Relative validity of a semi-quantitative, web-based FFQ used in the 'Snart Forældre' cohort – a Danish study of diet and fertility

Vibeke K Knudsen[1],*, Elizabeth E Hatch[2], Heidi Cueto[3], Katherine L Tucker[4], Lauren Wise[2], Tue Christensen[1] and Ellen M Mikkelsen[3]

[1]Division of Nutrition, National Food Institute, Technical University of Denmark, Mørkhøj Bygade 19, 2860 Søborg, Denmark: [2]Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA: [3]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark: [4]Department of Clinical Laboratory & Nutritional Sciences, University of Massachusetts Lowell, Lowell, MA, USA

Abstract

Objective: To assess the relative validity of a semi-quantitative, web-based FFQ completed by female pregnancy planners in the Danish 'Snart Forældre' study.
Design: We validated a web-based FFQ based on the FFQ used in the Danish National Birth Cohort against a 4 d food diary (FD) and assessed the relative validity of intakes of foods and nutrients. We compared means and medians of intakes, and calculated Pearson correlation coefficients and de-attenuated coefficients to assess agreement between the two methods. We also calculated the proportion correctly classified based on the same or adjacent quintile of intake and the proportion of grossly misclassified (extreme quintiles).
Setting: Participants (*n* 128) in the 'Snart Forældre' study who had completed the web-based FFQ were invited to participate in the validation study.
Subjects: Participants in the 'Snart Forældre' study, in total ninety-seven women aged 20–42 years.
Results: Reported intakes of dairy products, vegetables and potatoes were higher in the FFQ compared with the FD, whereas reported intakes of fruit, meat, sugar and beverages were lower in the FFQ than in the FD. Overall the de-attenuated correlation coefficients were acceptable, ranging from 0·33 for energy to 0·93 for vitamin D. The majority of the women were classified in the same or adjacent quintile and few women were misclassified (extreme quintiles).
Conclusion: The web-based FFQ performs well for ranking women of reproductive age according to high or low intake of foods and nutrients and, thus, provides a solid basis for investigating associations between diet and fertility.

Infertility is a common condition that affects approximately 12–15 % of all couples who attempt conception[1]. Identifying potentially modifiable factors that can improve fertility is an important public health goal, especially as there is a trend of delayed childbearing among Danish women[2,3]. Diet is a modifiable factor that may play an important role in fertility, potentially through effects on hormones[4–7], menstrual cycle patterns[8,9] and systemic inflammation[10]. A series of studies using data from the Nurses' Health Study has suggested associations of protein source[11], Fe[12], dairy products[13], *trans*-fatty acids[14] and B-vitamins[15] with ovulatory infertility. There are few studies on the role of diet on fecundability, the per cycle probability of conception, using a prospective cohort design[16–18]. The 'Snart Forældre' ('Soon Parents') study, initiated in 2011, is an extension of 'Snart Gravid' ('Soon Pregnant'), a prospective cohort study of

Danish pregnancy planners (2007–2011)[19–21]. A central aim of the 'Snart Forældre' study is to investigate the association between pre-pregnancy diet and fecundability. Unlike most previous studies of diet and fertility[11–15], our study will be able to assess the role of the woman's diet on a wide spectrum of fertility problems by measuring pre-conception diet in relation to the monthly probability of conception.

FFQ are widely used in epidemiological studies because they are less expensive than other dietary methods and easy to administer in large cohort studies. FFQ are often used in studies that assess diet–disease associations[22] to rank individuals according to high and low intake of foods and nutrients. Furthermore, the FFQ is easily transformed into a web-based format, which is an advantage in Internet-based studies like 'Snart Forældre', as the web-based data collection facilitates data processing and

Corresponding author: Email vkkn@food.dtu.dk

nutrient calculation in a large study population. The widespread use of the Internet (in 2012, 92 % of all 16–74-year-old citizens in Denmark had Internet access at home[23]) thus enables web-based studies.

To assess dietary intake among Danish women planning a pregnancy, we developed a web-based FFQ based on the FFQ used in the Danish National Birth Cohort study[24], updated for our target population of reproductive age women. The aim of the present study was to assess the relative validity of the web-based FFQ using a 4 d food diary as reference.

Materials and methods

'Snart Forældre'

The 'Snart Forældre' study is an Internet-based prospective cohort study of Danish pregnancy planners. Enrolment began in August 2011 and eligible women are 18–45 years old, in a relationship with a male partner, attempting to conceive and not using fertility treatment. Potential participants are recruited via a coordinated media strategy and enrolment is carried out via the study website. Using online questionnaires, data on health and lifestyle as well as demographic factors are collected at baseline and bimonthly for 12 months or until recognized conception. Participants are asked to complete the FFQ ten days after enrolment. The FFQ was implemented in September 2012; by 15 September 2014, a total of 1156 of 1402 'Snart Forældre' participants (82·5 %) had completed the FFQ.

