The Journal of Nutrition
Nutritional Epidemiology

Reproducibility of the Online Food4Me Food-Frequency Questionnaire for Estimating Dietary Intakes across Europe[1–3]

Steven J Marshall,[4,14] Katherine M Livingstone,[4,14] Carlos Celis-Morales,[4] Hannah Forster,[5] Rosalind Fallaize,[6] Clare B O'Donovan,[5] Clara Woolhead,[5] Cyril FM Marsaux,[7] Anna L Macready,[6] Santiago Navas-Carretero,[8,9] Rodrigo San-Cristobal,[8,9] Silvia Kolossa,[10] Lydia Tsirigoti,[11] Christina P Lambrinou,[11] George Moschonis,[11] Magdalena Godlewska,[12] Agnieszka Surwiłło,[12] Christian A Drevon,[13] Yannis Manios,[11] Iwona Traczyk,[12] J Alfredo Martínez,[8,9] Wim H Saris,[7] Hannelore Daniel,[10] Eileen R Gibney,[5] Lorraine Brennan,[5] Marianne C Walsh,[5] Julie A Lovegrove,[6] Mike Gibney,[5] and John C Mathers,[4]* on behalf of the Food4Me Study

[4]Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, United Kingdom; [5]University College Dublin Institute of Food and Health, University College Dublin, Belfield, Dublin, Republic of Ireland; [6]Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, University of Reading, Reading, United Kingdom; [7]Department of Human Biology, School of Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University Medical Centre, Maastricht, Netherlands; [8]Center for Nutrition Research, University of Navarra, Pamplona, Spain; [9]Biomedical Research Center Network (CIBER) in Physiopathology of Obesity and Nutrition, Institute of Health Carlos III, Madrid, Spain; [10]Research Center of Nutrition and Food Sciences (ZIEL), Biochemistry Unit, Technical University of Munich, Munich, Germany; [11]Department of Nutrition and Dietetics, Harokopio University, Athens, Greece; [12]The National Food & Nutrition Institute (NFNI), Warsaw, Poland; and [13]Department of Nutrition, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway

Abstract

Background: Accurate dietary assessment is key to understanding nutrition-related outcomes and is essential for estimating dietary change in nutrition-based interventions.

Objective: The objective of this study was to assess the pan-European reproducibility of the Food4Me food-frequency questionnaire (FFQ) in assessing the habitual diet of adults.

Methods: Participants from the Food4Me study, a 6-mo, Internet-based, randomized controlled trial of personalized nutrition conducted in the United Kingdom, Ireland, Spain, Netherlands, Germany, Greece, and Poland, were included. Screening and baseline data (both collected before commencement of the intervention) were used in the present analyses, and participants were included only if they completed FFQs at screening and at baseline within a 1-mo timeframe before the commencement of the intervention. Sociodemographic (e.g., sex and country) and lifestyle [e.g., body mass index (BMI, in kg/m^2) and physical activity] characteristics were collected. Linear regression, correlation coefficients, concordance (percentage) in quartile classification, and Bland-Altman plots for daily intakes were used to assess reproducibility.

Results: In total, 567 participants (59% female), with a mean ± SD age of 38.7 ± 13.4 y and BMI of 25.4 ± 4.8, completed both FFQs within 1 mo (mean ± SD: 19.2 ± 6.2 d). Exact plus adjacent classification of total energy intake in participants was highest in Ireland (94%) and lowest in Poland (81%). Spearman correlation coefficients (ρ) in total energy intake between FFQs ranged from 0.50 for obese participants to 0.68 and 0.60 in normal-weight and overweight participants, respectively. Bland-Altman plots showed a mean difference between FFQs of 210 kcal/d, with the agreement deteriorating as energy intakes increased. There was little variation in reproducibility of total energy intakes between sex and age groups.

Conclusions: The online Food4Me FFQ was shown to be reproducible across 7 European countries when administered within a 1-mo period to a large number of participants. The results support the utility of the online Food4Me FFQ as a reproducible tool across multiple European populations. This trial was registered at clinicaltrials.gov as NCT01530139. *J Nutr* 2016;146:1068–75.

Keywords: food-frequency questionnaire, reproducibility, online, dietary intakes, European

Introduction

Given that poor diet is a predominant cause of the growing burden of noncommunicable diseases, more effective strategies for improving diet are of increasing importance (1). In tandem, accurate dietary assessment tools are essential for evaluating the efficacy of lifestyle interventions (2), but all current methods of assessing habitual dietary intakes (including weighed dietary intakes, 24-h dietary recall, and FFQs) are subjective (3).

Manuscript received October 23, 2015. Initial review completed December 2, 2015. Revision accepted February 26, 2016.
First published online April 6, 2016; doi:10.3945/jn.115.225078.

Although weighed dietary recalls are considered the most accurate of the 3 (4), retrospective recalls (24-h recalls and FFQs) offer the advantages of lower costs and lower respondent burden (5) and are therefore widely used in large-scale epidemiologic and intervention studies.

With >70% of Europeans now Internet users (6), Internet-based diet and lifestyle interventions, including Internet-based FFQs, are an attractive, cost-effective, and scalable alternative to face-to-face interventions (7). However, self-reported dietary assessment is prone to respondent bias (8), which may limit reproducibility of the FFQ, resulting in poor measures of dietary change and in chance associations with disease outcomes (9, 10). It is therefore essential to evaluate the measurement error and reproducibility of FFQs to ensure confidence in the precision of any diet-related outcomes.

The online Food4Me FFQ used in this study was validated previously against a weighed food record over a 4-wk period (n = 49) and showed moderate agreement (correlation coefficient = 0.47) for assessing energy and nutrient intake (11) and a good agreement (0.60) against the EPIC (European Prospective Investigation into Cancer and Nutrition)-Norfolk printed FFQ (n = 113) (12). Furthermore, the reproducibility of the online Food4Me FFQ was assessed in the United Kingdom (n = 100) and showed good agreement, with mean cross-classification into "exact agreement plus adjacent" at 92% for both nutrient and food group intakes (11). The aim of our present investigation was to verify that the online Food4Me FFQ was reproducible across 7 European countries by comparing estimated intakes of foods, energy, and nutrients between screening and baseline in the Food4Me study.

