Public Health Nutrition: 20(15), 2660–2669 doi:10.1017/S1368980017001641

First evaluation steps of a new method for dietary intake estimation regarding a list of key food groups in adults and in different sociodemographic and health-related behaviour strata

Luis M Béjar*
Department of Preventive Medicine and Public Health, School of Medicine, University of Seville, Institute of Anatomy, 3rd floor, Sánchez-Pizjuán Avenue, Seville 41009, Spain

Submitted 22 February 2017: Final revision received 2 June 2017: Accepted 9 June 2017: First published online 9 August 2017

Abstract

Objective: A new method known as 'current-day dietary recall' (current-day recall) is based on an application for mobile phones called 'electronic 12 h dietary recall' (e-12HR). This new method was designed to rank participants into categories of habitual intake regarding a series of key food groups. The present study compared current-day recall against a previously validated short paper FFQ.
Design: Participants recorded the consumption of selected food groups using e-12HR during twenty-eight consecutive days and then filled out a short paper FFQ at the end of the study period. To evaluate the association and agreement between both methods, Spearman's correlation coefficients (SCC), cross-classification analysis and weighted kappa statistics (κ_w) were used.
Setting: Andalusia, Spain, Southern Europe.
Subjects: University students and employees over the age of 18 years.
Results: One hundred and eighty-seven participants completed the study (64·2 % female, 35·8 % male). For all particpants, for all food group intakes, the mean SCC was 0·70 (SCC ≥ 0·62 were observed for all strata); the mean percentage of participants cross-classified into categories of 'exact agreement + adjacent' was 90·1 % (percentages ≥ 87·8 % were observed for all strata); and the mean κ_w was 0·55 ($\kappa_w \geq 0{\cdot}53$ in ten of the twelve strata).
Conclusions: For the whole sample and for all strata thereof, the current-day recall has good agreement with the previously validated short paper FFQ for assessing food group intakes, rendering it a useful method for ranking individuals.

Keywords
Dietary assessment
Mobile phone application
24 h dietary recall
FFQ
Epidemiological methods

For most epidemiological applications, long-term (usual) intake is the conceptually relevant exposure[1–3]. A central feature of the long-term dietary intake of free-living individuals is variation from day to day in the foods and beverages (hereafter referred to as 'foods') that are ingested. Therefore, a single day, or a small number of days, provides a poor estimate of a person's true long-term dietary intake, but this estimate can be improved by using the average of multiple days of data for that person[1–5].

Short-term methods, such as food records and 24 h dietary recalls, allow for the collection of data that include all foods consumed by a person during a specific day. In order to determine usual dietetic intake (the long-term mean consumption of foods), it would also be necessary to repeat these measures multiple times on different days[5,6]. These repetitions would increase both the time and the amount of work for study participants, which may lead to deviations from their habitual intake and low rates of participation and compliance[7]. For this practical reason, in reality, it is rarely possible to measure a large number of days of dietary intake for an individual using these short-term tools[2,4,8]. In addition, daily variation in individual dietary intake is largely due to true day-to-day variation, but also has a component of errors in the measurement of food intake on a given day[2]. In fact, the weaknesses of food records and 24 h dietary recalls are well documented[2,4,6,8–10]. Because short-term methods are generally unrepresentative of habitual intake if only one or a few days are assessed, investigators have sought alternative methods for measuring long-term dietary intake (long-term methods), such as FFQ[2]. FFQ allow information to be collected regarding the consumption of a list of foods or food groups over prolonged periods of time (weeks or months), but these questionnaires do not take day-to-day intrapersonal variation into account. Furthermore, FFQ are subject to known measurement

*Corresponding author: Email lmbprado@us.es

errors[2,4,6,8–10]. Therefore, improvement of self-reporting that contributes to greater precision in the measurement of habitual dietary intake would represent a considerable boon for researchers[11], as well as for society as a whole, considering the important repercussions that the results and conclusions of studies in nutritional epidemiology can have on the general population[12].

The focus of the present study was not on the characterization of complete diet, but on the assessment of the consumption of key food groups[13–16]. For numerous epidemiological investigations, a detailed estimation of total food intake may lead to the collection of much redundant data, and this substantial collection of redundant information places an unnecessary burden on research participants and uses already scarce research resources[17]. However, the classification of individuals into categories of habitual intake of specific food groups is adequate to assess the relationships between relative rankings and non-communicable diseases[2,4,11,13,15,18–20] and to evaluate the effectiveness of personalized measures that are implemented to promote changes in dietary patterns with respect to selected food groups[2,11,13,15,18].

Traditional self-reporting instruments that evaluate dietary intake need to be replaced by new solutions or nutritional research and treatments for nutritional problems will remain restricted and deficient[21]. The use of mobile technologies could provide new opportunities for the design and development of more accurate tools to classify individuals according to the habitual consumption of selected food groups. For that, these new tools would allow data on the consumption of selected food groups over a large number of days to be collected for an individual, but at the same time with a reduced burden of time/work for him/her. The widespread use of the Internet on mobile phones in Spain (88·2 % of all Spaniards have accessed the Internet using their mobile phones in the last 3 months)[22] could facilitate the introduction of new methods of evaluation of usual dietary intake that include mobile technologies and the Internet[7].

When researching groups with regard to dietary habits, it is important to be aware of factors influencing the accuracy of collected data. Factors that could possibly affect reporting accuracy in dietary assessment are gender, age, education level, occupation, socio-economic position, BMI, health-related behaviours and psychological factors[7,16,23,24]. These factors may moderate accuracy in a number of ways. For example, BMI may impact social desirability and self-reported dietary intake (under-reporting or over-reporting); age may impact recall; and literacy may impact the ability of the respondent to understand and interpret dietary instruments[24].

