J Am Diet Assoc* **95**, 781–788; quiz 789–790.
  32. Verkleij-Hagoort AC, de Vries JHM, Stegers MPG, *et al.* (2007) Validation of the assessment of folate and vitamin B<sub>12</sub> intake in women of reproductive age: the method of triads. *Eur J Clin Nutr* **61**, 610–615.
  33. Health Council of the Netherlands (2006) *Guidelines for a Healthy Diet*. The Hague: Health Council of the Netherlands.
  34. Nevo (1996) *Dutch Food Composition Database 1996*. The Hague: Stichting Nevo.
  35. Nevo (2006) *Dutch Food Composition Database 2006*. The Hague: Stichting Nevo.
  36. Southgate DA & Durnin JV (1970) Calorie conversion factors. An experimental reassessment of the factors used in the calculation of the energy value of human diets. *Br J Nutr* **24**, 517–535.
  37. Shils ME, Shile M, Ross AC, *et al.* (2006) *Modern Nutrition in Health and Disease*, 10th ed., p. 982. Philadelphia, PA: Lippincott Williams & Wilkins.
  38. Molag ML, de Vries JH, Ocke MC, *et al.* (2007) Design characteristics of food frequency questionnaires in relation to their validity. *Am J Epidemiol* **166**, 1468–1478.
  39. Tooze JA, Subar AF, Thompson FE, *et al.* (2004) Psychosocial predictors of energy underreporting in a large doubly labeled water study. *Am J Clin Nutr* **79**, 795–804.
  40. Black AE, Prentice AM, Goldberg GR, *et al.* (1993) Measurements of total energy expenditure provide insights into the validity of dietary measurements of energy intake. *J Am Diet Assoc* **93**, 572–579.
  41. de Vries JH, de Groot LC & van Staveren WA (2009) Dietary assessment in elderly people: experiences gained from studies in the Netherlands. *Eur J Clin Nutr* **63**, Suppl. 1, S69–S74.
  42. Margetts BM, Thompson RL, Key T, *et al.* (1995) Development of a scoring system to judge the scientific quality of information from case-control and cohort studies of nutrition and disease. *Nutr Cancer* **24**, 231–239.