British Journal of Nutrition, page 1 of 27 doi:10.1017/S0007114520003220

Relative validity of FFQ to assess food items, energy, macronutrient and micronutrient intake in children and adolescents: a systematic review with meta-analysis

Luisa Saravia[1,2]*, Maria L. Miguel-Berges[1], Iris Iglesia[1,3], Marcus V. Nascimento-Ferreira[1,4], Guillermo Perdomo[2], Isabel Bove[5], Betzabeth Slater[6] and Luis A. Moreno[1]

[1]*GENUD (Growth, Exercise, Nutrition and Development) Research group, Instituto Agroalimentario de Aragón (IA2), Instituto de Investigación Sanitaria Aragón (IIS Aragón), Edif. del SAI (Servicio de Apoyo a la Investigación), Universidad de Zaragoza, Zaragoza 50009, Spain*
[2]*Escuela de Nutrición, Universidad de la República, Montevideo 11600, Uruguay*
[3]*Red de Salud Materno-infantil y del Desarrollo (SAMID), 5009 Zaragoza, Spain*
[4]*Youth/Child Cardiovascular Risk and Environmental (YCARE) Research Group, School of Medicine, University of Sao Paulo, São Paulo, CEP 01246-904, Brazil*
[5]*School of Psychology, Catholic University, Montevideo 11600, Uruguay*
[6]*Public Health Program, School of Public Health, University of Sao Paulo, Sao Paulo, CEP 01246-904, Brazil*

(Submitted 28 January 2020 – Final revision received 6 August 2020 – Accepted 12 August 2020)

Abstract

FFQ are one of the most widely used tools of research into nutritional epidemiology, and many studies have been conducted in several countries using this dietary assessment method. The present study aimed to evaluate the relative validity of FFQ, in comparison with other methods, in assessing dietary intake of children and adolescents, through a systematic review. Four electronic databases (Embase, PubMed, Scopus and Web of Science) found sixty-seven articles, which met the inclusion criteria (healthy children and adolescents from 3 to 18 years of age; journal articles written in English, Spanish and Portuguese between 1988 and March 2019; results showing the comparison between the FFQ with other methods of assessment of dietary intake). The articles were analysed by two independent reviewers. A meta-analysis was conducted using correlation coefficients as estimate effects between the FFQ and the reference standard method. Subgroup analysis and meta-regression were performed to identify the probable source of heterogeneity. In fifty-five of the sixty-seven studies, a single dietary assessment method was used to evaluate the FFQ; nine combined the two methods and three used three reference methods. The most widely used reference method was the 24-h recall, followed by the food record. The overall relative validity of the FFQ to estimate energy, macronutrient, certain micronutrient and certain food item intakes in children and adolescents may be considered weak. The study protocol was registered in PROSPERO under number CRD42016038706.

Key words: **FFQ: Dietary intake: Assessment**

Research interest in dietary aetiologies of chronic diseases has in the last few decades stimulated the development and validation of methods for dietary assessment for use in epidemiological studies[1]. In children and adolescents, a limited number of dietary assessment instruments have been found to be reproducible and valid[2]. When children and adolescents are the target population in dietary surveys, different considerations of respondents and observers must be taken into account. The cognitive abilities required to self-report food intake include an adequately developed concept of time, a good memory and attention span, and knowledge of the names of foods[3]. Studies on children's recall of food intake indicate that the instruments are susceptible to considerable error[4–10], including under-reporting, over-reporting and the incorrect identification of foods[11]. From the age of 8 years, there is a rapid increase in the ability of children to self-report food intake[12]. However, while cognitive abilities should be fully developed by adolescence, issues of motivation and body image may hinder a willingness to report[3].

Abbreviations: 24Hr, 24-h record; FR, food record; WFR, weight food record; DLW, doubly labelled water; CHO, carbohydrates; PICOS, population, intervention, comparator, outcome and setting; I^2, heterogeneity value; EI, energy intake.

* **Corresponding author:** Profesora Agregada Luisa Saravia, fax +598 96714664, email saravialuisa@gmail.com

After further refinement, revision and evaluation during the 1980s and 1990s, FFQ became one of the most widely used tools of research into nutritional epidemiology[13], and many epidemiological studies have been conducted in several countries around the world using this dietary assessment method[14–20]. The underlying principle of the FFQ approach is that average long-term diet, for example, intake over weeks, months or years, is conceptually relevant exposure rather than intake over a few specific days[1,21]. Therefore, the purpose of the FFQ is not only to quantify food or nutrient intake but also to classify individuals according to their food intake. For the relative validation of FFQ, different reference standard methods have been used. They include other dietary assessment tools, such as 24-h records (24Hr), food records (FR), weight food records (WFR) or biomarkers, and the doubly labelled water (DLW) method, in the case of energy intake (EI)[1].

The 24Hr is used to describe the average dietary intake of groups of individuals; participants are asked to recall and describe in detail and in an open-ended manner about the foods and beverages consumed over 1 d, preferably the day before, in detail and depth. This method requires a trained interviewer to ask the respondent to remember in detail all the foods and beverages they consumed during the previous 24 h[21,22]; usually, several 24Hr are used to capture within person variability.

The FR is a method that consists of a specially designed booklet, or a mobile app, in which individuals list every food and beverage consumed. This estimation is made using household measurements, such as cups or spoons, food photographs, or food models. FR can be completed over several days[22]. The limitations of this method are that individuals may change their eating behaviour during the time the recording lasts, the method does not take into account the long-term variety of consumption, possible changes in food habits and participants' burden, which can result in incomplete FR[1].

The WFR is similar to the FR method, except that the quantification of foods and beverages is by weight, taken and recorded by the participants rather than estimated[22]. Another method to describe the dietary intake of individuals is the observation method , which consists of training observers to estimate the types and amounts of food and beverages commonly served to groups of persons[23].

Nutritional biomarkers may be used as indicators of dietary exposure; therefore, any biochemical characteristic associated with the exposure that can be measured objectively can be used as a nutritional marker. More commonly, compounds found in foods – and their metabolites – are used as biomarkers, for example, serum-25, hydroxy vitamin D, vitamin C, vitamin E, retinol, carotenoids, skin carotenoids, urinary N, Ca and K excretions[24–29], although physical properties, such as stable isotope ratios, are also suitable[30].

Measurement of total energy expenditure using the DLW method has proven to be a useful tool with which to test the validity of EI measurements, based on the premises that, in subjects who are in energy balance, total EI is equivalent to total energy expenditure. This approach is limited to the validation of total EI rather than specific macronutrient and micronutrient intake[31].

All these methods have advantages and disadvantages for assessing individual or populational dietary intake. In epidemiological studies, where the objective is to assess different aspects of the nutritional status of the population, researchers must normally choose one of these methods because of the financial and time costs[1].

To address the measurement error associated with dietary questionnaires, large-scale epidemiological investigations often integrate substudies for the validation and calibration of the questionnaires and/or to administer a combination of different assessment methods (e.g. administration of different questionnaires and assessment of biomarker levels)[2,3,13,21,22,32].

In children and adolescents, FFQ have been used in several studies in Latin America[12,26,33–37] and all over the world[38–43], and some of them previously performed a relative validation of the specifically used tool.

However, to our knowledge, no systematic review to date has compared the relative validity of FFQ in relation with reference methods in children and adolescents. Systematic reviews with meta-analysis have the strength of increased statistical power for primary outcomes, the ability to reach agreement when original studies yield conflicting findings, improving effect size estimates and answering questions not addressed in original trials[44]. Therefore, the aim of the present study was to evaluate the relative validity of FFQ in assessing the dietary intake of children and adolescents, comparing the questionnaire with other forms of evaluating food consumption through a systematic review, considering energy, macronutrients (carbohydrates (CHO), protein, fat and fibre), certain micronutrients (Ca, Fe, Zn, vitamin A and vitamin C) and some food item (meat, milk, fruits and vegetables). We chose these foods and nutrients among others because of their relevance to children and adolescents nutrition and because they are more frequently studied.