Methods

Study design. The Food4Me study was a 6-mo, Internet-based, randomized controlled trial of personalized nutrition designed to improve diet and physical activity behaviors, which was conducted across 7 European countries (n = 1607). Recruitment was via the Food4Me website (13) from the following sites: University College Dublin (Ireland), Maastricht University (Netherlands), University of Navarra (Spain), Harokopio University (Greece), University of Reading (United Kingdom), National Food and Nutrition Institute (Poland), and Technical University of Munich (Germany). Individuals with ill health, food intolerances, or special nutritional requirements (e.g., pregnancy) were ineligible to participate. BMI (in kg/m^2) was estimated from self-reported body weight and height (14). Participants self-reported smoking habits and occupation. Physical activity level (ratio between total energy expenditure and basal metabolic rate) and sedentary behavior (min/d) were estimated from triaxial accelerometers (TracmorD; Philips Consumer Lifestyle). The Research Ethics Committees at each university or research center granted ethical approval for the study. All participants signed online consent forms. The Food4Me trial was registered as a randomized controlled trial at clinicaltrials.gov as NCT01530139. Full details on the study design are available elsewhere (14).

Food4Me FFQ. The Food4Me FFQ is an online, semiquantitative FFQ that was administered to individuals at screening, baseline, and follow-up time points after randomization. For the purposes of this reproducibility study, screening and baseline were used because no change in diet was expected. FFQs were available in the language of the country, with respondents asked to report mean consumption over the previous month for 157 items in the United Kingdom and Ireland [based on the 130-item printed EPIC-Norfolk FFQ (version CAMB/PQ/6/1205) (12, 15)], with additional country-specific foods added to capture intakes in the other 5 recruitment countries (e.g., "stroopwafels" was added to the Dutch FFQ). A total of 11 food categories were included: *1*) cereal; *2*) bread and savory biscuits; *3*) potatoes, rice, and pasta; *4*) meat and fish; *5*) dairy products; *6*) fats and spreads; *7*) sweets and snacks; *8*) soups, sauces, and spreads; *9*) drinks; *10*) fruit; and *11*) vegetables (**Supplemental Table 1**). Frequency of consumption of each food item was estimated by selecting one of the following options: never or <1 time/mo, 1–3 times/mo, 1 time/wk, 2–4 times/wk, 5–6 times/wk, 1 time/d, 2–3 times/d, 5–6 times/d, or >6 times/d. The online Food4Me FFQ included photographs of the foods, and participants selected the appropriate portion size from the following options: very small, small, small/medium, medium, medium/large, large, or very large. Food intake (g/d) was then calculated by multiplying portion size by frequency of consumption. For the purpose of comparing food group intakes, the 11 food categories were subdivided into 35 food groups based on previous validation by Forster et al. (12). Further details on the Food4Me FFQ are provided elsewhere (14).

Statistical analysis. Statistical analysis was performed with STATA (version 12; StataCorp LP) and MedCalc Statistical Software (version 12.2.1.0; MedCalc). ANOVA (continuous data) and logistic regression (categorical) tested for overall differences in anthropometric and sociodemographic characteristics (dependent variable) between countries (independent variable) and were adjusted for age and sex. Post hoc Tukey's tests and logistic regression (adjusted for age and sex) investigated differences in characteristics (dependent variable) between a given country and the overall mean for all countries (independent variable) (**Table 1**). FFQ reproducibility was determined by comparing dietary intakes at screening and baseline (mean ± SD: 2.7 ± 0.9 wk apart). Because the FFQ was designed to assess dietary intakes over a 1-mo period, participants were excluded from the current analysis if the time period between completion of FFQs was >1 mo (16). Participants with implausible energy intakes were excluded based on the upper limit of sustained energy expenditure defined by the Scientific Advisory Committee for Nutrition: energy intake >2.5 × basal metabolic rate (17). Multiple linear regression was used to determine differences in total energy, nutrient, and food group intakes (dependent variable) between FFQs (independent variable) and were adjusted for age, sex, country, time of FFQ completion, and total energy intake at screening. Normality of data was assessed with the Shapiro-Wilk test, and depending on the outcome, comparison of energy, nutrients, and food group intake was assessed with Pearson correlation coefficient or Spearman correlation coefficient (SCC, ρ). Energy-adjusted correlation coefficients were estimated with the residual method (18). Briefly, residuals from the regression analysis (energy intake as an independent variable and nutrient intake as a dependent variable) were added to the expected nutrient value for the mean energy intake of the sample (**Tables 2** and **3**). The coefficient of reproducibility between methods was calculated (19). Concordance (percentage) in quartile classification estimated the relative agreements between FFQs. Quartiles of intakes of nutrients and food groups were used to determine changes in classification between time points. The percentages of participants classified into the correct quartile (exact classification), adjacent quartile (exact classification plus adjacent), 2 quartiles apart (misclassification), or 3 quartiles apart (extreme misclassification) were estimated (**Supplemental Tables 2** and 3). Bland-Altman plots determined clinical relevance of any difference in total energy and nutrients between methods based on the mean difference between methods (bias), trends, variability, and widths of the limits of agreement (**Figure 1**). Reproducibility of total energy intakes was

[1] Supported by the European Commission under the Food, Agriculture, Fisheries and Biotechnology Theme of the 7th Framework Programme for Research and Technological Development (265494).
[2] Author disclosures: SJ Marshall, KM Livingstone, C Celis-Morales, H Forster, R Fallaize, CB O'Donovan, C Woolhead, CFM Marsaux, AL Macready, S Navas-Carretero, R San-Cristobal, S Kolossa, L Tsirigoti, CP Lambrinou, G Moschonis, M Godlewska, A Surwiłło, CA Drevon, Y Manios, I Traczyk, JA Martínez, WH Saris, H Daniel, ER Gibney, L Brennan, MC Walsh, JA Lovegrove, M Gibney, and JC Mathers, no conflicts of interest.
[3] Supplemental Tables 1–4 and Supplemental Figure 1 are available from the "Online Supporting Material" link in the online posting of the article and from the same link in the online table of contents at http://jn.nutrition.org.
[14] These authors contributed equally to this work.
*To whom correspondence should be addressed. E-mail: john.mathers@newcastle.ac.uk.