The research team has developed a new method known as 'current-day dietary recall' (current-day recall), which is based on an application for mobiles phones called 'electronic 12 h dietary recall' (e-12HR). The aim of the current study was to compare the current-day recall against a previously validated short paper FFQ and verify that the consumption data regarding a list of key food groups from both methods are comparable, in the whole sample and in different sociodemographic and health-related behaviour strata thereof (age group, gender, occupation, smoking status, physical activity status and BMI).

Methods

Study participants

The comparison of the current-day recall with a previously validated short paper FFQ was conducted in two centres, as described by Bejar *et al.*[12]: the Medical and Pharmaceutical Schools at the University of Seville (Andalusia, Spain, Southern Europe). Different events were organized in which the research team personally presented the project to the students and employees from both faculties. Of the 291 individuals who were interested, 207 were eligible. One hundred and ninety-four individuals decided to participate in the study and, of these, 187 completed both the application e-12HR and the short paper FFQ.

The period of participant recruitment spanned from October 2014 to December 2016. Participants were recruited to the study during the entire period of research, so that all days of the week and all seasons were included in the sample.

The inclusion criteria were the following: (i) being over 18 years of age; (ii) being a University of Seville student or employee from the Medical or Pharmaceutical School; and (iii) owning a mobile phone with access to the Internet and an Android operating system.

All of the personal data collected in the study remained anonymous and confidential and were treated according to current Spanish legislation[25]. To that end, each participant was assigned a personal alphanumeric code, so that no one, not even the research team, could link personal information to the results obtained.

The e-12HR mobile phone application

Participants downloaded the application e-12HR (previously called e-EPIDEMIOLOGY) to their personal mobile phones. The structure and functions of the application have been described by Bejar *et al.* in detail[12]. In brief, this application permitted the recording of each participant's daily consumption of a list of key food groups. This list consisted of twelve items which referred to ten different food groups: fruit, vegetables, legumes, chicken/turkey, fish, red meat (lamb, beef and pork), soft drinks, sweets, prepared foods and alcoholic beverages (see online supplementary material, Supplemental File 1, for the twelve screen captures of the application e-12HR). These key food groups were selected for the study because they provide consumption patterns ranging from nearly daily to sporadic, for the general population[1]. They were also considered to be markers for healthy

https://doi.org/10.1017/S1368980017001641 Published online by Cambridge University Press

(fruit, vegetables, legumes and fish) and unhealthy (soft drinks, sweets and prepared foods) dietary habits[26]. At the end of each recording day, after all meals were consumed, the participant could access the application and register the number of standard portions/servings that had been consumed during that day. After finishing the task on e-12HR, the data were automatically saved and sent to the research administrator's website via Wi-Fi or 3G/4G, after which time the user could not change his/her answers to the questionnaire. Participants were instructed in the use of e-12HR with a personal demonstration of how to use the application, as well as in estimation of standardized portion/serving sizes, and were reminded to maintain their habitual diet. The recording of food group intake was to be completed during twenty-eight consecutive days using the application.

The application can be modified (varying the food groups selected, standardized portions/servings, the time periods to be recorded by the application, etc.) to accommodate different sociocultural needs around the globe.

The current-day recall uses e-12HR and was designed to rank participants into categories of habitual intake regarding a series of key food groups, without recording the complete diet[13–16,27] (see the 'Codification of data' subsection). The current-day recall was designed as a dietary recall yet includes elements of an FFQ, as recording is focused on a selected list of food groups. The current-day recall is basically a simplified dietary recall which is completed at the end of the day and repeated many times (once per day) during the 28 d study period[12].

The application used to register daily consumption of selected food groups was based on a semi-quantitative short FFQ previously validated for the Spanish population[28].

As a reference, a short paper FFQ was filled out at the end of the study period, through personal interviews and at the convenience of the participants (see online supplementary material, Supplemental File 2, for a copy of the short paper FFQ). All participants completed the short paper FFQ within 1–4 d of finalizing the e-12HR application, with the exception of one participant who did so after 8 d. The short paper FFQ that was utilized was based on the same semi-quantitative short FFQ previously validated for the Spanish population[28].

All participants completed a questionnaire during these personal interviews, in which sociodemographic (date of birth, gender and occupation) and anthropometric data, as well as smoking and physical activity status, were collected. The anthropometric data were collected using a standard procedure (with these data, BMI (kg/m^2) was calculated using categories defined by the WHO)[29]. Also, during the personal interviews, the participants were asked how much time, on average, was necessary to complete the application each day. Participants could choose from one of the following options: approximately (1/2/3/4/5 or more) minutes per day. One hundred and sixty-eight participants (89·8 %) selected the option 'approximately 1 minute per day' and the remaining nineteen (10·2 %) chose 'approximately 2 minutes per day'. Thus, the time necessary to complete e-12HR was about 1 min/d.

Both the questionnaires used in the application and the short paper FFQ had the same items (see online supplementary material, Supplemental Files 1 and 2), the only difference being that in e-12HR, the questionnaire refers to daily consumption during twenty-eight consecutive days[12] while the short paper FFQ refers to consumption over the previous 28 d[13,18,30–32]. In order to make comparisons about the usefulness of each method, it was desirable to keep food records during the same period of time with each tool[7,30,33]. The period of time selected (28 d) is similar to that of previous comparison/validation studies[13,18,30–32].

Revision of data

Data collected from e-12HR were saved without modifications in a database. Subsequently, three sets of data were removed due to obvious inconsistencies: one participant had registered the consumption of 200 standardized portions of legumes in one day, another had registered the consumption of 250 standardized servings of soft drinks in one day, and another had registered the consumption of eighty-eight standardized servings of alcoholic beverages in one day.