Development of the FFQ

The semi-quantitative, web-based FFQ was based on the paper-based FFQ applied in the Danish National Birth Cohort[24] and the 'Diet, Cancer and Health' study[25], and was updated using information on usual dietary intake in Danish women aged 18–49 years[26] to capture intakes of foods and nutrients in this age group. In the web-based FFQ, participants are asked to record their usual intakes of foods and drinks in the previous year. The FFQ included approximately 220 foods and beverages. In addition, seven photographic series with varying portion sizes of foods and dishes typical in a Danish diet are included to help participants assess portion sizes. In the development phase, the FFQ was pilot-tested, resulting in a few alterations to the FFQ; some questions were reformulated for clarity, response categories were added and additional guidance on how to complete the FFQ was included. Feasibility was a main objective in the development of the web-based FFQ. Thus, the FFQ includes help buttons explaining portion sizes and describing additional examples of food items and dishes. Skip patterns are implemented to shorten the length of the questionnaire. To minimize missing values and errors in data entry, the web-based FFQ includes radio buttons, check boxes and drop-down menus for response options. In addition, a continuously updated progress bar at the bottom of each page illustrates the number of remaining questions. The FFQ does not record use of dietary supplements, as these are recorded in the baseline and follow-up questionnaires in the 'Snart Forældre' study. Nutrient contributions from supplements were therefore not included in the analyses.

Food diary

We validated the FFQ against a self-administered 4 d food diary (FD) containing pre-coded lines of common Danish dishes, foods and beverages, which is applied in the Danish National Survey of Diet and Physical Activity (DANSDA) and has been validated against objective measures and a weighed food record[27,28]. The FD is structured according to a typical Danish diet, with breakfast, lunch, dinner and three snacks, each with pre-coded lines of the most commonly eaten foods and beverages. Portion sizes are estimated using household measures (cups, glasses, slices, etc.) and twelve series of photographs with varying portion sizes of commonly eaten foods.

Validation study

We aimed to recruit 100 participants for the validation study. Eligible participants from the main study were consecutively invited to participate. In total, 128 women were invited in the period 17 August to 20 November 2013; from these, 100 (78 %) women agreed to participate and ninety-seven (76 %) completed the FD. Within 1–3 d after completion of the online FFQ the participants received the pre-coded food diaries, including instructions on how to complete them, the portion size photo series and prepaid return envelopes, by postal mail. In the instructions, participants were advised to report to the diary immediately after each meal throughout the 4 d. By random assignment, we asked the participants to complete the food diaries starting on the first coming Wednesday or first coming Sunday after they received the diaries, thereby including three weekdays and one weekend day. The FD were manually checked for completeness and scanned, and then intakes of foods and nutrients for each participant were calculated using GIES (General Intake Estimation System) software, developed at the National Food Institute, including standard recipes, information on portion sizes and data from the Danish Food Composition Tables (www.foodcomp.dk). This method has been described in detail elsewhere[29]. The data from the FFQ were cleaned and processed, all frequencies were converted to frequencies per day, and intakes of foods and nutrients were calculated using standard recipes, portion sizes and information from the Danish nutrient database (www.foodcomp.dk).

Statistical methods

Based on prior validation studies, participants with very low/very high energy intakes, i.e. below 2·5 MJ/d or above

25 MJ/d (<600 kcal or >6000 kcal), in either of the two diet assessment methods are generally excluded from analysis. However, none of the participants had energy intakes outside these limits and no dietary data were excluded from the present analyses. Intakes of foods and nutrients from both the FFQ and the FD were adjusted for total energy intake using the residual method[30].

We compared characteristics of the study population and the remaining 'Snart Forældre' cohort using means, medians and proportions. We compared median and mean food and nutrient intakes, as assessed by the two methods, using the paired t test when normally distributed (nutrients) and the Wilcoxon signed-rank test when non-normally distributed (foods). Agreement between the FFQ and the FD was assessed by Bland–Altman plots. Correlations between intakes were assessed by the Pearson correlation coefficient for energy-adjusted nutrients and foods. As day-to-day variation in pre-coded FD may attenuate the correlations between the mean intakes from the FD and intakes from the FFQ, de-attenuated coefficients, r_t, were calculated using the formula:

$$r_t = r_o \sqrt{(1 + intra_x / inter_x / n_x)},$$

where r_o is the observed correlation coefficient, $intra_x$ is the within-person variance for intakes estimated in the FD, $inter_x$ is the between-person variance and n_x is the number of days of diet registration[31]. Furthermore, the participants were divided into quintiles according to intake of foods and nutrients, and agreement between the two methods was assessed using cross-classification. Proportions of individuals who were categorized into the same or adjacent quintile were calculated, and the gross misclassification was defined as the percentage of individuals categorized into opposite lowest $v.$ highest quintiles across the two methods.

Results

A total of ninety-seven women completed both dietary assessment methods. Mean duration of completion of the FFQ was 29 min. The mean age of participants and their partners was 29 years and 31 years, respectively, in both the validation group and the 'Snart Forældre' group (Table 1). Participants in both groups were similar according to level of education, number of units of alcohol consumed per week, pack-years of smoking and most recent contraceptive used at study entry. Participants in the validation study had a lower mean BMI compared with 'Snart Forældre' participants (23·0 $v.$ 24·5 kg/m²) and a similar level of physical activity (median MET-h/week: 35·4 $v.$ 36·2, where MET = metabolic equivalent of task).