TABLE 1 Anthropometric and sociodemographic characteristics of European adults by country at the time of completing the screening Food4Me FFQ[1]

	Total	Country						P value
		Greece	Ireland	Netherlands	Poland	United Kingdom	Germany	
	(n = 567)	(n = 160)	(n = 70)	(n = 108)	(n = 153)	(n = 49)	(n = 27)	
Age, y	38.7 ± 13.4	38.3 ± 11.2	39.6 ± 13.2	42.7 ± 16.6	35.0 ± 12.1	38.4 ± 12.6	43.4 ± 15.4	<0.001
Sex, F	58.9	58.1	41.4*	50.9	70.1*	67.4	59.3	0.03
Ethnicity, Caucasian	97.5	99.4	97.1	95.4	100	87.8*	100	0.04
Occupation								
Professional and managerial	31.2	31.3	40.0	36.1	19.0	46.9	29.6	0.98
Intermediate occupation	29.1	28.1	21.4	17.6	46.4	12.2	33.3	0.47
Routine and manual	11.6	18.1*	14.3	8.3	5.9*	14.3	7.4	0.02
Student	17.1	7.5*	15.7	24.1	22.9	18.4	14.8	0.048
Not currently working	10.9	15.0*	8.6	13.9	5.9	8.2	14.8	0.04
Anthropometric measurements								
BMI, kg/m^2	25.4 ± 4.8	26.7 ± 5.5*	26.0 ± 4.6	24.4 ± 3.9*	24.7 ± 4.7	25.3 ± 4.3	24.5 ± 3.0	<0.001
Waist circumference, cm	85.5 ± 14.1	89.3 ± 14.8*	87.5 ± 14.1	84.6 ± 12.5	81.6 ± 14.2	84.2 ± 11.7	85.4 ± 13.0	<0.001
Body weight, kg	75.0 ± 15.4	76.9 ± 15.7	78.3 ± 16.3	74.7 ± 13.5	71.1 ± 16.2	72.7 ± 14.1	75.0 ± 12.1	0.13
Physical activity								
PAL	1.7 ± 0.2	1.7 ± 0.1	1.8 ± 0.2	1.7 ± 0.2	1.7 ± 0.2	1.7 ± 0.2	1.7 ± 0.2	0.07
SB, min/d	745 ± 78.0	744 ± 89.4	755 ± 72.1	753 ± 72.1	741 ± 79.2	725 ± 59.2	762 ± 66.7	0.36

[1] Values represent means ± SDs or percentages. *Post hoc Tukey tests (continuous data) and logistic regression (categorical) were used to test for significant differences between a given country and the overall study mean across all countries, $P < 0.05$. ANOVA and logistic regression were used to test for significant differences across countries in continuous and categorical variables, respectively. Analyses were adjusted for age and sex. PAL, physical activity level (ratio between total energy expenditure and basal metabolic rate); SB, sedentary behavior.

assessed according to age (<45 y and ≥45 y), sex, country, completion period between FFQs (short: 0–15.6 d; medium: 15.6–22.6 d; long: 22.6–31 d), and BMI at screening (underweight: BMI <18.5; normal weight: 18.5–24.9; overweight: 25 to 29.9; obese: ≥30) with the use of regression analyses, SCC, and concordance (percentage) in quartile classification (**Supplemental Table 4**).

Sensitivity analysis. Sensitivity analyses excluded participants who over- or underreported energy intakes (**Supplemental Figure 1**). Underreporting was operationalized as an energy intake <1.1 kcal/d multiplied by the predicted basal metabolic rate [with the use of the Henry equation (20)] (21), and energy intakes >4500 kcal/d were classified as overreporting (22).

Results

Of the 1607 randomly allocated participants, 1480 completed the FFQs at screening and at baseline, and 665 completed the FFQs within 1 mo of each other. Spain was excluded from all analyses due to insufficient numbers completing the FFQs within the 1-mo timeframe (n = 5). A further 93 participants were excluded based on implausible energy intakes. Individuals from Greece had a higher BMI and waist circumference, more participants in routine and manual work, fewer students, and more participants not currently working than the overall mean across all countries. Fewer Polish participants were in routine and manual employment and more Polish participants were females, whereas more Dutch participants were leaner than the overall mean. Fewer participants from the United Kingdom were Caucasian, whereas there were fewer female participants from Ireland than the overall mean. No significant differences in physical activity level, body weight, or sedentary behavior were identified (Table 1).

Reproducibility of nutrient intakes. Total energy intakes and intakes of protein, carbohydrates, total fat, SFAs, MUFAs, PUFAs, n–3 FAs, sugar, salt, calcium, folate, iron, carotene, riboflavin, fiber, sodium, and vitamins B-6, C, A, D, and E were lower at baseline than at screening ($P < 0.05$; Table 2). There were no significant differences between time points for percentage energy intakes from total fat, MUFAs, PUFAs, protein, carbohydrates, and sugars or for intakes of alcohol, vitamin B-12, thiamin, and retinol. Shapiro-Wilk tests revealed that data were not normally distributed; therefore, SCC was used to examine correlations. Unadjusted SCCs ranged from 0.59 for total fat (g/d) to 0.89 for alcohol (mean 0.67; $P < 0.001$), whereas energy-adjusted SCCs ranged from 0.59 for total fat to 0.89 for alcohol (0.69; $P < 0.001$; Table 2).