Methods

This systematic review followed the protocol of Preferred Reporting Items for Systematic Reviews and Meta-Analyses[45] and the methodology suggested by the Academy of Nutrition and Dietetics for the evidence analysis process in 2016[46] is shown in Table 1[44].

The evidence analysis question was to compare the relative validity of the FFQ with reference methods of assessing energy, nutrients and/or food intake in healthy children and adolescents. The study protocol was registered in PROSPERO (http://www.crd.york.ac.uk/PROSPERO) under number CRD42016038706[44].

The PICOS (population, intervention, comparator, outcome and setting) criteria were used to perform the systematic review and are shown in Table 2.

Search strategy and eligibility criteria

The search strategy was undertaken to identify all literature in English, Spanish and Portuguese, published between 1988 and October 2014. In March 2019, a second search was conducted using the same databases and the same keywords to update the previous search. The reason to start in 1988 is

British Journal of Nutrition

Table 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist

Section/topic	No.	Checklist item	Reported on page no.
Title			
Title	1	Identify the report as a systematic review, meta-analysis or both	1
Abstract			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number	1
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	1–3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes and study design (PICOS)	3
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g. web address), and, if available, provide registration information including registration number	4
Eligibility criteria	6	Specify study characteristics (e.g. PICOS, length of follow-up) and report characteristics (e.g. years considered, language, publication status) used as criteria for eligibility, giving rationale	4
Information sources	7	Describe all information sources (e.g. databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	4
Study selection	9	State the process for selecting studies (i.e. screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	4–6
Data collection process	10	Describe method of data extraction from reports (e.g. piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	4
Data items	11	List and define all variables for which data were sought (e.g. PICOS, funding sources) and any assumptions and simplifications made	Table 4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level) and how this information is to be used in any data synthesis	Table 5
Summary measures	13	State the principal summary measures (e.g. risk ratio, difference in means)	4–6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g. I^2) for each meta-analysis	5–6

PICOS, population, intervention, comparator, outcome and setting.

Table 2. PICOS (population, intervention, comparator, outcome and setting) of the systematic review

Population	Healthy children and adolescents aged 3–18 years
Intervention	Use of FFQ for ranking dietary intake of individuals, assessing energy, nutrients and/or food intake
Comparator	Other methods of dietary assessment (24 Hr, FR, WFR, biomarkers, DLW and observation)
Outcome	Meat, milk, fruits, vegetables, energy, CHO, protein, fat, fibre, Fe, Ca, Zn, vitamin A and vitamin C
Setting	Not applicable

24 Hr, 24-h record; FR, food record; WFR, weight food record; DLW, doubly labelled water; CHO, carbohydrates.

because in 1989, Willett[1] published his book Epidemiological Nutrition with very exhaustive research and analysis regarding FFQ and made his proposal of a questionnaire that was then widely used for many researchers. Articles were retrieved from four electronic databases: Embase, PubMed, Scopus and Web of Science. Keywords and their combinations, using Boolean commands, were used to retrieve as many papers as possible from the databases comprehensively. Included keywords were validation studies, diet surveys, questionnaires, diet records, FFQ, usual food intake, nutrition assessment, diet, dietary patterns, biomarkers, reproducibility, validity, child and adolescent. All references were downloaded from the computerised bibliographic software Refworks 2.0, provided by the University of Zaragoza (Spain). A sample query for PubMed is included in Table 3.

The systematic search process is illustrated in Fig. 1. In the first search, 11 097 papers were identified and, after removing the duplicates, 5362 titles of articles were analysed by two independent reviewers and 4841 articles were excluded. The reviewers considered as relevant to this review those papers comparing the FFQ with other methods of dietary assessment and if they were conducted in healthy children and adolescents aged 3–18 years. Identified systematic review articles were excluded after searching the bibliography for included relevant papers. The reviewers screened 521 abstracts. At first, eighty-one papers were analysed due to their relevance in relation to the research question and fifty-one articles fulfilled the inclusion criteria of this revision. Reference lists of identified studies and related reviews were hand-searched for relevant articles. The most common reasons to exclude original research papers was that the studies did not compare the FFQ with other methods of assessment of dietary intake (nine articles), they were systematic reviews (nine articles), participants were not the population in the study, for example, children under 3 years old, (ten articles) and the investigation was not published as a journal

Table 3. Example of the database search strategy, PubMed

Search ID#	Search terms	Search details	Results
1	diet records, food frequency questionnaires, recall, diet assessment method, children, toddler, adolescent, teenager	((("diet* record*"[tiab] OR "food frequency questionnaire*"[tiab] OR "recall*"[tiab] OR "diet* assessment method*"[tiab]))) AND (Child*[tiab] OR toddler*[tiab] OR Adolescent*[tiab] OR teenager*[tiab] NOT adult*[tiab] NOT pregnant*[tiab])	4145

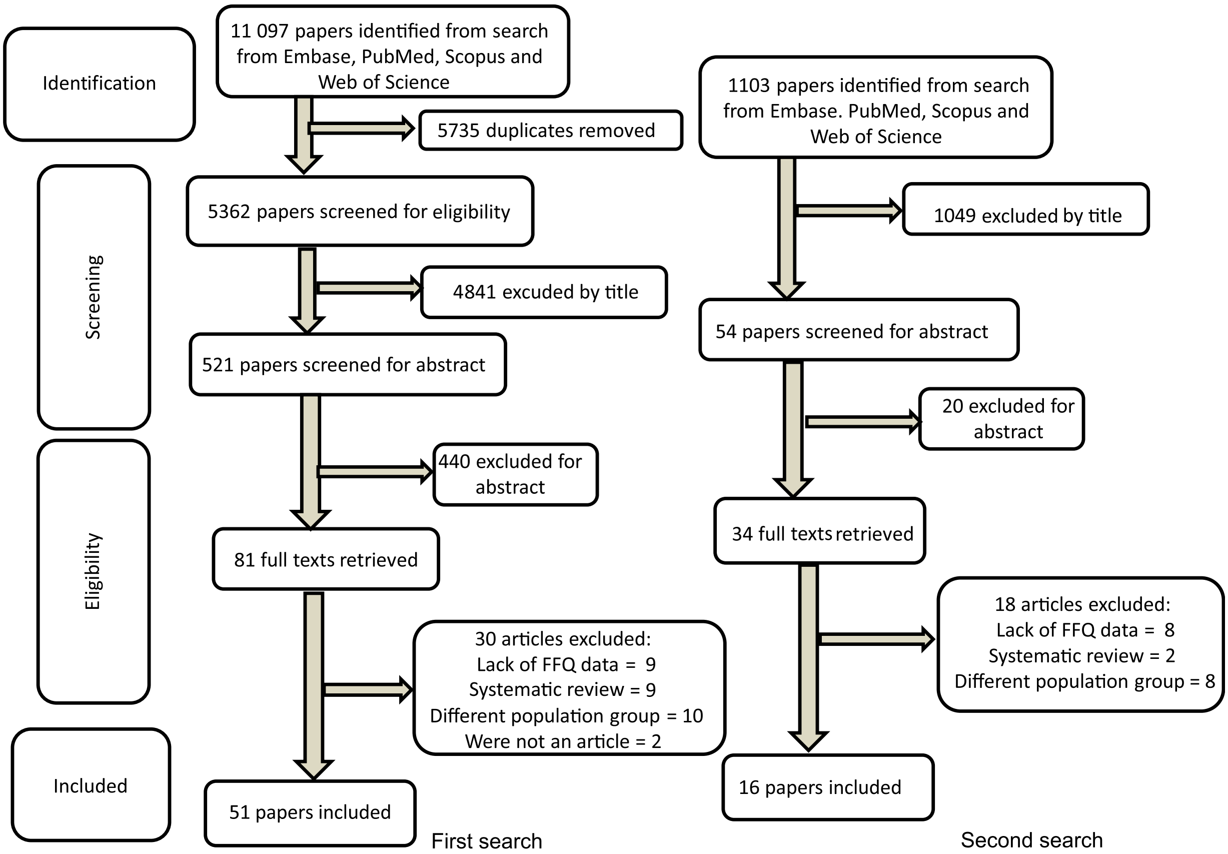

Fig. 1. Systematic research process of FFQ validity.

article (two: one a chapter in a book and the other a conference abstract). In case of different opinions between the two researchers, the discrepancies were discussed until a consensus was reached. In March 2019, using the same methodology, sixteen new articles were added and sixty-seven articles were finally included in the study.