The Bland–Altman plots for total energy intake and vegetable intake revealed some extreme values, but no systematic patterns were observed (Figs 1 and 2). Estimated intakes of total energy and macronutrients were generally higher when assessed by the FD compared with the FFQ

(Table 2), but only small differences were seen in the distribution of macronutrients, i.e. percentage of energy from carbohydrate, protein or fat, respectively. As the 95 % CI overlapped, no statistically significant differences in mean intakes were detected for most of the micronutrients. Exceptions included vitamin C, where intake was higher for the FFQ; and Ca, P, Mg, total Fe, Na and Zn, where higher intakes were observed for the FD compared with the FFQ. Pearson correlation coefficients ranged from 0·08 (Na) to 0·63 (carbohydrate/fibre, g/d); the de-attenuated correlation coefficients ranged from 0·13 (Na) to 0·93 (vitamin D; Table 2). Cross-classification analysis showed that added sugar, fibre, riboflavin and carbohydrate were most often correctly classified, with 79·4, 82·5, 79·4 and 76·3 % of participants, respectively, classified in the same or adjacent quintile (Table 3).

Median intakes of dairy products, vegetables and potatoes were higher in the FFQ assessment, while fruit, meat, sugar and beverages were lower in the FFQ, compared with the FD (Table 4). Pearson correlation coefficients for energy-adjusted foods ranged from 0·17 for fats (butter, spread, oil, etc.) to 0·61 for low-fat dairy products (Table 4). De-attenuated coefficients ranged from 0·25 (fats) to 0·75 (fish). Low-fat dairy products, cereals and beverages had the highest proportions of individuals correctly classified, with 80·4, 72·2 and 71·1 %, respectively, classified in the same or adjacent quintile (Table 5).

Discussion

The present validation study of a newly developed web-based FFQ among Danish women of reproductive age

Table 1 Characteristics of participants in the validation study (n 97) and participants in the 'Snart Forældre' cohort (n 734)

Characteristic	Validation study	'Snart Forældre'
Number of women	97	637
Age (years), mean	28·9	28·9
Parous (%)	43·3	40·9
Partner's age (years), mean	31·4	31·2
Height (cm), mean	169·2	168·5
Weight (kg), median	63·0	66·0
IQR	57–72	60–76
BMI (kg/m²), mean	23·0	24·5
Physical activity (MET-h/week), median	35·4	36·2
IQR	21·3–65·1	17·3–77·3
Years of education*,† (%)		
Short (<3 years)	16·5	24·9
Medium (3–4 years)	47·4	40·0
Long (>4 years)	36·1	33·1
Pack-years of ever smoking, mean	0·7	1·0
Alcohol (drinks/week), mean	3·1	2·7
Diabetes, yes (%)	0·0	0·3
OC as last contraception, yes (%)	55·7	53·1

IQR, interquartile range; MET, metabolic equivalent of task; OC, oral contraceptive.
*Years of education after 9 years of compulsory schooling.
†Missing data for thirteen (1·8 %) participants in the 'Snart Forældre' group.
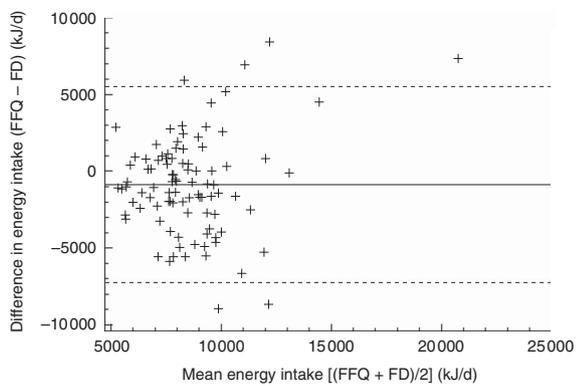
Fig. 1 Bland–Altman plot assessing the relative validity of the newly developed, semi-quantitative web-based FFQ among participants (*n* 97; women aged 20–42 years) from the 'Snart Forældre' study, Denmark, 2013. The difference in energy intake between the FFQ and the 4 d food dairy (FD) is plotted *v.* the mean energy intake from the two methods. ───── represents the mean difference (bias) and – – – – – represent the limits of agreement
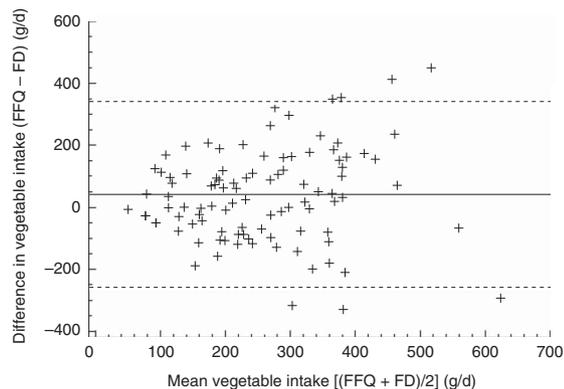

Fig. 2 Bland–Altman plot assessing the relative validity of the newly developed, semi-quantitative, web-based FFQ among participants (*n* 97; women aged 20–42 years) from the 'Snart Forældre' study, Denmark, 2013. The difference in vegetable intake between the FFQ and the 4 d food dairy (FD) is plotted *v.* the mean vegetable intake from the two methods. ───── represents the mean difference (bias) and – – – – – represent the limits of agreement

performed well relative to the reference method, a 4 d FD. Correlation coefficients for most of the foods and nutrients were in the range of what has previously been reported in validation studies of FFQ[32–34]. Direct comparisons of mean intakes of foods and nutrients across the two methods showed some discrepancies for selected nutrients, but the aim of the FFQ is to rank participants according to high *v.* low intake of the foods and nutrients under study. However, the overall agreement of the two methods was acceptable.