The percentage of participants whose dietary intakes were classified exactly at baseline, compared with screening, ranged from lowest for total fat to highest for alcohol (mean 50%; Supplemental Table 2). In total, 88% of participants were classified into the same or adjacent quartile, 10% were misclassified, and 2% were extremely misclassified.

Bland-Altman plots comparing intakes of energy, total fat, protein, and carbohydrates between time points are shown in Figure 1. The bias (mean difference) for total energy, carbohydrate, protein, and fat intake was 210 kcal/d, 11.4%, 9.1%, and 9.0%, respectively. A positive trend indicated a lower agreement in intakes between time points for those who reported higher energy intakes (>4500 kcal/d) and who were classified as overreporters in the sensitivity analyses. The amount consumed did not affect the agreement between intakes of carbohydrates, protein, and fat.

Reproducibility of food group intakes. Reported intakes of whole-meal bread, biscuits, other fruit, meat products and soups, sauces, and miscellaneous foods were lower at baseline than at screening ($P < 0.05$; Table 3). Unadjusted SCCs ranged from 0.42 for tinned fruit or vegetables to 0.89 for alcoholic beverages (mean 0.71, $P < 0.001$), whereas energy-adjusted

TABLE 2 Differences in total energy and nutrient intakes in European adults between screening and baseline as assessed with multiple linear regression and correlation coefficients[1]

	Time point[2]			Correlation coefficient[4]	
	Screening	Baseline	*P* value[3]	Crude	Energy adjusted
Total energy, kcal/d	2455 ± 685	2246 ± 730	<0.001	0.61	—
Total fat, g/d	96.4 ± 32.2	89.2 ± 32.8	<0.001	0.59	0.59
Total fat, % energy	35.4 ± 6.2	35.7 ± 5.8	0.423	0.61	0.61
SFAs, g/d	38.1 ± 14.6	35.1 ± 14.6	0.001	0.61	0.64
SFAs, % energy	13.9 ± 3.2	14 ± 3.1	0.78	0.65	0.65
MUFAs, g/d	36.7 ± 13.8	33.9 ± 13.2	<0.001	0.62	0.69
MUFAs, % energy	13.5 ± 3.5	13.6 ± 3.2	0.54	0.72	0.71
PUFAs, g/d	15.3 ± 5.4	14.3 ± 5.8	0.004	0.67	0.67
PUFAs, % energy	5.6 ± 1.3	5.8 ± 1.4	0.13	0.68	0.68
n–3 FAs, g/d	1.8 ± 0.7	1.7 ± 0.7	0.004	0.65	0.68
Protein, g/d	104 ± 34.3	95 ± 33.1	<0.001	0.63	0.68
Protein, % energy	17.1 ± 3.4	17.2 ± 3.4	0.49	0.71	0.70
Carbohydrates, g/d	288 ± 96.7	259 ± 96.1	<0.001	0.64	0.63
Carbohydrates, % energy	46.8 ± 7.6	46 ± 7.4	0.11	0.65	0.66
Total sugars, g/d	128 ± 47.8	117 ± 48.0	<0.001	0.66	0.72
Total sugars, % energy	21.1 ± 6.1	21 ± 5.9	0.83	0.73	0.73
Fiber, g/d	29.8 ± 12.1	26.8 ± 11.5	<0.001	0.71	0.73
Alcohol, g/d	10.4 ± 12.8	10.3 ± 13.7	0.83	0.89	0.89
Calcium, g/d	1230 ± 478	1110 ± 462	<0.001	0.63	0.69
Folate, μg/d	370 ± 131	338 ± 130	<0.001	0.65	0.70
Iron, mg/d	15.6 ± 5.1	14.2 ± 5	<0.001	0.62	0.63
Carotene, mg/d	6390 ± 5900	5550 ± 4100	0.005	0.7	0.71
Riboflavin, mg/d	2.3 ± 0.9	2.1 ± 0.9	0.001	0.71	0.76
Thiamin, mg/d	2.5 ± 2.3	2.4 ± 2.3	0.34	0.62	0.59
Vitamin B-6, mg/d	2.7 ± 0.9	2.5 ± 0.9	<0.001	0.67	0.69
Vitamin B-12, μg/d	7.7 ± 4.1	7.3 ± 4.1	0.06	0.73	0.75
Vitamin C, mg/d	167 ± 99.7	155 ± 94.3	0.04	0.73	0.76
Vitamin A, mg/d	1660 ± 1080	1510 ± 886	0.008	0.67	0.68
Retinol, μg/d	593 ± 451	582 ± 496	0.65	0.65	0.62
Vitamin D, μg/d	3.8 ± 2.3	3.5 ± 2	0.04	0.67	0.66
Vitamin E, mg/d	11.4 ± 4.3	10.4 ± 4.4	<0.001	0.67	0.70
Salt, g/d	7.2 ± 2.9	6.5 ± 2.7	<0.001	0.65	0.67
Sodium, mg/d	2900 ± 1140	2610 ± 1090	<0.001	0.65	0.67

[1] Values represent means ± SDs or percentages, $n = 567$.
[2] Mean difference between screening and baseline questionnaires was 2.7 ± 0.9 wk.
[3] Multiple linear regression between screening and baseline FFQs adjusted for country, time of FFQ completion, age, sex, and total energy at screening.
[4] Spearman correlation coefficient between screening and baseline FFQs. All results were significant to $P < 0.001$.

SCCs ranged from 0.45 for rice, pasta, grains, and starches to 0.87 for alcoholic beverages (mean 0.69; $P < 0.001$).

As shown in Supplemental Table 3, the percentage of participants correctly classified into the same quartile for food group intakes was lowest for rice, pasta, grains, and starches and highest for alcoholic beverages. For all food groups, the mean percentages of participants who were misclassified and extremely misclassified were 8% and 2%, respectively.