Data management and data extraction

From the selected articles, data were extracted into a table by the two independent reviewers. The summarised information was the reference standard method used to validate the FFQ, the aim and a brief description of the study, the characteristics of the population, a brief description of the statistical approach (correlation coefficients), and the results and conclusions of the study.

Quality assessment

All articles were independently assessed for quality by the two independent reviewers, using a standardised quality assessment checklist proposed by the Academy of Nutrition and Dietetics in 2012[47] to critically review research articles and grade the strength of evidence. All the papers qualified for this systematic review.

After a brief analysis of the results, it was decided to perform meta-analysis, including all those studies providing correlation coefficients (Pearson or Spearman) between the FFQ and the reference standard method. For the statistical tests, the α value was set at 0·05. Meta-analysis provides estimates of effect size[48]; the type of the effect size depends on the nature of the outcome (in this case, the magnitude of the association between methods). This facilitates the comparison of studies, irrespective of units of measurement or measurement scales. Meta-analysis provided a weighted average of the results of the individual studies in which the weight of the study depends on its precision. Meta-analysis was undertaken where the results could be quantitatively combined; they were performed using R Core Team (2017), a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria[49]. To confirm the correlation coefficient (*r*), Pearson correlation or Spearman's rank, the cut-off points were defined using Cohen's classification: less than 0·3, poor; 0·3–0·5, fair; 0·6–0·8, moderately strong; and at least 0·8, very strong[50,51].

Because of the high heterogeneity value (I^2) obtained in the meta-analysis, a meta-regression analysis was performed to assess the relationship between the FFQ and other reference methods for energy, macronutrients (CHO, protein, fat and fibre), micronutrients (Ca, Fe, Zn, vitamin A and vitamin C) and some food categories (meat, milk, vegetables and fruits)[49].

Finally, thirty-seven studies providing correlation coefficients, comparing the FFQ with another dietary assessment method (24Hr, FR and WFR) to estimate energy, CHO, protein, fat, fibre, Ca, Fe, Zn, vitamin A, vitamin C, meat, milk, fruit and vegetable intake were included in the meta-analysis[25,28,29,33,35,37,40,41,43,52–79].

For biomarkers and DLW, meta-analysis was not performed owing to the limited number of studies and because they often did not provide correlation coefficients[24–29,31,79–81].

The model chosen was the random effect meta-analysis model, in which the effect sizes in the included studies are assumed to represent a random sample from a particular distribution of these effect sizes[82]. A meta-regression analysis was performed to identify the probable source of heterogeneity.

Results

General description of studies included in the systematic review

Table 4 includes a summary of the main results from all the papers included in the systematic review[24–29,31,33–38,40,41,43,52–54,56–66,68–81,83–104]. Finally, sixty-seven articles met the inclusion criteria for the systematic review. Twenty-six studies showed results obtained in European countries (Belgium, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden, the UK and Multicentric Projects in Europe)[25,28,29,43,54,57,58,61,67,71,75,79–81,84,88,90,91,94–98,100,101], twenty-two presented data from American countries (the USA, Brazil, Bolivia, Guatemala and Peru)[24,26,27,31,33–37,56,60,62,72–74,76,87,89,92,93,102,103], nine studies were from Asian countries (Korea, Japan, Lebanon, Malaysia, China and Vietnam)[38,40,41,52,63–66,69], nine from Oceanian countries (Australia and New Zealand)[53,59,70,77,78,83,86,99,104] and one study included results from Colombia, Finland and the USA[85].

Thirty-eight (57 %) of the studies were published after 2010, and twenty-nine (43 %) between 1989 and 2009.

In the included studies, the number of participants ranged from 22 to 10 309. As the systematic review included children and adolescents, to describe the studies, a decision was made to classify them into three groups: thirty-seven studies focused on children (2 to ≤12 years old)[24,28,29,31,33,35,38,40,53,57–59,61,62,65,66,68,71,74,76–81,85,88,90–92,94–96,99,101,102,104], seventeen studies were on adolescents (12–19 years old)[26,34,36,43,52,54,56,63,64,67,75,84,86,98,100,103] and thirteen studies on both groups[25,27,37,41,60,70,72,73,83,87,89,93,105].

In twenty-five (37 %) studies, the FFQ respondents were the caregivers[24,28,29,31,33,35,37,40,57,59,62,65,76,78–81,86,88,90,91,94,99,101,104], thirty-five (55 %) were older children or adolescents[26,28,30,32,34,36,37,39–41,44,46–55,61,64–66,69–71,75,77–79,81,82,85,86,88], and five (8 %) were caregivers and children or adolescents[27,38,61,66,92].

Thirty-seven (55 %) of the FFQ were quantitative[24,25,27,28,38,41,54,59–61,63–66,69,71,73,75,77–79,81,84,87,88,91–93,95,99–103,105], twenty-three (34 %) were semiquantitative[26,29,31,33–37,40,43,52,58,70,72,74,76,80,83,86,88,90,94,96] and seven (11 %) were qualitative[53,56,57,62,85,89,98].

The number of food items in each FFQ ranged from 5 to 227, depending on the nutrient or nutrients being measured. The frequencies of the intake categories were variable, as two (3 %) of the studies reported to use eleven categories[58,66]; three (4 %) reported to use ten categories[29,78,96]; five (7 %) studies used nine categories[28,31,37,89,91]; ten (15 %) studies used eight categories[36,38,52,57,62,65,68,95,100]; and ten (15 %) studies used seven categories[26,35,43,53,56,61,64,71,85,97]. Nine (13 %) studies used six categories[69,77,81,84,87,88,94,98,103]; three (4 %) studies used five categories[79,86,90]; two (3 %) studies used four categories[34,92]; and two (3 %) studies used three categories[40,88]. In sixteen (24 %) of the studies, the used categories were not shown[24,25,27,33,41,54,59,60,63,74–76,80,93,104,106] and in five (7 %) of the studies the categories varied for each food item[69,71,82,100,101]. Fifteen (22 %)[28,31,37,40,54,56,61,63,74,86,88,90,91,94,99] studies used 1 year as the target period that the respondent was asked to recall; eight (12 %)[26,34,35,40,64,69,75,82] studies used 6 months as the target period; three (4 %)[24,29,96] used 3 months as the target period; one (1 %)[79] study used 2 months as the target period; fifteen (22 %)[38,57,62,66,69,72,75,78,80,81,84,89,92,93,100] studies used 1 month as the target period; twelve (18 %)[27,33,34,43,60,73,77,87,98,102,103,104] studies used less than 1 month; and in thirteen (19 %)[25,52,53,58,59,64,67,68,71,85,95,101,105] studies, the target period was not shown.

Reference methods

The number of studies using the different relative validation approaches is reported in Fig. 2. Most studies (fifty-seven of sixty-seven) used a single dietary assessment method to compare the FFQ; eight studies combined two methods and two used three reference methods. The most widely used reference method was the 24Hr, twenty-three studies (35 %)[33,35,40,41,54,56,57,60,62–64,72–74,76,77,84,87,91,93,101,103,104]; followed by the FR, twenty-two studies (33 %)[33,36,37,51,52,57,58,60,64,68,69,70,82,84,85,87,89,93,97,98,107]. Others methods were used in a minor numbers of studies (32 %)[24–29,31,43,66,69,75,78–81,89,92,95–97,100,102].