Total energy intake differed between the two methods, with higher energy intake in the FD relative to the FFQ. Although total energy intake may have been slightly underestimated in the FFQ, good correlations for

macronutrients were observed (0·70, 0·56 and 0·63 for carbohydrate, protein and fat, respectively), indicating that potential under-reporting of energy may not be specific to certain foods or food groups. Also the macronutrient distribution showed good agreement with data for energy distributions in the Danish population[26]. This, together with relatively strong correlations, ensures a solid basis for testing hypotheses of associations between diet and fertility. FFQ may overestimate energy intake compared with food records, possibly as a consequence of a high number of questions on intake. In the present questionnaire, questions at the food group level were used to adjust the questions on more detailed intake (e.g. reported overall fruit intake was used to adjust the total intake of apples, pears, bananas, etc.). Furthermore, as the current FFQ used electronic skip patterns, participants were not asked irrelevant questions and this may have reduced over-reporting.

Food and nutrient intakes were adjusted for total energy intake. Adjusting for total energy is often done in epidemiological studies of the association between diet and disease, and therefore it is appropriate to include energy adjustment in validation studies too. Furthermore, adjusting for total energy intake has been shown to bring the other nutrients into alignment. Among micronutrients, highest correlation coefficients were seen for vitamin D, Mg and Se. High correlation coefficients between FFQ and 24 h recall have been reported for vitamin D by Elorriaga *et al.*[33] and Streppel *et al.*[34]. Toft *et al.*[32] and Johansson *et al.*[35] found correlations for saturated fat of the same magnitude as in the present study when comparing FFQ and 24 h recalls. Among the micronutrients we evaluated, the lowest correlations were for vitamin A ($r = 0·24$) and Na ($r = 0·13$). Vitamin A has been poorly correlated across methods in other studies[35–37], likely due to the large variability in intake of food sources rich in vitamin A[35]. Because of the high day-to-day variability, habitual intake may not be well captured in the 4 d food record[38]. Intake of Na is known to be poorly assessed by most dietary assessment methods[39–41]. Added salt in cooking or at the table was not recorded in either method and Na intake, as assessed here, originates only from processed foods and dishes in the nutrient database. Because food items that are similar in contents of other nutrients can vary greatly in salt content, there can be large differences in salt intake assessed by a food record and a long-term measure of usual intake (FFQ).

Among food groups, the strongest correlations were seen for intakes of low-fat dairy products, meat and beverages. Streppel *et al.*[34], Paalanen *et al.*[37] and Macedo-Ojeda *et al.*[42] also found good correlations for dairy products and meat, while in contrast Streppel *et al.* reported poorer correlation for intake of beverages[34]. The validity of intake of food groups is not as easily compared with results from other studies as is validity of nutrients, as food groups are classified in various ways. However, fruit, vegetables and

Table 2 Mean daily intakes of macro- and micronutrients* from the 4 d food diary (FD) and the newly developed, semi-quantitative, web-based FFQ; relative differences and 95 % CI for the differences between the FD and FFQ; and Pearson correlation coefficients and de-attenuated coefficients between intakes from the FD and FFQ. Participants (n 97; women aged 20–42 years) from the 'Snart Forældre' study, Denmark, 2013