The data collected from the short paper FFQ were manually introduced without modifications into a separate database by the research team.

Codification of data

For the e-12HR: for each of the ten key food groups, the data from the 28 d of e-12HR use were recorded as daily consumption. These data were later used to calculate the average consumption of said items in that period.

For the short paper FFQ: for the same food groups, the frequency of consumption was categorized into six categories ranging from 'less than once a week' to '3 times or more a day' (see online supplementary material, Supplemental File 2).

To compare the data from the e-12HR and the short paper FFQ, and taking into account that the current-day recall was designed to rank participants into categories of habitual intake, the data from the e-12HR were transformed to include them in one of the same categories of habitual consumption from the FFQ. For example, suppose that a participant consumes an average of 0·45 standard rations of vegetables daily during 28 d using e-12HR. This average consumption represents 3·15 standard portions per week ($0{\cdot}45 \times 7 = 3{\cdot}15$), which would be classified in the category '3–4 times a week.' This was made possible because both the short paper FFQ and e-12HR used the same standardized portion/serving sizes (see online supplementary material, Supplemental Files 1 and 2).

https://doi.org/10.1017/S1368980017001641 Published online by Cambridge University Press

Statistical analysis

The association between dietary intake methods (current-day recall and short paper FFQ) was assessed using Spearman's correlation coefficients (SCC)[34]. Cohen's cut-offs were used to interpret the SCC values. According to these cut-offs, $r = \pm 0{\cdot}5$ is considered strong, $r = \pm 0{\cdot}30$ moderate and $r = \pm 0{\cdot}10$ is weak[35].

The relative agreement between the two methods was assessed using cross-classification analysis[34]. The participants were classified by the two methods into quintiles of 'exact agreement', 'exact agreement + adjacent', 'slight disagreement', 'strong disagreement' and 'extreme disagreement'[30–32] (see Table 3 and online supplementary material, Supplemental Table 1).

The inter-rater agreement of two assessment methods was analysed by the weighted kappa statistic (κ_w)[34], assigning partial credit to scores using the Stata prerecorded weights[36]: if there was complete agreement, a weight of 1·00 was assigned; for cases cross-classified into adjacent categories, 0·80 was assigned; for cases cross-classified two categories apart, 0·60; for three categories apart, 0·40; for four categories apart, 0·20; and for cases cross-classified into extreme categories, 0·00. Values of κ_w over 0·80 indicate very good agreement; between 0·61 and 0·80, good agreement; beween 0·41 and 060, moderate agreement; between 0·21 and 0·40, fair agreement; and <0·20, poor agreement[37].

All statistical analysis was performed using the statistical software package Stata version MP 13.1 and a $P < 0{\cdot}05$ was considered statistically significant[36].

Table 1 Characteristics of participants in the study; university students and employees over the age of 18 years, Andalusia, Spain, Southern Europe, October 2014 to December 2016

Characteristic	*n*	%	Mean	SD
Participants who completed the study	187	100·0		
Number of days completed through the e-12HR				
28 d	123	65·8		
26 d	31	16·6		
25 d	9	4·8		
24 d	14	7·5		
23 d	4	2·1		
22 d	2	1·1		
20 d	4	2·1		
Age (years)			28·2	10·9
Age group				
<25 years	103	55·1		
≥25 years	84	44·9		
Gender				
Female	120	64·2		
Male	67	35·8		
Occupation				
Student	119	63·6		
Employee	68	36·4		
Smoking status				
No	151	80·7		
Yes	36	19·3		
Physical activity status (min/week)				
≥150	60	32·1		
<150	127	67·9		
BMI (kg/m^2)			23·6	4·9
Underweight	10	5·3		
Normal range	121	64·7		
Overweight	42	22·4		
Obese	14	7·5		

e-12HR, electronic 12-h dietary recall.

Results

One hundred and ninety-four individuals participated in the study, but seven participants did not complete the application and the FFQ. These individuals' data were not used for posterior analysis.

Information on the number of days completed with the e-12HR application can be consulted in Table 1.

Among the participants, the mean age was 28·2 years; 55·1 % were <25 years old and 44·9 % were ≥25 years old; 64·2 % were females and 35·8 % were males; 63·6 % were students and 36·4 % were employees; 19·3 % were smokers and 80·7 % were non-smokers; and 32·1 % had a physical activity status of 150 min or more/week[38] and 67·9 % had less than 150 min/week[38]. The mean BMI was 23·6 kg/m^2, with 5·3 % of the participants in the underweight range (BMI < 18·5 kg/m^2), 64·7 % in the healthy weight range (BMI = 18·5–24·9 kg/m^2), 22·4 % being overweight (BMI = 25·0–29·9 kg/m^2) and 7·5 % obese (BMI ≥ 30·0 kg/m^2; Table 1).

No significant statistical differences were found in any of the variables studied between the participants who completed the study and those who did not.

For all particpants, for all food group intakes, the mean SCC was 0·70 (by strata, ranging from 0·62 (≥25 years) to 0·75 (<25 years); Table 2).

For all participants, for all food group intakes, the mean percentage of individuals in categories of 'exact agreement' was 56·6 % (by strata, ranging from 47·6 % (employees) to 63·9 % (<25 years)); the mean percentage of participants cross-classified into categories of 'exact agreement + adjacent' was 90·1 % (by strata, ranging from 87·8 % (employees) to 91·8 % (<25 years)); the mean percentage of participants misclassified was 9·8 % ('slight disagreement' (7·6 %), 'strong disagreement' (2·2 %) and 'extreme disagreement' (0·0 %)). 'Slight disagreement', by strata, ranged from 6·6 % (males) to 8·7 % (≥25 years). 'Strong disagreement', by strata, ranged from 1·4 % (<25 years and students) to 3·7 % (employees; Table 3 and see the online supplementary material, Supplemental Table 1, for full details).