Table 4. Description of the main findings of the systematic review

	Author, country, year	No. of participants – age group	Reference method	Outcomes	No. of of food items	Category frequencies	Respondent of the FFQ	Quantitative or qualitative	Target period	Results
1	Kunaratnam, *et al.*, Australia, 2018[105]	2–5 year (*n* 62)	3 FR	F	17	Not shown	Caregivers	Quantitative	2 weeks	There was poor agreement between FFQ–DQI and 3-d FR–DQI scores (r 0·36, $P<0·01$). The 3-d-FR–DQI scores were positively associated with CHO, β-carotene, Ca, protein, total fat and negatively associated with sugar, starch and vitamin C.
2	Leong *et al.*, New Zealand, 2018[78]	5 years (*n* 100)	3 WFR	F & N	123	10	Caregivers	Quantitative	4 weeks	Mean correlations between the FFQ and WDR were acceptable for nutrient and food group intakes (r 0·34 and r 0·41, respectively).
3	Rodriguez *et al.*, Peru, 2017[37]	0–14 years (*n* 120)	6 FR	N	150	9	Caregivers	Semiquantitative	1 year	Age-adjusted correlations among children aged 0–7 years were highest for vitamin C (0·66), total fat (0·67), and lowest for retinol (–0·06). High correlations were observed in children aged 8–14 years who participated in FFQ administration with their caregiver (*n* 23) or alone (*n* 1). Ca (0·54) performed well, while nutrients such as total CHO (–0·30) and Zn (–0·05) had lower correlations. Age-adjusted correlations among children aged 8–14 years was substantially lower when the caregiver responded to the FFQ alone, with most ranging from –0·50 to 0·10.
4	Söderberg *et al.*, Sweden, 2017[29]	5–7 years (*n* 85)	3 FR + 2 blood extractions for biomarkers	N	16	10	Caregivers	Semiquantitative	3 months	The correlation between all three instruments was moderate to strong. SFFQ2 and the 3D record correlated moderately to S-25(OH)D. Bland–Altman analysis showed that Ca was underestimated by on average of 29 mg/d, (LOA 808 and –865 mg/d).
5	Nyström *et al.*, Sweden, 2017[79]	5·5 (SD 0·1) years (*n* 38)	2 weeks DLW + 4 × 24Hr	F & N	Between 42 and 86	5	Caregivers	Quantitative	2 months	The mean value of energy intake (EI) calculated using the FFQ was statistically different from TEE.
6	Vioque *et al.*, Spain, 2016[28]	4–5 years (*n* 169)	3 × 24Hr, 1 blood extraction for biomarkers	F & N	105	9	Caregivers	Quantitative	1 year	The correlation for validity of the FFQ compared with the 24Hr and biomarkers was higher than 0·20.

Table 4. (*Continued*)

	Author, country, year	No. of participants – age group	Reference method	Outcomes	No. of of food items	Category frequencies	Respondent of the FFQ	Quantitative or qualitative	Target period	Results
7	Yum & Lee, Korea, 2016[52]	12–18 years (*n* 153)	8 × FR	N	71	8	Adolescents	Semiquantitative	Not shown	Spearman's correlation coefficients ranged from 0·27 (for vitamin A and fibre) to 0·90 (for energy). The proportions of subjects in the opposite categories between the first FFQ and the 8-d food record data were generally low within the range of 0·74 % (for energy and CHO) to 13·2 % (for β-carotene)
8	Rahmawaty *et al.*, Australia, 2016[83]	9–13 years (*n* 22)	7 × FR	N	131	Varied depending on the food item	Children	Semiquantitative	6 months	Bland–Altman plots showed an acceptable limit of agreement between the FFQ and the average 7-d FR
9	De Cock *et al.*, Belgium, 2016[84]	14–16 years (*n* 99)	3 × 24 Hr	F	42	6	Adolescents	Quantitative	1 month	Mean differences were small for beverage intake but large for snack intake, except for healthy snack ratio
10	Moghames *et al.*, Lebanon, 2016[40]	5–10 years (*n* 111)	4 × 24Hr	N	112	3	Caregivers	Semiquantitative	1 year	Energy and nutrient intakes estimated by the FFQ were like those obtained by 24Hr
11	Saeedi *et al.*, New Zealand, 2016[53]	9–10 years (*n* 50)	4 × FR	F	28	7	Children	Qualitative	Not shown	In validity analyses, 70 % of food items/groups had an ICC between 0·3 and 0·5
12	Fatihah *et al.*, Malaysia, 2015[38]	7–12 years (236 in Phase 1; 209 in Phase 2)	3 × FR	F & N	94	8	Caregivers and adolescents	Quantitative	1 month	Mean intake of macronutrients in FFQ1 and 3FR correlated well, although the FFQ intake data tended to be higher (20·4 %) (CC from 0·497 to 0·310)
13	Julián-Almárcegui *et al.*, Spain, 2015[54]	14·4 (SD 2·4 years (*n* 84 for swimmers; *n* 57 for controls)	2 × 24Hr	N	19	Not shown	Adolescents	Quantitative	1 year	Pearson correlations 0·52 for controls and 0·47 for swimmers after correcting for intra-variability. Cross-classification analysis 73·7 % for controls and 63·1 % of swimmers were classified correctly
14	Marcinkevage *et al.*, Guatemala, 2015[33]	Caregivers (*n* 145) of 6–11 years children	3 × 24Hr	N	108	Not shown	Caregivers	Semiquantitative	1 week	Agreement by FFQ and 24Hr ranged from 62·0 % for cholesterol to 95·9 % for vitamin B_{12} across all three FFQ
15	Saloheimo *et al.*, Colombia, Finland, USA, 2015[85]	9·9 years (*n* 321)	3 × FR	F	23	7	Children	Qualitative	Not shown	Validity CC were below 0·5 for 22/23 food groups, and they differed among country sites. For validity, gross misclassification was <5 % for 22/23 food groups. Over- or underestimation did not appear for 19/23 food groups
16	Aguilar *et al.*, USA, 2014[27]	5–17 years (*n* 50)	3 × 24Hr, three blood extraction, six palm scans for biomarkers	F & N	27	Not shown	Caregivers and adolescents	Quantitative	1 week	Each serving of averaged total F & V reported from the FFQ was 3·798 (P < 0·001) increase in RRS intensity counts

Table 4. (*Continued*)