Subgroup analysis: reproducibility of total energy intakes. As summarized in Supplemental Table 4, energy intake was lower at baseline than at screening for Greece, Poland, and Germany. Correlations in energy intakes between time points were highest for the Netherlands and lowest for Greece, whereas the percentage energy intakes correctly classified were lowest in Germany and the United Kingdom and highest in the Netherlands. Energy intake was lower at baseline than at screening for those with short and medium time between assessments but not for the longest. For participants with the longest period between completing FFQs, SCCs of energy intakes were poorest (Supplemental Table 3). Energy intake was lower at baseline than at screening for normal-weight and overweight participants but not for obese participants. SCCs were lower and the percentage of individuals misclassified was higher in overweight and obese participants than normal-weight participants (Supplemental Table 3). Energy intake was lower at baseline than at screening for participants aged ≥45 and <45 y. SCCs for energy intakes between time points were higher for participants aged ≥45 y, with similar proportions of individuals correctly classified and extremely misclassified. Energy intakes at baseline were lower than at screening for both males and females. Although more females than males were correctly classified into the same quartile, more females than males were misclassified (Supplemental Table 3).

Sensitivity analysis. Analyses were repeated in valid reports ($n = 437$) after the removal of overreporters ($n = 8$) and underreporters ($n = 122$). Supplemental Figure 1 summarizes

TABLE 3 Differences in the Food4Me FFQ food group intakes (g/d) in European adults between screening and baseline as assessed with multiple linear regression and correlation coefficients[1]

	Time point[2]			Correlation coefficient[4]	
	Screening	Baseline	*P* value[3]	Crude	Energy adjusted
Rice, pasta, grains, and starches	76.2 ± 57.8	70.2 ± 56.5	0.08	0.52	0.45
Savory food (lasagna, pizza)	36.6 ± 33.3	34.7 ± 35.4	0.34	0.65	0.65
White bread (rolls, tortillas, crackers)	53 ± 95.4	44.2 ± 73.9	0.07	0.76	0.76
Whole meal, brown breads, and rolls	103 ± 131	86.3 ± 102	0.01	0.75	0.69
Breakfast cereals and porridge	56.9 ± 73	52.8 ± 73.4	0.35	0.81	0.80
Biscuits	28.1 ± 46.1	22.4 ± 40.8	0.03	0.61	0.60
Cakes, pastries, and buns	15.7 ± 17.4	14.6 ± 16.8	0.34	0.57	0.54
Milk	185 ± 215	170 ± 199	0.21	0.7	0.66
Cheeses	38.5 ± 36.7	35.7 ± 35.5	0.17	0.64	0.67
Yogurts	70.9 ± 89.4	76.6 ± 119	0.27	0.66	0.61
Ice cream, creams, and desserts	21.9 ± 22	21.5 ± 25.4	0.74	0.61	0.59
Eggs and egg dishes	30.8 ± 49.4	29.2 ± 41.9	0.55	0.75	0.68
Fats and oils (e.g., butter, low-fat spreads)	19.7 ± 17.3	18.5 ± 15.1	0.16	0.7	0.69
Potatoes and potato dishes	55.4 ± 56.6	53.1 ± 51.5	0.46	0.74	0.71
Chipped, fried, and roasted potatoes	14.8 ± 16.8	15.5 ± 17.5	0.49	0.77	0.75
Peas, beans, lentils, and vegetable dishes	31.9 ± 33.8	33.1 ± 47.5	0.56	0.79	0.78
Green vegetables	43.6 ± 49.9	38.9 ± 39.5	0.07	0.68	0.70
Carrots	22.6 ± 36.2	19.4 ± 20.5	0.11	0.67	0.66
Salad vegetables (e.g., lettuce)	51.2 ± 57.4	47.5 ± 46.6	0.06	0.77	0.78
Other vegetables (e.g., onions)	55.2 ± 50.2	51.8 ± 47.3	0.24	0.75	0.74
Tinned fruit or vegetables	2.2 ± 8.8	1.9 ± 6.3	0.45	0.42	0.46
Bananas	41.1 ± 50.5	37.6 ± 43.8	0.26	0.81	0.82
Other fruit (e.g., apples, pears, oranges)	246 ± 214	218 ± 196	0.02	0.8	0.81
Nuts and seeds, herbs, and spices	4.8 ± 7.6	4.9 ± 9.4	0.91	0.68	0.67
Fish and fish products/dishes	48.3 ± 40.2	47 ± 42.2	0.60	0.75	0.73
Bacon and ham	18.1 ± 24.9	17.8 ± 27.3	0.81	0.76	0.73
Red meat (e.g., beef, veal, lamb, pork)	38.4 ± 36	36.8 ± 33.4	0.40	0.74	0.73
Poultry (chicken and turkey)	26.2 ± 36.2	22.7 ± 24.5	0.05	0.59	0.58
Meat products (e.g., burgers and sausages)	46 ± 53.1	40 ± 37.8	0.03	0.65	0.64
Alcoholic beverages	134 ± 173	139 ± 207	0.69	0.89	0.87
Sugars, syrups, preserves, and sweeteners	4.7 ± 10.3	4.5 ± 9.1	0.66	0.85	0.81
Confectionary and savory snacks	16.7 ± 20.8	15.9 ± 21.7	0.56	0.71	0.64
Soups, sauces, and miscellaneous foods	103 ± 80.9	92.7 ± 75.4	0.03	0.71	0.68
Teas and coffees	593 ± 505	579 ± 484	0.67	0.81	0.68
Other beverages (e.g., fruit juices, squash)	238 ± 289	223 ± 277	0.35	0.75	0.74

[1] Values represent means ± SDs, $n = 567$.
[2] Mean difference between screening and baseline questionnaires was 2.7 ± 0.9 wk.
[3] Multiple linear regression between screening and baseline FFQs adjusted for country, time of FFQ completion, age, sex, and energy intake at screening.
[4] Spearman correlation coefficient between screening and baseline FFQs. All results were significant to $P < 0.001$.