For all participants, for all food group intakes, the mean κ_w value was 0·55 (by strata, ranging from 0·45 (≥25 years) to 0·63 (<25 years); Table 4).

Discussion

The present study demonstrates the development of a new method for the assessment of habitual dietary intake, called current-day recall, which is based on e-12HR, and

https://doi.org/10.1017/S1368980017001641 Published online by Cambridge University Press

Table 2 Spearman correlatuion coefficients derived from e-12HR and the short paper FFQ, overall and according to sociodemographic and health-related behaviour strata; university students and employees over the age of 18 years, Andalusia, Spain, Southern Europe, October 2014 to December 2016

		Sociodemographic strata					
		Age group (years)		Gender		Occupation	
Comparison	All	<25	≥25	Female	Male	Student	Employee
Fruit	0·81	0·83	0·79	0·78	0·86	0·81	0·81
Vegetables	0·76	0·86	0·63	0·77	0·75	0·84	0·61
Legumes	0·55	0·60	0·49	0·52	0·60	0·57	0·51
Chicken/turkey	0·65	0·70	0·52	0·67	0·61	0·67	0·53
Fish	0·62	0·75	0·47	0·64	0·61	0·72	0·43
Red meat	0·68	0·71	0·66	0·63	0·67	0·69	0·68
Soft drinks	0·76	0·79	0·71	0·76	0·76	0·77	0·75
Sweets	0·69	0·78	0·57	0·67	0·73	0·73	0·60
Prepared foods	0·61	0·65	0·56	0·53	0·73	0·61	0·60
Alcoholic beverages	0·82	0·86	0·78	0·76	0·80	0·85	0·77
Mean	0·70	0·75	0·62	0·67	0·71	0·73	0·63
		Health-related behaviour strata					
		Smoking		Physical activity (min/week)		BMI (kg/m^2)	
Comparison	All	No	Yes	≥150	<150	<25	≥25
Fruit	0·81	0·82	0·72	0·78	0·81	0·78	0·87
Vegetables	0·76	0·77	0·72	0·83	0·72	0·74	0·79
Legumes	0·55	0·54	0·54	0·65	0·49	0·55	0·54
Chicken/turkey	0·65	0·63	0·77	0·68	0·65	0·70	0·53
Fish	0·62	0·61	0·66	0·53	0·68	0·60	0·67
Red meat	0·68	0·65	0·77	0·80	0·63	0·68	0·68
Soft drinks	0·76	0·73	0·84	0·78	0·75	0·76	0·75
Sweets	0·69	0·72	0·57	0·78	0·64	0·70	0·67
Prepared foods	0·61	0·62	0·56	0·63	0·60	0·60	0·62
Alcoholic beverages	0·82	0·78	0·80	0·80	0·83	0·84	0·77
Mean	0·70	0·69	0·70	0·73	0·68	0·70	0·69

e-12HR, electronic 12-h dietary recall.
$P<0·001$ for all data.

Table 3 Cross-classification analysis* derived from e-12HR and the short paper FFQ, overall and according to sociodemographic and health-related behaviour strata; university students and employees over the age of 18 years, Andalusia, Spain, Southern Europe, October 2014 to December 2016

		Sociodemographic strata					
		Age group (years)		Gender		Occupation	
Comparison	All	<25	≥25	Female	Male	Student	Employee
Exact agreement†	56·6	63·9	47·7	55·8	58·2	61·8	47·6
Exact agreement + adjacent‡	90·1	91·8	88·1	90·0	90·3	91·5	87·8
Slight disagreement§	7·6	6·8	8·7	8·3	6·6	7·1	8·5
Strong disagreement‖	2·2	1·4	3·2	1·8	3·1	1·4	3·7
Extreme disagreement¶	0·0	0·0	0·0	0·0	0·0	0·0	0·0
		Health-related behaviour strata					
		Smoking		Physical activity (min/week)		BMI (kg/m^2)	
Comparison	All	No	Yes	≥150	<150	<25	≥25
Exact agreement†	56·6	56·7	59·2	58·3	55·8	57·4	54·8
Exact agreement + adjacent‡	90·1	90·5	88·9	90·7	89·8	90·4	89·5
Slight disagreement§	7·6	7·5	8·3	7·0	8·0	7·6	8·0
Strong disagreement‖	2·2	2·1	2·8	2·3	2·2	2·1	2·5
Extreme disagreement¶	0·0	0·0	0·0	0·0	0·0	0·0	0·0

e-12HR, electronic 12-h dietary recall.
*Data pesented are mean agreement (%) for ten different food groups: fruit, vegetables, legumes, chicken/turkey, fish, red meat (lamb, beef and pork), soft drinks, sweets, prepared foods and alcoholic beverages.
†Exact agreement: cases cross-classified into the same category.
‡Exact agreement + adjacent: cases cross-classified into the same or adjacent category.
§Slight disagreement: cases cross-classified two categories apart.
‖Strong disagreement: cases cross-classified three or four categories apart.
¶Extreme disagreement: cases cross-classified into extreme categories.

https://doi.org/10.1017/S1368980017001641 Published online by Cambridge University Press

Table 4 Weighted kappa statistics derived from e-12HR and the short paper FFQ, overall and according to socio-demographic and health-related behaviour strata; university students and employees over the age of 18 years, Andalusia, Spain, Southern Europe, October 2014 to December 2016