	Author, country, year	No. of participants – age group	Reference method	Outcomes	No. of of food items	Category frequencies	Respondent of the FFQ	Quantitative or qualitative	Target period	Results
17	Appannah *et al.*, Australia, 2014[86]	14 years (*n* 783)	3 × FR	F & N	227	5	Caregivers	Semiquantitative	1 year	Pearson's CC between participants' *z*-scores for the DP identified in the FFQ and FR was 0·35 for girls and 0·49 for boys ($P < 0·05$)
18	Bel-Serrat *et al.*, Europe, 2014[113]	2–9 years (*n* 2508)	2 × 24Hr	F	43	8	Caregivers	Qualitative	1 month	The CEHQ-FFQ provided higher intake estimates than the 24Hr. De-attenuated Pearson CC ranged from 0·001 to 0·48 in children aged 2–6 years and from 0·01 to 0·44 in children aged 6–9 years
19	Flood *et al.*, Australia, 2014[59]	2–5 years (*n* 77)	3 × FR	F & N	17	Not shown	Caregivers	Quantitative	Not shown	Kappa ranged from 0·37 to 0·85. Spearman's rank CC was >0·5
20	Martinez & Estima, Brazil, 2013[56]	15–19·9 years (*n* 109)	4 × 24Hr	F & N	50	7	Adolescents	Qualitative	1 year	EI ICC = 0·53. Values for crude data were high in rice, sugars (0·71) and meat groups (0·77)
21	Pampaloni *et al.*, Italy, 2013[58]	9–10 years (*n* 75)	7 × FR	N	21	11	Children	Semiquantitative	Not shown	Mean dietary Ca intakes were 725·6 mg/d (95 % CI 683·2, 768·1) from 7 FR and 892·4 mg/d (95 % CI 844·6, 940·2) from the FFQ
22	Hunsberger *et al.*, USA, 2012[60]	10–17 years (*n* 99)	3 × 24Hr	F & N	41	Not shown	Adolescents	Quantitative	1 week	The 24Hr estimated higher levels of SAF and added sugar consumption; the de-attenuated correlations of these measures ranged from 0·478 to 0·768
23	Lillegaard *et al.*, Norway, 2012[61]	9 years (*n* 733); 13 years (*n* 904)	4 × FR	F	23	7	Adolescents	Quantitative	1 year	The median Spearman's CC between the two methods was 0·36 among the 9-year-olds and 0·32 among the 13-year-olds
24	Mulasi-Pokhriyal *et al.*, USA, 2012[87]	9–13 years (*n* 164) and 14 to 18 years (*n* 171)	2 × 24Hr	F & N	77	6	Children and adolescents	Quantitative	1 week	Among all children, Block FFQ estimates for vitamin A, vitamin C, vegetables and fruits were significantly higher than those assessed through the 24Hr ($P < 0·001$)
25	Del Pino & Friedman, Brazil, 2011[35]	6–10 years (*n* 91)	3 × 24Hr	N	90	7	Caregivers	Semiquantitative	6 months	The FFQ overestimated all nutrients. CC with the values obtained by 24Hr were mostly above 0·50
26	Huybrechts *et al.*, Belgium, 2011[80]	2–10 years (*n* 10·309), 8 European countries	Urine biomarkers	N	43	Not shown	Caregivers	Semiquantitative	1 month	Significant positive correlation between milk consumption frequencies and the ratios of K/Cr (0·16 ($P < 0·001$)) and a weaker with the ratios of UCa/Cr (0·07 ($P < 0·001$))
27	Kobayashi *et al.*, Japan, 2011[66]	3–11 years (*n* 50) and 12 years (*n* 53)	4 × WFR	N	75	11	Caregivers and adolescents	Quantitative	1 month	The correlation coefficient in sex-, age- and energy-adjusted value revealed that the largest number of subjects with high (0·50 or more) value was obtained by the CFFQ in the YC group

Table 4. (*Continued*)

	Author, country, year	No. of participants – age group	Reference method	Outcomes	No. of of food items	Category frequencies	Respondent of the FFQ	Quantitative or qualitative	Target period	Results
28	Nurul-Fadhilah *et al.*, Malaysia, 2012[63]	Adolescents (*n* 170)	3 × 24Hr	N	124	Not shown	Adolescents	Quantitative	1 year	Estimated mean intake for most nutrients assessed by the FFQ was higher as compared with the three 24Hr ($P < 0·05$)
29	Sahashi *et al.*, Japan, 2011[65]	6 years (*n* 47)	2 × FR (3 d each)	N	162	8	Caregivers	Quantitative	6 months	The validity correlation ranged from 0·05 for α tocopherol to 0·59 for retinol. The median correlation was 0·40
30	Scagliusi *et al.*, Brazil, 2011[62]	6–9 years (*n* 61)	2 × 24Hr	N	50	8	Caregivers	Qualitative	1 month	Energy-adjusted and de-attenuated CC ranged from −0·03 for vitamin C to 0·93 for Ca. The mean coefficient was 0·46
31	Xia *et al.*, China, 2011[64]	12–18 years (*n* 168)	9 × 24Hr	F & N	86	7	Adolescents	Quantitative	Not shown	The relative validity results indicate that the crude Spearman's CC of FFQ1 and the 24Hr ranged 0·41–0·65
32	Araujo *et al.*, Brazil, 2010[36,109]	12–19 years (*n* 169)	3 × FR	N	90	8	Adolescents	Semiquantitative	6 months	The Pearson CC ranged 0·33–0·46, and the mean agreement varied from 62 to 143 %
33	Dutman *et al.*, the Netherlands, 2010[81]	4–6 years (*n* 30)	7 d DLW	N	85	6	Caregivers	Quantitative	1 month	The Pearson EI:EE was 0·62
34	Huybrechts *et al.*, Belgium, 2010[88]	2·5–6·5 years (*n* 510)	3 × FR	N	47	6	Caregivers	Semiquantitative	1 year	Pearson's correlations varied among the four main components of the DQI (from 0·39 to 0·74)
35	Slater *et al.*, Brazil, 2010[26]	Adolescents (*n* 80)	2 × 24Hr + blood biomarkers	F	94	7	Adolescents	Semiquantitative	6 months	The highest validity coefficient was obtained for the vegetable group, as assessed by the FFQ (*r* 0·873)
36	Vereecken *et al.*, Belgium, 2010[104]	14·6 years (*n* 48)	4 × 24Hr	N	22	3	Adolescents	Quantitative	Not shown	Spearman's correlations between the first FFQ and the YANA-Cs were on average 0·46, with significant correlations for energy and all nutrients (≥0·32), except for the percentage of energy from fat (0·18)
37	Vereecken *et al.*, Belgium, 2010[104]	3·5 years (*n* 216)	3 × FR	F & N	14	8	Children	Quantitative	Not shown	At group level, good agreement was found for energy, fat and protein intake, an overestimation was found for CHO (5·6 %) and fibre (13·3 %), and an underestimation was found for Ca (9 %)
38	Watanabe *et al.*, Japan, 2010[69]	12–13 years (*n* 63)	7 × WFR	F & N	82	6	Adolescents	Quantitative	1 month	For validity, the CC of EI for the whole day was 0·31
39	Di Noia & Contento, USA, 2009[89]	10 – 14 years (*n* 156)	3 d observed meals	F	5	9	Adolescents	Qualitative	1 month	The 5-A-Day FFQ intake was significantly correlated with observed intake (*r* 0·39; $P < 0·01$)
40	Hong *et al.*, Vietnam, 2010[41]	11–15 years (*n* 180)	4 × 24Hr	N	160	Not shown	Adolescents	Quantitative	6 months	Coefficients for nutrient intakes between the mean of the 3 FFQ and mean of 4 24Hrs were mostly about 0·40

Table 4. (*Continued*)

	Author, country, year	No. of participants – age group	Reference method	Outcomes	No. of of food items	Category frequencies	Respondent of the FFQ	Quantitative or qualitative	Target period	Results
41	Huybrechts *et al.*, Belgium, 2009[90]	Preschool children (*n* 650)	3 × FR	F	13	5	Caregivers	Semiquantitative	1 year	The largest corrected Spearman's correlations (>0·6) were found for the intake of potatoes, grains, fruit, milk products, cheese, sugared drinks and fruit juice, while the lowest correlations (<0·4) were found for bread products and meat products
42	Stiegler *et al.*, Germany, 2009[91]	9–11 years (*n* 82)	1 × 24Hr	F & N	82	9	Caregivers	Quantitative	1 year	Reported consumption measured with the FFQ was 42 % lower for butter, 40 % lower for desserts and 39 % lower for tea than values established from the 24Hr
43	Watson *et al.*, Australia, 2009[70]	9–16 years (*n* 224)	4 × FR	N	120	Varied for each food item	Children and adolescents	Semiquantitative	6 months	Correlation coefficients for comparative validity ranged from 0·03 for retinol to 0·56 for Mg for transformed, energy-adjusted, deattenuated nutrient data, with correlation coefficients greater than 0·40 for total fat, saturated fat, monounsaturated fat, CHO, sugars, riboflavin, vitamin C, folate, β-carotene, Mg, Ca and Fe
44	Zemel *et al.*, USA, 2009[92]	7–10 years (*n* 139)	7 × WFR	N	41	4	Caregivers and children	Quantitative	1 month	Ca intake was about 300 mg/d higher by CCFFQ compared with WFR
45	Papadopoulou *et al.*, Greece, 2008[43]	15·3 years (*n* 250)	3 × WFR	F & N	54	7	Adolescents	Semiquantitative	1 week	The Pearson's coefficients ranged from 0·83 for EI to 0·34 for folate intake
46	Vereecken *et al.*, Europe, 2008[71]	11–12 years (*n* 112 for Belgian) and (*n* 114 for Italian)	7 × FR	F & N	14	7	Children	Quantitative	Not shown	Spearman's correlations between the FFQ items and the FR varied between 0·13 and 0·67
47	Harnack *et al.*, USA, 2006[93]	11–14 years (*n* 248)	3 × 24Hr + Youth Risk Behaviour Survey	N	10	Not shown	Adolescents	Quantitative	1 month	The correlation between Ca intakes was 0·43
48	Huybrechts *et al.*, Belgium, 2006[94]	2·5–6·5 years (*n* 509)	3 × FR	N	47	6	Caregivers	Semiquantitative	1 year	Mean Ca intakes were 838 (SD 305) and 777 (SD 296) mg/d for FR and FFQ respectively, indicating a mean difference of 60·9 (SD 294·4) mg/d ($P < 0·001$) Pearson's correlation was 0·52
49	Perez-Cueto *et al.*, Bolivia, 2006[34]	Adolescents (*n* 82)	3 × FR	F & N	72	4	Adolescents	Semiquantitative	1 week	Nutrient estimates obtained from the Q1 and Q2 and the 3R were not statistically different ($P > 0·05$)
50	Haraldsdóttir *et al.*, Europe, 2005[95]	11·2–11·6 years (mean) (*n* 60 for Norway), (*n* 56 for Denmark), (*n* 43 for Iceland) and (*n* 46 for Portugal)	1 × WFR + 6 × FR + 2 × 24Hr	F	5	8	Adolescents	Quantitative	Not shown	Spearman's rank correlations for F & V intake according to the FFQ part and the 7-d FR ranged between *r* 0·40–0·53. Test–retest Spearman's rank correlations for the FFQ part were *r* 0·47–0·84