the delta between time points for percentage energy from fat, carbohydrates, and protein in the total cohort and in valid reporters. This difference between time points is consistently smaller for the valid reporters in comparison with the whole cohort. After exclusion of misreporters, differences between time points in reported intakes of total fat, SFAs, MUFAs, PUFAs, n–3 FAs, protein, calcium, carotene, riboflavin, vitamins C and A, biscuits, other fruit and soups, sauces, and miscellaneous foods were not significant. For nutrients, SCC ranged from 0.60 for total fat and SFA g/d to 0.91 for alcohol and for food groups from 0.52 for rice, pasta, grains, and starches to 0.91 for alcoholic beverages ($P < 0.001$). Bland-Altman analysis on valid reports produced a higher agreement between time points for total energy intake (bias reduced from 210 to 88.5 kcal/d), carbohydrates (11.4–5.3%), protein (9.2–2.3%), and fat (9.5–2.4%). The coefficient of reproducibility in valid reports was reduced by 780 kcal/d for energy intake, 14.4% for percentage energy from carbohydrates, 12.7% for protein intake, and 13.3% for fat intake.

Discussion

Main findings. Our main findings indicate that the online Food4Me FFQ is reproducible for estimation of nutrient and food group intakes by adults across 7 European countries.

Comparison with other studies. An earlier study investigated the reproducibility of the online Food4Me FFQ by asking 100 participants within a single country (United Kingdom) to complete the FFQ on 2 occasions 4 wk apart. In that study, Fallaize et al. (11) reported higher mean correlation coefficients than in the present study for total energy intake (0.77 compared with 0.61), nutrients (0.75 compared with 0.67), and food group

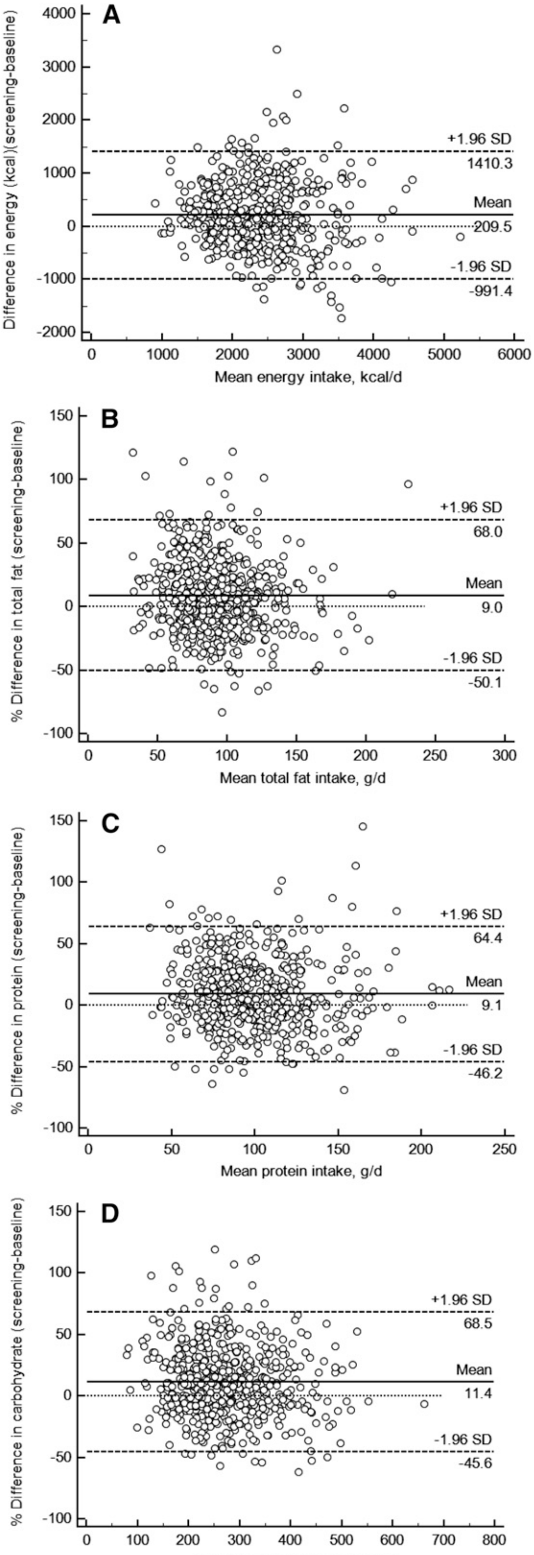

FIGURE 1 Bland-Altman plots for reproducibility between screening and baseline intakes of total energy (A), fat (B), protein (C), and carbohydrates (D) (n = 567) in European adults. The solid line represents the mean difference, the dashed line represents the limits of agreement, and the dotted line represents the trend in agreement.

intakes (0.75 compared with 0.71). Cross-classification analysis for nutrients was also higher, with 92% of participants classified into the same or adjacent quartile, compared with 88% in the current study. Bland-Altman analysis indicated a lower mean difference for total energy intake in the study by Fallaize et al. (11) than in ours (135 kcal/d compared with 210 kcal/d), but the removal of misreporters lowered the mean difference in the current study to 89 kcal/d. In the current study, the online Food4Me FFQ was administered to a much larger and more diverse group of participants across 7 European countries who, in addition to completing the FFQ, were responding to a wider range of questionnaires. Furthermore, FFQ reproducibility in the study by Fallaize et al. (11) was assessed in conjunction with validation against a 4-d weighed food diary, which may have increased the participants' awareness of their habitual intake, and, thus, they may have been more likely to report similar intakes. The observed lower agreement between repeated administrations of the FFQ in the current study may be because the participants were less focused on the FFQ per se. Previous studies of the reproducibility of FFQs have reported correlation coefficients for total energy intake of 0.66 and 0.65 (8, 23, 24), which are very similar to our observations. The much higher correlation of 0.82 reported by Beasley et al. (25) was for an Internet-based FFQ repeated within a short time interval (1 wk) and thus subject to less variation (26). The shortest interval between FFQ administrations in the current study (0–15.65 d) produced a correlation of 0.64, lower than the 0.82 reported by Beasley et al. (25). However, reproducibility in Beasley et al. (25) was also accompanied by a validation study against a 4-d weighed food diary, which may have improved correlations by increasing the participants' awareness of their diet. Cross-classification analyses in the current study showed agreements that were comparable with previous studies for energy, nutrients, and food groups (27–29). We observed that reported energy intakes were lower in the second FFQ, which confirms findings from other reproducibility studies (11, 25, 28, 30) and may be attributed to the learning effect of repeated measure. Alternatively, this observation may be due to fatigue caused by overburdening participants who had recently completed the initial FFQ (31). However, when misreporters were excluded, most differences between screening and baseline were no longer significant.