		Sociodemographic strata					
		Age group (years)		Gender		Ocupation	
Comparison	All	<25	≥25	Female	Male	Student	Employee
Fruit	0·67	0·70	0·64	0·64	0·73	0·67	0·67
Vegetables	0·62	0·71	0·51	0·62	0·62	0·68	0·50
Legumes	0·45	0·56	0·33	0·42	0·50	0·52	0·34
Chicken/turkey	0·53	0·60	0·40	0·55	0·50	0·58	0·41
Fish	0·51	0·66	0·34	0·50	0·53	0·61	0·35
Red meat	0·57	0·63	0·49	0·55	0·51	0·59	0·51
Soft drinks	0·58	0·63	0·49	0·57	0·59	0·61	0·51
Sweets	0·49	0·59	0·35	0·45	0·57	0·55	0·35
Prepared foods	0·50	0·57	0·41	0·42	0·61	0·54	0·43
Alcoholic beverages	0·60	0·68	0·52	0·54	0·60	0·67	0·50
Mean	0·55	0·63	0·45	0·53	0·58	0·60	0·46
		Health-related behaviour strata					
		Smoking		Physical activity (min/week)		BMI (kg/m^2)	
Comparison	All	No	Yes	≥150	<150	<25	≥25
Fruit	0·67	0·69	0·57	0·70	0·65	0·64	0·74
Vegetables	0·62	0·63	0·56	0·68	0·58	0·59	0·66
Legumes	0·45	0·44	0·50	0·50	0·43	0·45	0·45
Chicken/turkey	0·53	0·52	0·59	0·57	0·51	0·58	0·42
Fish	0·51	0·49	0·59	0·38	0·58	0·52	0·48
Red meat	0·57	0·55	0·59	0·65	0·52	0·58	0·52
Soft drinks	0·58	0·55	0·64	0·62	0·55	0·55	0·62
Sweets	0·49	0·50	0·44	0·57	0·45	0·49	0·47
Prepared foods	0·50	0·51	0·45	0·53	0·49	0·52	0·46
Alcoholic beverages	0·60	0·57	0·61	0·56	0·62	0·61	0·56
Mean	0·55	0·55	0·55	0·58	0·54	0·55	0·54

e-12HR, electronic 12-h dietary recall.
$P < 0·0001$ for all data.

compares this method with a previously validated short paper FFQ, in the whole sample and in different socio-demographic and health-related behaviour strata thereof (age group, gender, occupation, smoking status, physical activity status and BMI).

For all particpants, for all food group intakes, the mean SCC was 0·70 (high correlations were observed for all strata, SCC ≥ 0·62). Cross-classification analysis showed that 56·6 % of the participants were correctly classified into the same category (percentages ≥ 47·6 % were observed for all strata) and 90·1 % were classified into categories of 'exact agreement + adjacent' (percentages ≥ 87·8 % were observed for all strata). Just 2·2 % were misclassified into categories of 'strong disagreement' (percentages ≤ 3·7 % were observed for all strata) and 0·0 % were misclassified into an opposite category. For all participants, for all food group intakes, the mean κ_w value of 0·55 was moderate (by strata, the values showed moderate agreement for all strata ($\kappa_w = 0·45$–$0·60$), except <25 years with a good agreement ($\kappa_w = 0·63$); $\kappa_w \geq 0·53$ in ten of the twelve strata). The data collected through both methods could have been analysed on a continuous scale (using another type of statistical analysis). However, as the current-day recall has been designed to rank individuals according to their habitual intake of selected food groups rather than to assess their absolute level of intake, the research team preferred to analyse the collected data by organizing them into categories[12]. Cross-classification analysis and κ_w are dependent on the number of categories used. In order to limit this dependence when evaluating agreement and misclassification, the six original categories could have been reorganized into three[34]. In any case, the research team preferred to use the six original categories for the analysis instead of three, since the presentation of information classified in the original six categories provides more compact and precise information on the capability of both methods to assign individuals according to the distribution of food consumption, when compared with reorganizing the information using only three categories.

In the comparison of both methods (current-day recall and short paper FFQ), the results are satisfactory in the sample groups and strata. However, lower values in the different analyses that were used (SCC, cross-classification analysis and κ_w) were found among the age group of ≥25 years and employees. A large proportion of employees are aged 25 years or older. Taking this into account, there are a number of possible reasons that explain the values observed: on the one hand, age can affect memory[24];

https://doi.org/10.1017/S1368980017001641 Published online by Cambridge University Press

on the other, as compared with adults, young people and adolescents are more comfortable and efficient with the use of mobile technologies, and they have expressed their preference for these methods[39,40]; also, in another study performed with university students, the results suggested that there was no fatigue effect when using an application for collection of food consumption during 3 weeks[40], an effect which was indeed noticed among older participants in a 28 d monitoring study.