Table 4. (*Continued*)

	Author, country, year	No. of participants – age group	Reference method	Outcomes	No. of of food items	Category frequencies	Respondent of the FFQ	Quantitative or qualitative	Target period	Results
51	Andersen *et al.*, Norway, 2004[96]	11·9 years (mean) (*n* 114)	2 × 24Hr + 7-d FR	F	16	10	Adolescents	Semiquantitative	3 months	Spearman's CC between the FFQ and the FR varied from 0·21 for fruit and potatoes to 0·32 for the total intake of fruit and vegetables
52	Jensen *et al.*, USA, 2004[72]	10–18 years (*n* 162)	2 × 24Hr	N	80	Varied for each food item	Adolescents	Semiquantitative	1 month	The correlation between Ca intakes as estimated by the second FFQ *v.* the average of the two 24Hr was 0·54 for the total sample
53	Vereecken & Maes, Belgium, 2003[106]	11–18 years (*n* 7072 for relative validity); 11–12 years (*n* 101 for second validity)	1 group: 1 × 24Hr and the other group: 7 × FR	F	15	7	Adolescents	Quantitative	Not shown	Spearman's correlations between the FFQ and the FR varied between 0·10 for crisps and 0·65 for semi-skimmed milk. Agreement varied between 34 % for the narrower definition of vegetables and 72 % for whole-fat milk. Gross misclassification varied between 1 % for chips and 21 % for diet soft drinks
54	Lietz *et al.*, UK, 2002[25]	12·3 (SD 0·3 years (*n* 67)	7 × WFR + 7 d urine biomarkers	N	Not shown	Not shown	Adolescents	Quantitative	Not shown	The median Spearman's CC for the nutrients examined was 0·31 and increased to 0·48 after adjustment for total energy
55	Yaroch *et al.*, USA, 2000[73]	11–17 years (*n* 22)	3 × 24 -Hr	N	110	8	Adolescents	Quantitative	2 weeks	The natural log-transformed energy-adjusted, deattenuated CC between the second FFQ and the mean from three recalls exceed 0·50 for most nutrients
56	Lambe *et al.*, Europe, 2000[98]	16 (SD 1) years (*n* 179)	1 × FR (3 d) + 1 × FR (14 d)	F	32	6	Adolescents	Qualitative	2 weeks	The mean between method differences was 0·02 (±0·06) portions/d with limits of agreement of –0·10 to 0·14
57	Field *et al.*, USA, 1999[74]	Children (*n* 109)	3 × 24Hr	F & N	97	Not shown	Adolescents	Semiquantitative	1 year	For most nutrients and food, the correlations between the FFQ and the 24Hr recalls remained greater among the junior high school students (fourth to fifth grade range: *r* 0·0–0·42; sixth to seventh grade range: *r* 0·07–0·76)
58	Robinson *et al.*, UK, 1999[75]	15 years (*n* 47)	7 × WFR + Food Checklist	F & N	83	Not shown	Adolescents	Quantitative	1 month	Energy and macronutrient intakes determined by FFQ1 were higher than those recorded in the WFR (all $P < 0·001$)
59	Taylor & Goulding, New Zealand, 1998[99]	3–6 years (*n* 67)	4 × FR	N	35	Not shown	Caregivers	Quantitative	1 year	The FFQ correctly identified 68 % of children with recorded intakes less than 800 mg
60	Samuelson *et al.*, Sweden, 1996[100]	15 years (*n* 218)	7 × WFR	F & N	29	8	Adolescents	Quantitative	1 month	There was relatively good correlation between the FFQ and the 7-WFR results

Table 4. (*Continued*)

	Author, country, year	No. of participants – age group	Reference method	Outcomes	No. of of food items	Category frequencies	Respondent of the FFQ	Quantitative or qualitative	Target period	Results
61	Bellú *et al.*, Italy, 1996[101]	9·3 years (mean) (*n* 323)	1 24Hr	N	116	Varied for each food item	Caregivers	Quantitative	Not shown	EI was found to be higher. For females, protein (–9 %) and cholesterol (–11 %) were found to be significantly lower with FFQ than 24Hr. For all other nutrients, the mean difference was <6 %
62	Crawford *et al.*, USA, 1994[102]	9–10 years (*n* 60)	1 × 24Hr + 1 × FR + observation	N	42	Varied for each food item	Adolescents	Quantitative	5 d	Comparisons of the intakes of energy and selected macronutrients showed different ranges and median percentage absolute errors for each dietary assessment method
63	Kaskoun *et al.*, USA, 1994[31]	4·2–6·9 years (*n* 45)	14 d DLW	N	111	9	Caregivers	Semiquantitative	1 year	TEI by FFQ (9·12 (SD 2·28) MJ/d) was significantly higher than TEE (5·74 (SD 1·13) MJ/d; $P < 0.001$)
64	Stein *et al.*, USA, 1994[76]	44–60 months (*n* 173)	7 × 24Hr	N	24	Not shown	Caregivers	Semiquantitative	6 months	Changes in nutrient density correlated poorly ($r < 0.15$) for all nine nutrients
65	Byers *et al.*, USA, 1993[24]	Caregivers of 6 to10 years (*n* 97)	1 blood biomarker	F & N	111	Not shown	Caregivers	Quantitative	3 months	The dietary reports of intakes of thirty-five fruits and vegetables showed Spearman's rank-order correlations of 0·30 with serum carotenoids and 0·34 with serum vitamin C
66	Frank *et al.*, USA, 1992[103]	12–17 years (*n* 1108)	7 × 24Hr	F	64	6	Adolescents	Quantitative	1 week	A mean 50 % agreement for both FFQ and 24Hr was observed
67	Jenner *et al.*, Australia, 1989[77]	11–12 years (*n* 225)	14 × 24Hr	N	175	6	Caregivers and children	Quantitative	1 week	Agreement between the reference method and the first two or three diet records in the series was relatively good

24Hr, 24-h record; CC, correlation coefficient; CCFFQ, Calcium Counts FFQ; CEHQ, Children's Eating Habits Questionnaire; CFFQ, children FFQ; CHO, carbohydrates; DP, dietary patterns; DLW, doubly labelled water; DQI, dietary quality index; EE, energy expenditure; F, food; FR, food record; F&V, fruits and vegetables; ICC, intraclass correlation coefficient; N, nutrient; RRS, resonance Raman spectroscopy; SFFQ, short FFQ; TEE, total energy expenditure; TEI, total energy intake; WDR, weighed diet record; WFR, weight food record; YANA-C, 24-h dietary recall instrument 'Young Adolescents' Nutrition Assessment on Computer.