	FD	FFQ	Difference	Relative difference (%)	95 % CI for the difference	Pearson correlation coefficient	De-attenuated coefficient
Energy (kJ)	8986	8129	857	9·5	212, 1502	0·29	0·33
Carbohydrate (g)	244	224	20	8·5	13·7, 27·8	0·63	0·70
Carbohydrate (E%)	44	47	−3	−7·1	−4·6, −1·7	0·60	0·67
Protein (g)	84	76	8·0	9·6	5·2, 10·9	0·49	0·56
Protein (E%)	16·6	16·9	−0·3	−1·8	−1·0, 0·4	0·49	0·55
Fat (g)	86·3	73·6	12·7	14·8	9·9, 15·6	0·56	0·63
Fat (E%)	38·3	36·7	1·5	4·0	0·2, 2·8	0·59	0·65
Saturated fat (g)	32·2	29·0	3·2	9·9	1·9, 4·5	0·51	0·61
Monounsaturated fat (g)	33·1	27·3	5·8	17·5	4·2, 7·4	0·52	0·59
Polyunsaturated fat (g)	14·2	11·5	2·7	19·1	2·0, 3·4	0·41	0·49
Added sugar (g)	54·6	36·6	17·9	32·8	11·0, 24·8	0·41	0·47
Fibre (g)	20·9	22·0	−1·1	−5·2	−2·1, −0·0	0·63	0·70
Vitamin A (RE)	1082	1091	−8·8	−0·8	−135, 118	0·17	0·24
Vitamin D (µg)	4·5	3·8	0·7	14·7	−0·1, 1·4	0·50	0·93
Vitamin E (mg)	9·6	9·1	0·6	5·9	−0·2, 1·4	0·43	0·49
Thiamin (mg)	1·2	1·2	0·0	2·1	−0·0, 0·1	0·32	0·41
Riboflavin (mg)	1·7	1·7	−0·0	−0·3	−0·1, 0·1	0·51	0·58
Vitamin B_6 (mg)	1·5	1·5	−0·1	−3·5	−0·1, 0·0	0·40	0·48
Folate (µg)	338·5	363·2	−24·7	−7·3	−46·3, −3·0	0·43	0·49
Vitamin B_{12} (µg)	5·7	5·6	0·1	1·5	−0·3, 0·5	0·41	0·56
Vitamin C (mg)	110·9	131·8	−20·9	−18·8	−31·9, −9·8	0·36	0·44
Ca (mg)	1135·4	1066·4	69·0	6·1	0·6, 137·4	0·41	0·45
P (mg)	1487·8	1442·1	45·7	3·1	−9·1, 100·5	0·52	0·57
Mg (mg)	343·7	332·3	11·3	3·3	−0·2, 22·9	0·61	0·69
Na (mg)	3573·4	2783·4	790·0	22·1	629·8, 950·2	0·08	0·13
K (mg)	3093	3268	− 174	−5·6	−296, −52	0·50	0·58
Total Fe (mg)	10·5	10·1	0·4	4·0	0·1, 0·7	0·48	0·58
Non-haem Fe (mg)	9·6	9·3	0·3	1·5	0, 0·6	0·47	0·51
Zn (mg)	11·1	10·8	0·3	2·4	−0·1, 0·6	0·45	0·60
Iodine (µg)	259·1	236·2	22·9	8·8	−10·4, 56·2	0·23	0·44
Se (µg)	52·0	46·2	5·8	11·1	2·7, 8·9	0·53	0·62

E%, percentage of total energy; RE, retinol equivalents.
*All nutrients except total energy were energy-adjusted using the residual method[30].

fish are well-defined food groups and the correlations for these groups were comparable to what have previously been reported[32,33,42]. We found low correlation between methods for intake of fats (e.g. butter, spread, oil). Detailed information on type of fats used in cooking (frying, etc.) was recorded in the FFQ because the association between fatty acids and fecundability is a central hypothesis of the 'Snart Forældre' study. In contrast, in the FD, fats used in cooking were assessed using standard recipes, which may explain the somewhat poor correlation. The relatively high correlation of saturated fat between the two methods may be explained by other sources of saturated fat such as dairy products, meat, etc.

The relatively high correlations for both food groups and nutrients were further supported by cross-classification analysis. Very few women were grossly misclassified and the majority of women, for most foods and nutrients, were either identically classified or within one adjacent category. These results are in accordance with results from other validation studies of FFQ[32,43] and demonstrate the ability of the FFQ in ranking individuals

according to low and high intake, respectively. In both the correlations and the cross-classification, results tended to be stronger for nutrients than for foods. Nutrient intake is expected to be more evenly distributed due to contributions from various foods, while larger day-to-day variation has been reported for intake of foods[44]. This is further evidenced by the greater improvement in correlations for foods after adjusting for day-to-day variability.

When comparing intakes assessed in this population with those in a nationally representative sample of Danish women[26], intakes for the participants in 'Snart Forældre' were similar to those of the larger sample of Danish women, with the exception of higher intakes of fish and vegetables, and lower intake of potatoes, among 'Snart Forældre' participants. However, the age range in the national sample was 18–75 years, and thus included women of older age than those in 'Snart Forældre'. Furthermore, participants in 'Snart Forældre' are likely to belong to a somewhat select group, with greater awareness of healthy eating habits, as pregnancy planning has been associated with healthier lifestyle[45], and as

Table 3 Cross-classification analysis of nutrient intakes estimated by the 4 d food diary and the newly developed, semi-quantitative, web-based FFQ: number and percentage in the same or adjacent quintile, and number and percentage grossly misclassified (opposite quintile). Participants (n 97; women aged 20–42 years) from the 'Snart Forældre' study, Denmark, 2013

	Same or adjacent quintile		Opposite quintile	
	n	%	n	%
Energy (kJ/d)	56	57·7	5	5·2
Carbohydrate (g/d)	74	76·3	1	1·0
Carbohydrate (E%)	72	74·2	1	1·0
Protein (g/d)	66	68·0	2	2·1
Protein (E%)	57	58·8	2	2·1
Fat (g/d)	66	68·0	1	1·0
Fat (E%)	67	69·1	1	1·0
Saturated fat (g/d)	64	66·0	3	3·1
Monounsaturated fat (g/d)	66	68·0	1	1·0
Polyunsaturated fat (g/d)	64	66·0	3	3·1
Added sugar (g/d)	77	79·4	1	1·0
Fibre (g/d)	80	82·5	2	2·1
Vitamin A (RE/d)	58	59·8	3	3·1
Vitamin D (μg/d)	64	66·0	4	4·1
Vitamin E (mg/d)	65	67·0	2	2·1
Thiamin (mg/d)	60	61·9	2	2·1
Riboflavin (mg/d)	77	79·4	2	2·1
Vitamin B_6 (mg/d)	57	58·8	3	3·1
Folate (μg/d)	63	64·9	1	1·0
Vitamin B_{12} (μg/d)	67	69·1	3	3·1
Vitamin C (mg/d)	63	64·9	1	1·0
Ca (mg/d)	73	75·3	1	1·0
P (mg/d)	74	76·3	2	2·1
Mg (mg/d)	74	76·3	–	–
Na (mg/d)	64	66·0	8	8·2
K (mg/d)	67	69·1	4	4·1
Total Fe (mg/d)	62	63·9	–	–
Non-haem Fe (mg/d)	67	69·0	1	1·0
Zn (mg/d)	66	68·0	3	3·1
Iodine (μg/d)	55	56·7	6	6·2
Se (μg/d)	66	68·0	3	3·1