The ability of a method to discriminate among individuals is most directly evaluated by comparing individual estimates of food group intake based on this method with those measured by another more accurate method; that is, a gold standard. But there is no perfect measure of dietary intake, with the implication that validation studies are not possible[2–4,6,8–10]. Thus, validation studies never compare an operational method with absolute truth; rather, they compare one method with another method that is judged to be superior. Given that neither method is perfect, it is crucial that errors from both methods be as independent (i.e. uncorrelated) as possible to avoid spuriously high estimates of validity[2,6,7,11,13,41]. For this reason, from the outset the research team discounted using a 24 h dietary recall as a reference method in the present study, since the current-day recall is basically a simplified dietary recall. Among the available comparison methods for validating a dietary recall (such as the current-day recall), food records are likely to have the least correlated errors. However, the research team did not consider the option of using a food record as a reference method in the present study. Validation of a long-term method utilizing a short-term method is challenging when the reference method does not accurately reflect the usual food intake. In addition, a food group that is not consumed on a daily basis is more critical when episodically consumed foods are related and compared (the key food groups in e-12HR were selected for analysis because they provide consumption patterns ranging from nearly daily to sporadic)[1]. When used as a standard to assess the validity of the current-day recall method, food records should, in principle, be kept for a sufficient number of days to represent average intake and cover the interval of time corresponding to the method being evaluated (28 d in the present study). The process of keeping a food record multiple times is both burdensome and time-consuming for study participants and habitual food intake may be altered as a result; to the extent that this is a departure from usual food habits, this will tend to reduce the correlation between the two methods. Other problems with food records are low compliance and low participation rates in dietary studies[2,4–9]. Additionally, keeping a food record will heighten awareness of food and thus might increase the accuracy in completing the e-12HR application[2,7,11,33]. An alternative to the use of the food records as a standard for evaluating the current-day recall may be an FFQ. Because errors are more likely to be correlated with this method (both rely on memory and perception of portion/serving sizes), it is probably suboptimal. However, in many situations, such as when participants are less than highly motivated or the amount of work for study participants may be excessive (for example, e-12HR during a period of 28 d and three to seven weighed or estimated food records over the same time period), FFQ may be the only reasonable option[2,30]. Moreover, FFQ may provide a more realistic instrument to assess long-term intake because they also capture infrequently consumed foods; while short-term instruments like food records have presumably less bias, they must be repeated many times[42]. Therefore, although not an established reference method, the research team considered the previously validated short paper FFQ to be an appropriate reference in this first evaluation of the current-day recall[26,30].

While the current-day recall demonstrated good agreement with the reference method (for the whole sample and for all strata thereof), some disagreement was observed between the two tools: cross-classification analysis showed that 9·8 % of the participants were incorrectly classified by two to four categories (by strata, ranging from 8·2 % (<25 years) to 12·2 % (employees)). Several factors must be taken into account. On the one hand, both methods present some similarities: (i) the same difficulties in the precise estimation of portion/serving size, although it has been reported that frequency of consumption seems to have a greater impact on dietary intake than portion sizes[2]; and (ii) the same difficulties in the interpretation of questions, since both methods use the same questions to measure the frequency of consumption. For example, both ask: 'How many portions of vegetables have you eaten? (1 portion = approx. 150 g)'. However, there are important differences between both methods. First, with e-12HR, this question is answered at the end of each day during the study period, while the short paper FFQ is completed at the end of 28 d. This minimizes dependence on the memory of the participant in e-12HR (this method relies on short-term memory) in comparison to the FFQ (this method relies on long-term memory), keeping in mind that the recollection of past consumption of foods can be influenced by more recent food consumption[2,4,8,9]. Second, the current-day recall is not limited by day-to-day variability in dietary intake and may accurately assess intakes of foods that are eaten infrequently. This day-to-day variability interferes with the precise determination of habitual dietary intake[43], especially in the case of FFQ, where data are collected only once at the end of an extended time period.

One inherent limitation to most FFQ is that they are paper-based forms. Thus, errors such as skipped questions or multiple marks are common. Furthermore, data from paper forms must be entered into analysis software. Web-based FFQ offer straightforward solutions to these limitations of paper FFQ[11,19]. In their most simple application, paper FFQ match web-based FFQ[6,44]; this allows the

https://doi.org/10.1017/S1368980017001641 Published online by Cambridge University Press

flexibility of using either a paper or a computerized questionnaire interchangeably, but the benefits from computer administration are limited to direct data entry, real-time error checking and rapid analysis[41]. Other advantages include reducing paper waste, postage costs, and the space, security and organization required for paper file storage[45]. FFQ may be administered according to the needs of the study and the target population. The research team considered that, in the present study, the potential disadvantages of developing a web-based FFQ, in comparison with a paper-based FFQ, outweighed its potential benefits, keeping in mind two inherent study characteristics: (i) the paper-based FFQ used is very short and simple (containing only twelve items); and (ii) the sample is made up of students and employees from the Medical and Pharmacy Schools at the University of Seville. The simplicity of the short paper FFQ reduced the chance for errors, the amount of paper consumed and storage space issues. The relatively easy access to the sample population made it possible to complete the short paper FFQ in person, making it unnecessary to mail it. In this case, the costs associated with data entry were minimal compared with the potential costs of developing a web-based FFQ[12]. In recent years, many well-established FFQ have been developed into web-based versions and there is a growing body of evidence demonstrating that data from web-based FFQ are comparable with data from printed versions[30].

Since use of the paper FFQ, instead of food record or 24 h dietary recall, was the preferred method of reference in the present study, the research team considered this to be an evaluation study rather than a validation study of the current-day recall. Once the results indicate, for the whole sample and for all strata thereof, that the current-day recall has reasonable ranking ability for selected food group intake estimates[34] and is highly comparable with the previously validated short paper FFQ[30], the research team will plan a validation study in which both methods (current-day recall and short paper FFQ) will be compared with 3–7 d weighed or estimated food records[30,31], although assuming lower compliance and participation rates in this future study. This will help more thoroughly evaluate the potential validity of e-12HR as a research tool for habitual intake estimation of key food groups.

Strengths and limitations

The strengths of the present study include using more than one statistical method in order to give credence to the results[34]. In addition, the recording of food group intake was to be completed during twenty-eight consecutive days using the application and at the end of this period using the short paper FFQ, minimizing the likelihood of changes in dietary intake[7,30,46]. Repeated applications of traditional short-term instruments, such as food records and 24 h dietary recalls, can modify habitual intake due to the excessive workload for participants[7]. Despite repeated use, the modification of habitual intake seems unlikely through use of the e-12HR, due to the reduced workload that using this application presents (1 min/d).