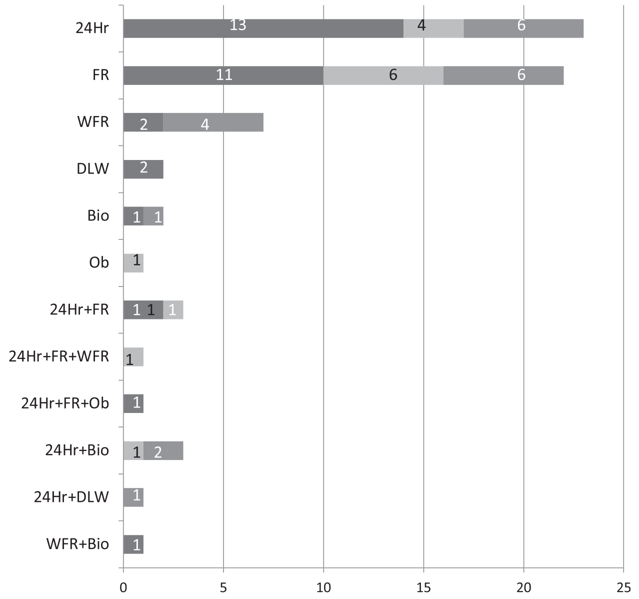

Fig. 2. Number of identified studies using the comparative dietary assessment method. ■, Nutrients; ■, food; ■, food and nutrients. 24Hr, 24-h record; FR, food record; WFR, weight food record; DLW, doubly labelled water; Bio, biomarkers; Ob, observation.

Conclusions of the systematic review provided by authors of sixty-seven studies

In relation to the results obtained by the authors of the selected studies, using Cohen's cut points[50,51], 2 % of the articles[26] concluded that the FFQ showed a very strong relative validity (>0·8) to assess food and nutrient intake in children and adolescents; 7 % of the studies[35,56,66,72,81] concluded that the FFQ showed a moderately strong relative validity (0·6–0·8) to assess dietary intake in children and adolescents; 31 %[24,25,36,38,41,53,54,57,61,62,64,65,67,69,73,78,86,89,93,94,104] showed a fair relative validity (0·3–0·5) to assess intake and 4 %[28,76,96] stated that the FFQ had poor relative validity for dietary assessment in this population group. Of the studies, 20 %[37,43,52,58–60,70,71,74,88,90,95,105] obtained different values depending on the food or nutrient assessed and 36 %[27,29,31–33,39,62,68,74,76,78,79,82–84,86,90,91,97–102] did not show results.

Summary of the results

Quality. Study quality was assessed using the standardised quality assessment checklist[47] proposed by the Academy of Nutrition and Dietetics in 2012. It was observed that all the studies, except two[43,101], were of a high quality (Table 5)[24–29,31,33–38,40,41,43,52–54,56–66,68–81,83–104]. However, these two studies were included in the systematic review, as some of the checklist items did not apply to them; one was also included in the corresponding meta-analysis[43].

The complete information extracted from the data analysis is available upon request from the corresponding author.

Meta-analysis. Meta-analysis was performed for energy, macronutrients and some micronutrients and food sources.

The meta-analysis performed included thirty-seven studies comparing the FFQ with another dietary assessment method (24Hr, FR and WFR) to estimate energy, CHO, protein, fat, fibre, Ca, Fe, Zn, vitamin A, vitamin C, meat, milk, fruit and vegetable intake[25,28,29,33,35,37,38,40–43,52–54,56–65,67–79].

Since we did not find statistical differences between younger children or caregivers and adolescents as respondents of the FFQ, the age of the category of the respondent was not taken in account for the meta-analysis.

Fig. 3 shows the results obtained for energy, Fig. 4(a–d) shows the results for CHO, protein, fat and fibre, Fig. 5(a–e) shows the results for Ca, Fe, Zn, vitamin A, and vitamin C, and Fig. 6(a–d) shows the results of the four food categories: meat, milk, fruits and vegetables.

For most nutrients and food categories, the correlations were very similar whether the reference method was 24Hr, FR or WFR (although slightly better in the case of WFR), according to Figs. 3, 4(a–d), 5(a) and (b), and 6(b–d). The exceptions were Zn (with correlations of 0·42 for the 24Hr, 0·22 for the FR and 0·52 for the WFR), vitamin A (with correlations of 0·50 for the 24Hr, 0·27 for the FR and 0·50 for the WFR), vitamin C (with correlations of 0·42 for the 24Hr, 0·32 for the FR and 0·51 for the WFR) and meat (with correlations of 0·41 for the 24Hr, 0·24 for the FR and 0·44 for the WFR). For energy, Fig. 3 shows that the correlations were similar when the reference method was the 24Hr of 0·48, the FR of 0·44 and the WFR of 0·47. For milk, Fig. 6(b) shows that the correlations were similar when the reference method was the 24Hr of 0·58, the FR of 0·56 and the WFR of 0·57. For Fe, Fig. 5(b) shows that the correlations were similar when the reference method was the 24Hr of 0·45, the FR of 0·42 and the WFR of 0·44. For energy, milk and Fe, the method that yielded the highest correlations to validate the FFQ was the 24Hr (0·58 for milk, 0·48 for energy and 0·45 for Fe). For Ca, both FR and WFR had the same correlation coefficient of 0·52, which was like the correlation of 0·45 for the 24Hr, according to Fig. 5(a). For vitamin A, both 24Hr and WFR had similar correlation coefficient, which was 0·50.

In all cases, the overall correlation coefficients between the FFQ and the reference methods were between 0·35 (meat) and 0·56 (milk). The level of heterogeneity (I^2) was significant in all cases and stand between 73 % for vitamin A and 99 % for meat. This finding is consistent with the relative validity of the FFQ assessed with other methods described in the studies included in the meta-analysis.

Meta-regression. The study showed a high heterogeneity value (I^2) (values of I^2 over or equal to 75 %) for different variables obtained in the meta-analysis, so a search for outliers, sensitivity studies and meta-regression analysis was performed. In most cases, meta-regression was not significant for any of the analysed variables (publication year, publication world region, reference method, number of food item of the FFQ, period that was assessed with the FFQ and respondent of the FFQ).

Positive effects were found for Ca and protein with the variable publication year (Akaike's information criterion = –23·669 for Ca and Akaike's information criterion = –19·878 for protein). Negative effects were found for energy with the variable food item (Akaike's information criterion = –23·103).

Table 5. Quality of studies that were included in the systematic review

Author	Year	Overall quality	Was the research question clearly stated?	Was the selection of study subjects/patients free from bias?	Were study groups comparable?	Was method of handling withdrawals described?	Was blinding used to prevent introduction of bias?	Were intervention/therapeutic regimens/exposure factor or procedure and any comparison described in detail? Were intervening factors described?	Were outcomes clearly defined and the measurements valid and reliable?	Was the statistical analysis appropriate for the study design and type of outcome indicators?	Are conclusions supported by results with biases and limitations taken into consideration?	Is bias due to study's funding or sponsorship unlikely?
Kunaratman	2018	+	Yes	Unclear	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Leong	2018	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Rodriguez	2017	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Söderberg	2017	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Nyström	2017	+	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Vioque	2016	+	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
Yum	2016	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Rahmawaty	2016	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
De Cock	2016	+	Yes	Unclear	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Moghames	2016	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Saeedi	2016	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Fatitah	2015	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Julian	2015	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Marcinkevage	2015	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Saloheimo	2015	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Aguilar	2014	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Appannah	2014	+	Yes	Yes	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bel-Serrat	2013	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Flood	2014	+	Yes	Yes	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Martinez	2013	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Pampaloni	2013	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Unclear
Hunsberger	2012	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	No
Lillegaard	2012	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Mulasi-Pokhriyal	2012	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Del Pino	2011	+	Yes	Unclear	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Huybrechts	2011	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Kobayashi	2011	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
Nurul-Fadhilah	2011	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Sahashi	2012	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Scagliusi	2011	+	Yes	Unclear	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Xia	2011	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
Araujo	2010	+	Yes	Unclear	N/A	N/A	Unclear	Yes	Yes	Yes	Yes	Yes
Dutman	2010	+	Yes	Yes	N/A	N/A	Unclear	Yes	Yes	Yes	Yes	Yes
Huybrects	2010	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Slater	2010	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Vereecken	2010	+	Yes	Unclear	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Vereecken	2010	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Watanabe	2010	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Di Noia	2009	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Hong	2010	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	No