E%, percentage of total energy.

participants in health surveys in general may be more health-conscious compared with the background population[46]. However, this is unlikely to affect the internal validity of analyses of diet and fertility, as participation is not expected to be associated with both exposure (diet) and outcome (fertility) because the outcome is not known at the time of exposure assessment[47].

When assessing the validity of a method (FFQ) using another method as reference (e.g. FD), it is important that the errors in the two methods are uncorrelated[48]. Assessing usual dietary intake by FFQ is highly dependent on the participant's pattern memory, while the FD is recorded at the time of consumption and thus not dependent on memory. Further, filling in the FFQ retrospectively would have no influence on the participant's habitual intake, while recording the diet on four consecutive days (FD) may alter the participant's diet. Thus, errors associated with FFQ and FD are to some extent considered to be independent, leading to true values of correlation[49].

Under-reporting is a problem in most dietary survey methods and cannot be avoided in the present study, but any resulting error in dietary measurement should be non-differential with respect to disease outcome[50]. Such misclassification would attenuate the true association, e.g. bias effect estimates towards the null for extreme categories.

A potential limitation of the present validation study is the lack of an objective measure, e.g. a biomarker of dietary intake independent of self-reported intake, with which to compare the FFQ, to further examine the validity of the FFQ. However, the purpose of the FFQ is to rank the participants according to intake rather than to assess the absolute intake. Therefore, a relative validation method, such as the FD, may be appropriate in assessing validity. Furthermore, biomarkers are not always useful because they may measure

Table 4 Median (and interquartile range) daily intakes of food groups (g/d) estimated by the 4 d food diary (FD) and the newly developed, semi-quantitative, web-based FFQ; P values for differences in intake between the FD and FFQ; and Pearson correlation coefficients and de-attenuated coefficients between intakes from the FD and FFQ. Participants (n 97; women aged 20–42 years) from the 'Snart Forældre' study, Denmark, 2013

	FD		FFQ		P value*	Pearson correlation coefficient	De-attenuated coefficient
	Median	IQR	Median	IQR			
High-fat dairy	39	17–69	66	40–111	<0·001	0·42	0·49
Low-fat dairy	140	32–224	243	109–344	<0·001	0·61	0·66
Cereals	185	144–217	174	151–217	0·566	0·35	0·43
Vegetables	228	148–306	248	173–370	0·010	0·38	0·44
Fruit	174	99–239	121	90–200	0·001	0·43	0·53
Potatoes	32	1–68	76	54–104	<0·001	0·26	0·48
Meat	91	61–143	77	62–101	0·002	0·41	0·66
Fish	24	3–48	24	16–33	0·388	0·38	0·75
Poultry	20	2–50	25	16–43	0·456	0·32	0·54
Eggs	23	13–45	22	16–33	0·295	0·47	0·62
Fats	32	24–40	31	25–38	0·649	0·17	0·25
Sugar	39	21–69	21	13–33	<0·001	0·21	0·28
Beverages	1844	1412–2519	1399	1074–1799	<0·001	0·46	0·48
Juice	9	0–64	18	16–45	0·648	0·29	0·36

IQR, interquartile range.
*Wilcoxon signed-rank test.

Table 5 Cross-classification analysis of food group intakes (g/d) estimated by the 4 d food diary and the newly developed, semi-quantitative, web-based FFQ: number and percentage in the same or adjacent quintile, and number and percentage grossly misclassified (opposite quintile). Participants (n 97; women aged 20–42 years) from the 'Snart Forældre' study, Denmark, 2013

	Same or adjacent quintile		Opposite quintile	
	n	%	n	%
High-fat dairy	64	66·0	2	2·1
Low-fat dairy	78	80·4	–	–
Cereals	70	72·2	4	4·1
Vegetables	65	67·0	2	2·1
Fruit	64	66·0	2	2·1
Potatoes	63	64·9	1	1·0
Meat	64	66·0	3	3·1
Fish	68	70·1	2	2·1
Poultry	67	69·1	2	2·1
Eggs	65	67·0	2	2·1
Fats	54	55·7	3	3·1
Sugar	66	68·0	2	2·1
Beverages	69	71·1	–	–
Juice	66	68·0	2	2·1

short-term intake and the study's purpose is measuring dietary intake over a longer period. A strength of the study is that a large proportion of the 'Snart Forældre' participants invited to the validation study completed the FFQ (82 %), indicating the study results will be generalizable to the entire 'Snart Forældre' cohort.