The small number of individuals in some of the subgroups is one limitation of the study; for example, non-smokers (n 36). Other study limitations are that the population was highly educated, which may limit the generalizability of study results to other populations. Additionally, although the use of an FFQ was the preferred choice of reference method, it introduces several limitations discussed above, such as difficulty in the precise estimation of portion/serving size and reliance on memory. It is important to emphasize that the high association and agreement between the current-day recall and the validated short paper FFQ scores do not indicate that the current-day recall is 'accurate' since there is no 'true' measure of dietary intake[1]. Ideally, validation studies would include the use of nutritional biomarkers, but currently few biomarkers exist for specific foods[47].

In future studies of the current-day recall, a third version of e-12HR will be used (the second version is currently in use) which includes several improvements, such as an adaptation to iOS (which will help increase the sample size) and the inclusion of photographs to help participants estimate portion size[12].

Conclusions

The fact that the results in the present study differed between strata indicates that there can be no single measure of agreement of a given method for all subjects. The good agreement with the reference method (for the whole sample and for all strata thereof) indicates the utility of the current-day recall for ranking individuals according to their consumption of the food groups selected for the study. The current-day recall can be a useful tool to analyse possible associations with risks for chronic diseases and for evaluating the effects of intervention studies, when the characterization of complete diet is not required. With the growing popularity of mobile phones among Spaniards, this instrument is likely to be accepted by the population. However, future studies should explore the potential validity of the current-day recall in more representative samples and employ 3–7 d weighed or estimated food records as a reference method.

Acknowledgements

Acknowledgements: The author would like to thank the participants in this study. *Financial support*: This work was partly supported by funding from the University of Seville's Research Plan. The University of Seville had no role in the design, analysis or writing of this article. *Conflict of interest:* None. *Authorship:* The sole author had responsibility for all parts of the manuscript. *Ethics of human subject*

participation: This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving human subjects were approved by the Research Ethics Committee at the University of Seville. Written informed consent was obtained from all subjects.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1368980017001641

References

1. Carroll RJ, Midthune D, Subar AF *et al.* (2012) Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. *Am J Epidemiol* **175**, 340–347.
2. Willett W (editor) (2013) *Nutritional Epidemiology*, 3rd ed. New York: Oxford University Press.
3. Stumbo PJ (2013) New technology in dietary assessment: a review of digital methods in improving food record accuracy. *Proc Nutr Soc* **72**, 70–76.
4. Rutishauser IH (2005) Dietary intake measurements. *Public Health Nutr* **8**, 1100–1107.
5. Shamah-Levy T, Rodríguez-Ramírez S, Gaona-Pineda EB *et al.* (2016) Three 24-hour recalls in comparison with one improve the estimates of energy and nutrient intakes in an urban Mexican population. *J Nutr* **146**, 1043–1050.
6. Tucker KL, Smith CE, Lai CQ *et al.* (2013) Quantifying diet for nutrigenomic studies. *Annu Rev Nutr* **33**, 349–371.
7. Svensson A & Larsson C (2015) A mobile phone app for dietary intake assessment in adolescents: an evaluation study. *JMIR Mhealth Uhealth* **3**, e93.
8. Gibson R (editor) (2005) *Principles of Nutritional Assessment*, 2nd ed. New York: Oxford University Press.
9. Martín-Moreno JM & Gorgojo L (2007) Assessment of dietary intake at the population level through individual questionnaires: methodological shadows and lights. *Rev Esp Salud Publica* **81**, 507–518.
10. Dhurandhar NV, Schoeller D, Brown AW *et al.* (2015) Energy balance measurement: when some thing is not better than nothing. *Int J Obes (Lond)* **39**, 1109–1113.
11. Kristal AR, Kolar AS, Fisher JL *et al.* (2014) Evaluation of web-based, self-administered, graphical food frequency questionnaire. *J Acad Nutr Diet* **114**, 613–621.
12. Bejar L, Sharp B & García-Perea M (2016) The e-EPIDEMIOLOGY mobile phone app for dietary intake assessment: comparison with a food frequency questionnaire. *JMIR Res Protoc* **5**, e208.
13. Cleghorn CL, Harrison RA, Ransley JK *et al.* (2016) Can a dietary quality score derived from a short-form FFQ assess dietary quality in UK adult population surveys? *Public Health Nutr* **19**, 2915–2923.
14. Branscum P, Sharma M, Kaye G *et al.* (2010) An evaluation of the validity and reliability of a food behavior checklist modified for children. *J Nutr Educ Behav* **42**, 349–352.
15. Saeedi P, Skeaff SA, Wong JE *et al.* (2016) Reproducibility and relative validity of a short food frequency questionnaire in 9–10 year-old children. *Nutrients* **8**, E271.
16. Saloheimo T, González SA, Erkkola M *et al.* (2015) The reliability and validity of a short food frequency questionnaire among 9–11-year olds: a multinational study on three middle-income and high-income countries. *Int J Obes Suppl* **5**, Suppl. 2, S22–S28.
17. Giabbanelli PJ & Adams J (2016) Identifying small groups of foods that can predict achievement of key dietary recommendations: data mining of the UK National Diet and Nutrition Survey, 2008–12. *Public Health Nutr* **19**, 1543–1551.
18. Buch-Andersen T, Pérez-Cueto FJ, Toft U *et al.* (2016) Relative validity and reproducibility of a parent-administered semi-quantitative FFQ for assessing food intake in Danish children aged 3–9 years. *Public Health Nutr* **19**, 1184–1194.
19. Knudsen VK, Hatch EE, Cueto H *et al.* (2016) Relative validity of a semi-quantitative, web-based FFQ used in the 'Snart Forældre' cohort – a Danish study of diet and fertility. *Public Health Nutr* **19**, 1027–1034.
20. Medin AC, Carlsen MH & Andersen LF (2016) Associations between reported intakes of carotenoid-rich foods and concentrations of carotenoids in plasma: a validation study of a web-based food recall for children and adolescents. *Public Health Nutr* **19**, 3265–3275.
21. Archer E, Hand GA & Blair SN (2013) Validity of US nutritional surveillance: National Health and Nutrition Examination Survey caloric energy intake data, 1971–2010. *PLoS One* **8**, e76632.
22. National Statistics Institute (2016) Survey on the Equipment and Use of Information and Communication Technologies (ICT-H) in Households, year 2016. http://www.ine.es/jaxi/Datos.htm?path=/t25/p450/base_2011/a2016/l0/&file=04011.px (accessed February 2017).
23. Tabacchi G, Filippi AR, Breda J *et al.* (2015) Comparative validity of the ASSO-Food Frequency Questionnaire for the web-based assessment of food and nutrients intake in adolescents. *Food Nutr Res* **59**, 26216.
24. Alcantara I, Haardörfer R, Gazmararian JA *et al.* (2015) Relative validation of fruit and vegetable intake and fat intake among overweight and obese African-American women. *Public Health Nutr* **18**, 1932–1940.
25. Spanish Government (1999) Law 15/1999, of Protection of Personal Data. https://www.boe.es/boe/dias/1999/12/14/pdfs/A43088-43099.pdf (accessed February 2017).
26. Henriksson H, Bonn SE, Bergström A *et al.* (2015) A new mobile phone-based tool for assessing energy and certain food intakes in young children: a validation study. *JMIR Mhealth Uhealth* **3**, e38.
27. Vereecken CA, Rossi S, Giacchi MV *et al.* (2008) Comparison of a short food-frequency questionnaire and derived indices with a seven-day diet record in Belgian and Italian children. *Int J Public Health* **53**, 297–305.
28. Rodríguez IT, Ballart JF, Pastor GC *et al.* (2008) Validation of a short questionnaire on frequency of dietary intake: reproducibility and validity. *Nutr Hosp* **23**, 242–252.
29. World Health Organization (1998) *Obesity: Preventing and Managing the Global Epidemic. Report of a WHO Consultation on Obesity, Geneva, 3–5 June 1997*. Geneva: WHO; available at http://whqlibdoc.who.int/hq/1998/WHO_NUT_NCD_98.1_(p1-158).pdf
30. Forster H, Fallaize R, Gallagher C *et al.* (2014) Online dietary intake estimation: the Food4Me food frequency questionnaire. *J Med Internet Res* **16**, e150.
31. Fallaize R, Forster H, Macready AL *et al.* (2014) Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. *J Med Internet Res* **16**, e190.
32. Marshall SJ, Livingstone KM, Celis-Morales C *et al.* (2016) Reproducibility of the online Food4Me food-frequency questionnaire for estimating dietary intakes across Europe. *J Nutr* **146**, 1068–1075.
33. Rangan AM, O'Connor S, Giannelli V *et al.* (2015) Electronic Dietary Intake Assessment (e-DIA): comparison of a mobile phone digital entry app for dietary data collection with 24-hour dietary recalls. *JMIR Mhealth Uhealth* **3**, e98.
34. Masson LF, McNeill G, Tomany JO *et al.* (2003) Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutr* **6**, 313–321.