Previous FFQ reproducibility studies using repeated assessments within 1 mo have reported coefficient ranges of 0.58–0.86 for energy intake between several countries (11, 23, 25, 28, 29, 32). Intercountry variations in SCCs in the Food4Me FFQ were similar, suggesting that this dietary assessment tool has wide applicability across several European countries. The disparity between the cross-classifications and SCCs in the United Kingdom may have been due to the presence of dietary misreporters, and following exclusion of misreporters, these measures of reproducibility were more closely aligned. Our sex-dependent findings are consistent with a previous study (33), reporting higher reproducibility for a 240-item FFQ in males than in females (Pearson correlation coefficient = 0.70 and 0.65, respectively). The reproducibility of the online Food4Me FFQ was similar for both older and younger participants. The lower reported energy intake at baseline compared with screening was significant for both normal-weight and overweight participants but not for obese participants. This is probably due to a smaller sample size of obese individuals (n = 79) compared with normal-weight (n = 296) and overweight (n = 192) individuals, because when assessed by SCC, reproducibility was lowest in the obese group. These findings confirm previous results, in which obese individuals were more likely to misreport their dietary intakes (34, 35). Self-administered dietary assessment tools should thus be interpreted with caution when applied to a population of predominantly obese subjects.

Previous studies on the validation and reproducibility of the Food4Me FFQ excluded under- and overreporters before the main analysis (11, 12). The current study included the whole cohort. The percentage of people underreporting (21.5%) was higher than that of overreporters (1.4%), a common occurrence that has been previously reported (36). A sensitivity analysis after removal of misreports improved the reproducibility of the Food4Me FFQ.

Strengths and limitations. The main strength of this study is the large number of participants from 7 European countries, which enabled stratification according to country, age, sex, obesity status, and time interval between FFQs. However, by excluding participants who did not complete FFQs within a 1-mo period, we had too few participants from Spain ($n = 5$) to allow comparisons with this country. Nonetheless, another strength of this study is that it was possible to assess the FFQ reproducibility between valid reporters and misreporters in a European population. As recommended by Cade et al. (16), we applied the cutoff of <1 mo between repeated FFQs to avoid confounding by real temporal changes in food intake. With a short time between the FFQs, it is conceivable that participants might remember and therefore replicate their previous FFQ responses (16). However, the comprehensive nature of the online Food4Me FFQ would make this unlikely, and a 1-mo period is considered an optimal time period to assess reproducibility (16), while minimizing any influence of dietary change over time (11).

In conclusion, the Food4Me FFQ is moderately reproducible when administered to a large cohort of European adults. Variations in reproducibility between countries were small, thus providing confidence in the utility of the method for reporting intakes of energy, nutrients, and food groups across multiple European countries.

Acknowledgments

CAD, YM, IT, JAM, WHS, HD, ERG, LB, JAL, MG, and JCM contributed to the research design; JCM was the proof-of-principle study leader; CC-M, HF, RF, CBO, CW, CFMM, ALM, SN-C, RS-C, SK, LT, CPL, GM, MG, AS, ERG, LB, MCW, and JCM contributed to the developing of the standardized operating procedure for the study; CC-M, HF, RF, CBO, CW, CFMM, ALM, SN-C, RS-C, SK, LT, CPL, MG, AS, MCW, and JCM conducted the intervention; CC-M, CFMM, and WHS contributed to physical activity measurements; and SJM and KML drafted the paper and performed the statistical analysis for the manuscript. All authors contributed to a critical review of the manuscript during the writing process, and read and approved the final version of the paper.