Conclusion

In conclusion, the present validation study showed that a web-based FFQ designed for women of reproductive age is appropriate for collecting data on dietary intake among Danish pregnancy planners. Further modification and testing of the FFQ would be needed to extend its use to other groups such as men, different age groups, or groups with different cultural backgrounds.

Acknowledgements

References

1. Smith S, Pfeifer SM & Collins JA (2003) Diagnosis and management of female infertility. *JAMA* **290**, 1767–1770.
2. Statistics Denmark (2014) Births. http://www.dst.dk/en/Statistik/emner/foedsler/foedsler.aspx (accessed September 2014).
3. Martin SP (2000) Diverging fertility among US women who delay childbearing past age 30. *Demography* **37**, 523–533.
4. Aubertin-Leheudre M, Gorbach S, Woods M et al. (2008) Fat/fiber intakes and sex hormones in healthy pre-menopausal women in USA. *J Steroid Biochem Mol Biol* **112**, 32–39.
5. Gann PH, Chatterton RT, Gapstur SM et al. (2003) The effects of a low-fat/high-fiber diet on sex hormone levels and menstrual cycling in premenopausal women: a 12-month randomized trial (the diet and hormone study). *Cancer* **98**, 1870–1879.
6. Gaskins AJ, Mumford SL, Zhang C et al. (2009) Effect of daily fiber intake on reproductive function: the BioCycle Study. *Am J Clin Nutr* **90**, 1061–1069.
7. Goldin BR, Adlercreutz H, Gorbach SL et al. (1982) Estrogen excretion patterns and plasma levels in vegetarian and omnivorous women. *N Engl J Med* **307**, 1542–1547.
8. Pedersen AB, Bartholomew MJ, Dolence LA et al. (1991) Menstrual differences due to vegetarian and nonvegetarian diets. *Am J Clin Nutr* **53**, 879–885.
9. Barr SI (1999) Vegetarianism and menstrual cycle disturbances: is there an association? *Am J Clin Nutr* **70**, 3 Suppl., 549S–554S.
10. Ruder EH, Hartman TJ, Blumberg J et al. (2008) Oxidative stress and antioxidants: exposure and impact on female fertility. *Hum Reprod Update* **14**, 345–357.
11. Chavarro JE, Rich-Edwards JW, Rosner BA et al. (2008) Protein intake and ovulatory infertility. *Am J Obstet Gynecol* **198**, 210.e1–7.
12. Chavarro JE, Rich-Edwards JW, Rosner BA et al. (2006) Iron intake and risk of ovulatory infertility. *Obstet Gynecol* **108**, 1145–1152.
13. Chavarro JE, Rich-Edwards JW, Rosner B et al. (2007) A prospective study of dairy foods intake and anovulatory infertility. *Hum Reprod* **22**, 1340–1347.
14. Chavarro JE, Rich-Edwards JW, Rosner BA et al. (2007) Dietary fatty acid intakes and the risk of ovulatory infertility. *Am J Clin Nutr* **85**, 231–237.
15. Chavarro JE, Rich-Edwards JW, Rosner BA et al. (2008) Use of multivitamins, intake of B vitamins, and risk of ovulatory infertility. *Fertil Steril* **89**, 668–676.
16. Mumford SL, Sundaram R, Schisterman EF et al. (2014) Higher urinary lignan concentrations in women but not men are positively associated with shorter time to pregnancy. *J Nutr* **144**, 352–358.
17. Ruder EH, Hartman TJ, Reindollar RH et al. (2014) Female dietary antioxidant intake and time to pregnancy among couples treated for unexplained infertility. *Fertil Steril* **101**, 759–766.
18. Vujkovic M, de Vries JH, Lindemans J et al. (2010) The preconception Mediterranean dietary pattern in couples undergoing *in vitro* fertilization/intracytoplasmic sperm injection treatment increases the chance of pregnancy. *Fertil Steril* **94**, 2096–2101.

19. Huybrechts KF, Mikkelsen EM, Christensen T *et al.* (2010) A successful implementation of e-epidemiology: the Danish pregnancy planning study 'Snart-Gravid'. *Eur J Epidemiol* **25**, 297–304.

20. Mikkelsen EM, Hatch EE, Wise LA *et al.* (2009) Cohort profile: the Danish Web-based Pregnancy Planning Study – 'Snart-Gravid'. *Int J Epidemiol* **38**, 938–943.

21. Wise LA, Rothman KJ, Mikkelsen EM *et al.* (2010) An internet-based prospective study of body size and time-to-pregnancy. *Hum Reprod* **25**, 253–264.

22. Sauvageot N, Alkerwi A, Albert A *et al.* (2013) Use of food frequency questionnaire to assess relationships between dietary habits and cardiovascular risk factors in NESCAV study: validation with biomarkers. *Nutr J* **12**, 143.

23. Danske medier (2011) *Danskernes brug af internettet 2012*. Copenhagen: Association of Danish Media.

24. Olsen SF, Mikkelsen TB, Knudsen VK *et al.* (2007) Data collected on maternal dietary exposures in the Danish National Birth Cohort. *Paediatr Perinat Epidemiol* **21**, 76–86.