35. Cohen J (editor) (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York: Lawrence Erlbaum Associates.
36. StataCorp LP (2015) *Stata Statistical Sofware: Release 13.1*. College Station, TX: StataCorp LP; available at http://www.stata.com.
37. Viera AJ & Garrett JM (2005) Understanding interobserver agreement: the kappa statistic. *Fam Med* **37**, 360–363.
38. World Health Organization (2011) Global Recommendations on Physical Activity for Health: 18–64 years old. http://www.who.int/dietphysicalactivity/physical-activity-recommendations-18-64years.pdf?ua=1 (accessed May 2017).
39. Daugherty BL, Schap TE, Ettienne-Gittens R *et al.* (2012) Novel technologies for assessing dietary intake: evaluating the usability of a mobile telephone food record among adults and adolescents. *J Med Internet Res* **14**, e58.
40. Hongu N, Pope BT, Bilgiç P *et al.* (2015) Usability of a smartphone food picture app for assisting 24-hour dietary recall: a pilot study. *Nutr Res Pract* **9**, 207–212.
41. Bassett JK, English DR, Fahey MT *et al.* (2016) Validity and calibration of the FFQ used in the Melbourne Collaborative Cohort Study. *Public Health Nutr* **19**, 2357–2368.
42. Gerdessen JC, Souverein OW, van 't Veer P *et al.* (2015) Optimising the selection of food items for FFQs using Mixed Integer Linear Programming. *Public Health Nutr* **18**, 68–74.
43. Casperson SL, Sieling J, Moon J *et al.* (2015) A mobile phone food record app to digitally capture dietary intake for adolescents in a free-living environment: usability study. *JMIR Mhealth Uhealth* **3**, e30.
44. Matthys C, Pynaert I, De Keyzer W *et al.* (2007) Validity and reproducibility of an adolescent web-based food frequency questionnaire. *J Am Diet Assoc* **107**, 605–610.
45. Falomir Z, Arregui M, Madueno F *et al.* (2012) Automation of food questionnaires in medical studies: a state-of-the-art review and future prospects. *Comput Biol Med* **42**, 964–974.
46. Watanabe M, Yamaoka K, Yokotsuka M *et al.* (2011) Validity and reproducibility of the FFQ (FFQW82) for dietary assessment in female adolescents. *Public Health Nutr* **14**, 297–305.
47. Hedrick V, Dietrich A, Estabrooks P *et al.* (2012) Dietary biomarkers: advances, limitations and future directions. *Nutr J* **11**, 109.

https://doi.org/10.1017/S1368980017001641 Published online by Cambridge University Press