References

1. Beaglehole R, Bonita R, Horton R, Adams C, Alleyne G, Asaria P, Baugh V, Bekedam H, Billo N, Casswell S, et al. Priority actions for the non-communicable disease crisis. Lancet 2011;377:1438–47.
2. Adamson AJ, Mathers JC. Effecting dietary change. Proc Nutr Soc 2004;63:537–47.
3. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sorensen TI, Speakman JR, Jeansonne M, Allison DB. Energy balance measurement: when something is not better than nothing. Int J Obes (Lond) 2015;39:1109–13.
4. Boushey C, Coulston AM. Nutrition in the prevention and treatment of disease. Amsterdam: Academic Press; 2008.
5. Subar AF. Developing dietary assessment tools. J Am Diet Assoc 2004;104:769–70.
6. Miniwatts Marketing Group. World Internet usage and population statistics [cited 2015 Apr 11]. Available from: http://www.internetworldstats.com/stats.htm.
7. Steele RM, Mummery WK, Dwyer T. A comparison of face-to-face or Internet-delivered physical activity intervention on targeted determinants. Health Educ Behav 2009;36:1051–64.
8. Fernández-Ballarth JD, Lluis Pinol J, Zazpe I, Corella D, Carrasco P, Toledo E, Perez-Bauer M, Angel Martinez-Gonzalez M, Salas-Salvado J, Martin-Moreno JM. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. Br J Nutr 2010;103:1808–16.
9. Marks GC, Hughes MC, van der Pols JC. Relative validity of food intake estimates using a food frequency questionnaire is associated with sex, age, and other personal characteristics. J Nutr 2006;136:459–65.
10. Schatzkin A, Kipnis V, Carroll RJ, Midthune D, Subar AF, Bingham S, Schoeller DA, Troiano RP, Freedman LS. 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 2003; 32:1054–62.
11. Fallaize R, Forster H, Macready AL, Walsh MC, Mathers JC, Brennan L, Gibney ER, Gibney MJ, Lovegrove JA. Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. J Med Internet Res 2014;16:e190.
12. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, Macready AL, Lovegrove JA, Mathers JC, Gibney MJ, et al. Online dietary intake estimation: the Food4Me food frequency questionnaire. J Med Internet Res 2014;16:e150.
13. Food4Me. An integrated analysis of opportunities and challenges for personalised nutrition [cited 2015 Dec 12]. Available from: http://www.food4me.org/.
14. Celis-Morales C, Livingstone KM, Marsaux CFM, Forster H, O'Donovan CB, Woolhead C, Macready AL, Fallaize R, Navas-Carretero S, San-Cristobal R, et al. Design and baseline characteristics of the Food4Me study: a web-based randomised controlled trial of personalised nutrition in seven European countries. Genes Nutr 2015;10:450.
15. Bingham SA, Gill C, Welch A, Cassidy A, Runswick SA, Oakes S, Lubin R, Thurnham DI, Key TJA, Roe L, et al. Validation of dietary assessment methods in the UK arm of epic using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin c and carotenoids as biomarkers. Int J Epidemiol 1997;26: S137–51.
16. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires—a review. Public Health Nutr 2002;5:567–87.
17. Scientific Advisory Committee on Nutrition. [Internet] London: The Stationery Office; 2011; Dietary reference values for energy 2011 [cited 2016 Feb 19]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/339317/SACN_Dietary_Reference_Values_for_Energy.pdf.
18. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65:1220S–8S.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
20. Henry CJK. Basal metabolic rate studies in humans: measurement and development of new equations. Public Health Nutr 2005;8:1133–52.
21. Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, Prentice AM. Critical-evaluation of energy-intake data using fundamental principles of energy physiology: 1. Derivation of cutoff limits to identify under-recording. Eur J Clin Nutr 1991;45:569–81.
22. Hébert JR, Peterson KE, Hurley TG, Stoddard AM, Cohen N, Field AE, Sorensen G. The effect of social desirability trait on self-reported dietary measures among multi-ethnic female health center employees. Ann Epidemiol 2001;11:417–27.
23. Vereecken CA, De Bourdeaudhuij I, Maes L, Grp HS. The Helena online food frequency questionnaire: reproducibility and comparison with four 24-h recalls in Belgian-Flemish adolescents. Eur J Clin Nutr 2010;64:541–8.

24. Johansson L, Solvoll K, Opdahl S, Bjorneboe GEA, Drevon CA. Response rates with different distribution methods and reward, and reproducibility of a quantitative food frequency questionnaire. Eur J Clin Nutr 1997;51:346–53.
25. Beasley JM, Davis A, Riley WT. Evaluation of a web-based, pictorial diet history questionnaire. Public Health Nutr 2009;12:651–9.
26. Tsubono Y, Nishino Y, Fukao A, Hisamichi S, Tsugane S. Temporal change in the reproducibility of a self-administered food frequency questionnaire. Am J Epidemiol 1995;142:1231–5.
27. Filippi AR, Amodio E, Napoli G, Breda J, Bianco A, Jemni M, Censi L, Mammina C, Tabacchi G. The web-based ASSO-food frequency questionnaire for adolescents: relative and absolute reproducibility assessment. Nutr J 2014;13:119.
28. Labonté ME, Cyr A, Baril-Gravel L, Royer MM, Lamarche B. Validity and reproducibility of a web-based, self-administered food frequency questionnaire. Eur J Clin Nutr 2012;66:166–73.
29. Nurul-Fadhilah A, Teo PS, Foo LH. Validity and reproducibility of a food frequency questionnaire (FFQ) for dietary assessment in Malay adolescents in Malaysia. Asia Pac J Clin Nutr 2012;21:97–103.
30. Chen CM, Huang Z-P, Zhang X, He L, Willett W, Wang J-L, Hasegawa K, Chen J-S. Reproducibility and validity of a Chinese food frequency questionnaire. Biomed Environ Sci 2010;23:1–3.
31. Wong JE, Parnell WR, Black KE, Skidmore PML. Reliability and relative validity of a food frequency questionnaire to assess food group intakes in New Zealand adolescents. Nutr J 2012;11:65.
32. Sarmento RA, Antonio JP, Riboldi BP, Montenegro KR, Friedman R, de Azevedo MJ, de Almeida JC. Reproducibility and validity of a quantitative FFQ designed for patients with type 2 diabetes mellitus from southern Brazil. Public Health Nutr 2014;17:2237–45.
33. Kesse-Guyot E, Castetbon K, Touvier M, Hercberg S, Galan P. Relative validity and reproducibility of a food frequency questionnaire designed for French adults. Ann Nutr Metab 2010;57:153–62.
34. Johansson L, Solvoll K, Bjorneboe GEA, Drevon CA. Under- and overreporting of energy intake related to weight status and lifestyle in a nationwide sample. Am J Clin Nutr 1998;68:266–74.
35. Pfrimer K, Vilela M, Resende CM, Scagliusi FB, Marchini JS, Lima NKC, Moriguti JC, Ferriolli E. Under-reporting of food intake and body fatness in independent older people: a doubly labelled water study. Age Ageing 2015;44:103–8.
36. Lutomski JE, van den Broeck J, Harrington J, Shiely F, Perry IJ. Sociodemographic, lifestyle, mental health and dietary factors associated with direction of misreporting of energy intake. Public Health Nutr 2011;14:532–41.