25. Overvad K, Tjonneland A, Haraldsdottir J *et al.* (1991) Development of a semiquantitative food frequency questionnaire to assess food, energy and nutrient intake in Denmark. *Int J Epidemiol* **20**, 900–905.

26. Pedersen AN, Fagt S, Groth MV *et al.* (2010) *Danskernes Kostvaner 2003–2008. Hovedresultater* (*Dietary Habits in Denmark 2003–2008. Main Results*). Søborg: DTU Fødevareinstituttet.

27. Biltoft-Jensen A, Groth MV, Matthiessen J *et al.* (2009) Diet quality: associations with health messages included in the Danish Dietary Guidelines 2005, personal attitudes and social factors. *Public Health Nutr* **12**, 1165–1173.

28. Knudsen VK, Gille MB, Nielsen TH *et al.* (2011) Relative validity of the pre-coded food diary used in the Danish National Survey of Diet and Physical Activity. *Public Health Nutr* **14**, 2110–2116.

29. Biltoft-Jensen A, Matthiessen J, Rasmussen LB *et al.* (2009) Validation of the Danish 7-day pre-coded food diary among adults: energy intake v. energy expenditure and recording length. *Br J Nutr* **102**, 1838–1846.

30. Willett WC, Howe GR & Kushi LH (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 4 Suppl., 1220S–1228S.

31. Willett WC (2013) *Nutritional Epidemiology*, 3rd ed. Oxford: Oxford University Press.

32. Toft U, Kristoffersen L, Ladelund S *et al.* (2008) Relative validity of a food frequency questionnaire used in the Inter99 study. *Eur J Clin Nutr* **62**, 1038–1046.

33. Elorriaga N, Irazola VE, Defago MD *et al.* (2014) Validation of a self-administered FFQ in adults in Argentina, Chile and Uruguay. *Public Health Nutr* (Epublication ahead of print version).

34. Streppel MT, de Vries JH, Meijboom S *et al.* (2013) Relative validity of the food frequency questionnaire used to assess dietary intake in the Leiden Longevity Study. *Nutr J* **12**, 75.

35. Johansson I, Hallmans G, Wikman A *et al.* (2002) Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. *Public Health Nutr* **5**, 487–496.

36. Liu L, Wang PP, Roebothan B *et al.* (2013) Assessing the validity of a self-administered food-frequency questionnaire (FFQ) in the adult population of Newfoundland and Labrador, Canada. *Nutr J* **12**, 49.

37. Paalanen L, Männistö S, Virtanen MJ *et al.* (2006) Validity of a food frequency questionnaire varied by age and body mass index. *J Clin Epidemiol* **59**, 994–1001.

38. Basiotis PP, Welsh SO, Cronin FJ *et al.* (1987) Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. *J Nutr* **117**, 1638–1641.

39. Labonte ME, Cyr A, Baril-Gravel L *et al.* (2012) Validity and reproducibility of a web-based, self-administered food frequency questionnaire. *Eur J Clin Nutr* **66**, 166–173.

40. Christensen SE, Moller E, Bonn SE *et al.* (2014) Relative validity of micronutrient and fiber intake assessed with two new interactive meal- and web-based food frequency questionnaires. *J Med Internet Res* **16**, e59.

41. George GC, Milani TJ, Hanss-Nuss H *et al.* (2004) Development and validation of a semi-quantitative food frequency questionnaire for young adult women in the southwestern United States. *Nutr Res* **24**, 29–43.

42. Macedo-Ojeda G, Vizmanos-Lamotte B, Marquez-Sandoval YF *et al.* (2013) Validation of a semi-quantitative food frequency questionnaire to assess food groups and nutrient intake. *Nutr Hosp* **28**, 2212–2220.

43. Pinto E, Severo M, Correia S *et al.* (2010) Validity and reproducibility of a semi-quantitative food frequency questionnaire for use among Portuguese pregnant women. *Matern Child Nutr* **6**, 105–119.

44. Palaniappan U, Cue RI, Payette H *et al.* (2003) Implications of day-to-day variability on measurements of usual food and nutrient intakes. *J Nutr* **133**, 232–235.

45. Backhausen MG, Ekstrand M, Tyden T *et al.* (2014) Pregnancy planning and lifestyle prior to conception and during early pregnancy among Danish women. *Eur J Contracept Reprod Health Care* **19**, 57–65.

46. Hennekens CH & Buring JE (1987) *Epidemiology in Medicine*, 1st ed. Boston, MA/Toronto: Little, Brown and Company.

47. Olsen J, Melbye M, Olsen SF *et al.* (2001) The Danish National Birth Cohort – its background, structure and aim. *Scand J Public Health* **29**, 300–307.

48. Kaaks R, Ferrari P, Ciampi A *et al.* (2002) Uses and limitations of statistical accounting for random error correlations, in the validation of dietary questionnaire assessments. *Public Health Nutr* **5**, 969–976.

49. Margetts BM & Nelson M (editors) (1997) *Design Concepts in Nutritional Epidemiology*, 2nd ed. New York: Oxford University Press.

50. 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.