Table 5. (*Continued*)

Author	Year	Overall quality	Was the research question clearly stated?	Was the selection of study subjects/patients free from bias?	Were study groups comparable?	Was method of handling withdrawals described?	Was blinding used to prevent introduction of bias?	Were intervention/therapeutic regimens/exposure factor or procedure and any comparison described in detail? Were intervening factors described?	Were outcomes clearly defined and the measurements valid and reliable?	Was the statistical analysis appropriate for the study design and type of outcome indicators?	Are conclusions supported by results with biases and limitations taken into consideration?	Is bias due to study's funding or sponsorship unlikely?
Huybrechts	2009	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Stiegler	2009	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Watson	2009	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Zemel	2009	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Papadopoulou	2008	–	Yes	Yes	N/A	No	Unclear	Yes	Yes	Unclear	Yes	Unclear
Vereecken	2008	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Harnack	2006	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Huybrechts	2006	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Perez-Cueto	2006	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Haraldsdóttir	2005	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Andersen	2004	+	Yes	Yes	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Jensen	2004	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Vereecken	2003	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Lietz	2002	+	Yes	Yes	N/A	No	Yes	Yes	Yes	Yes	Yes	Yes
Yaroch	2000	+	Yes	No	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
Lambe	2000	+	Yes	Yes	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Unclear
Samuelson	2000	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes
Field	1999	+	Yes	Yes	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Unclear
Robison	1999	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Unclear
Taylor	1998	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Unclear
Bellú	1996	–	No	Yes	N/A	N/A	Unclear	No	Yes	Unclear	Unclear	Yes
Crawford	1994	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Kaskoun	1994	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Unclear
Stein	1994	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
Byers	1993	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Unclear
Frank	1992	+	Yes	Yes	N/A	No	Unclear	Yes	Yes	Yes	Yes	Yes
Jenner	1989	+	Yes	Yes	N/A	Yes	Unclear	Yes	Yes	Yes	Yes	Yes

+, Positive quality; N/A, not applicable; –, negative quality.

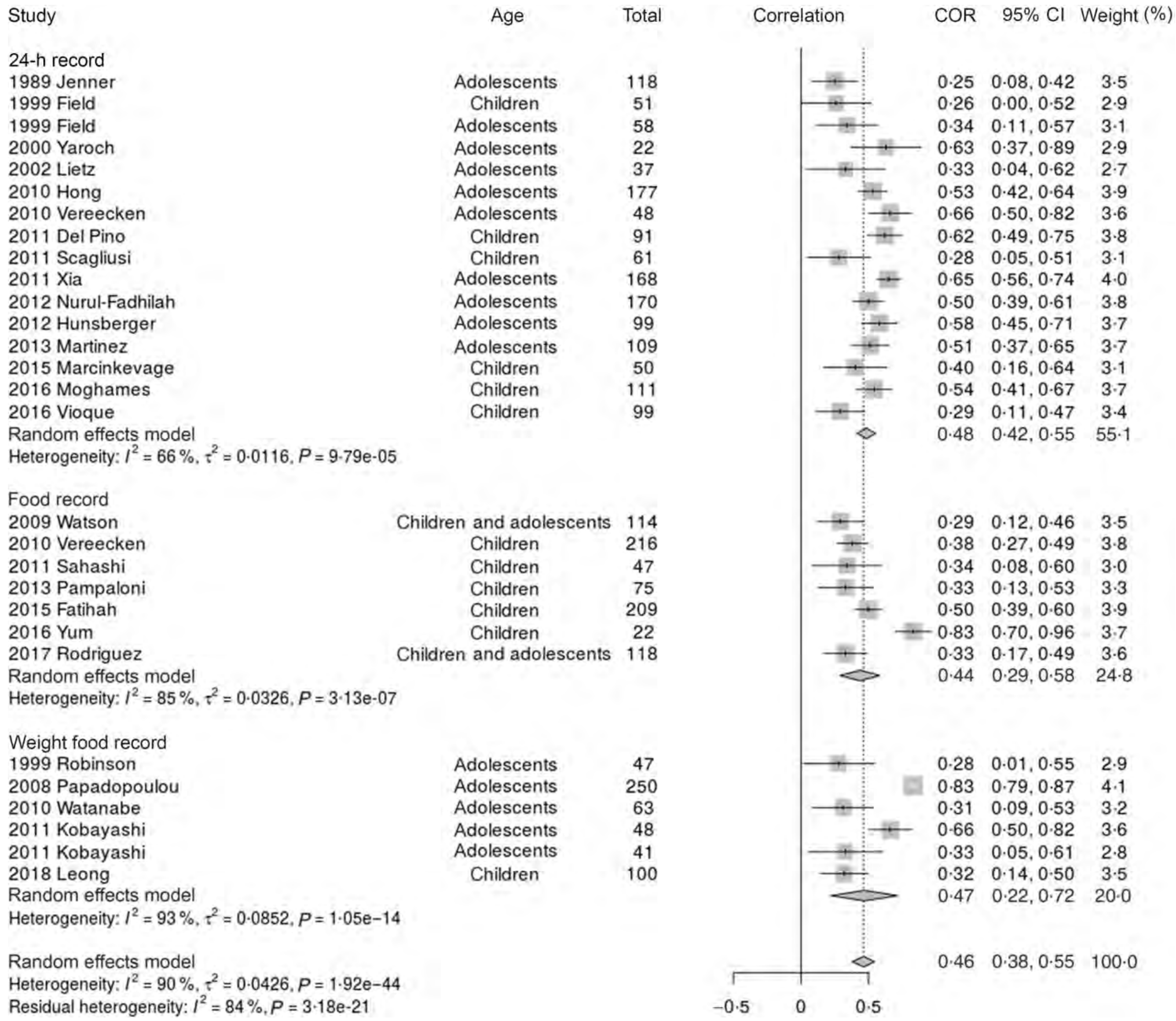

Fig. 3. Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate energy. COR, correlation.

Discussion

To the best of our knowledge, this is the first systematic review and meta-analysis assessing the relative validity of FFQ to estimate energy, macronutrients, certain micronutrients and some food item intake in children and adolescents. In the meta-analysis reported in this article, the overall relative validity of energy, macronutrients (CHO, protein, fat and fibre), certain micronutrients (Ca, Fe, Zn, vitamin A and vitamin C) and some food categories (meat, milk, vegetables and fruits) intake estimation using the FFQ may be considered as weak (correlation coefficients between 0·35 and 0·56). However, three reference standard methods were used across studies. For most nutrients and food categories, the correlations were similar whether the reference method was 24Hr, FR or WFR.

It has been frequently said that there is no perfect measurement of dietary intake, with the implication that validation studies are not possible. The lack of a perfect standard is, however, not unique to dietary intake assessment method; all measurements entail uncertainty, although they differ in their magnitude. Thus, relative validation studies never compare an operational method with the absolute truth, but instead they compare one method with another method[1].

Relative validity of EI assessed with the FFQ when using the 24Hr, FR and WFR as the reference method can be considered as weak (correlation coefficients = 0·46). However, correlation coefficients cannot identify whether the FFQ is under- or over-estimating EI. Most of the studies assessing the relative validity of the FFQ to estimate EI report the difference in mean daily EI between methods but do not report the agreement or bias between them.

There are two studies assessing the validity of the FFQ against the DLW method[31,81], which is considered the reference standard for studies validating EI; both studies produced different

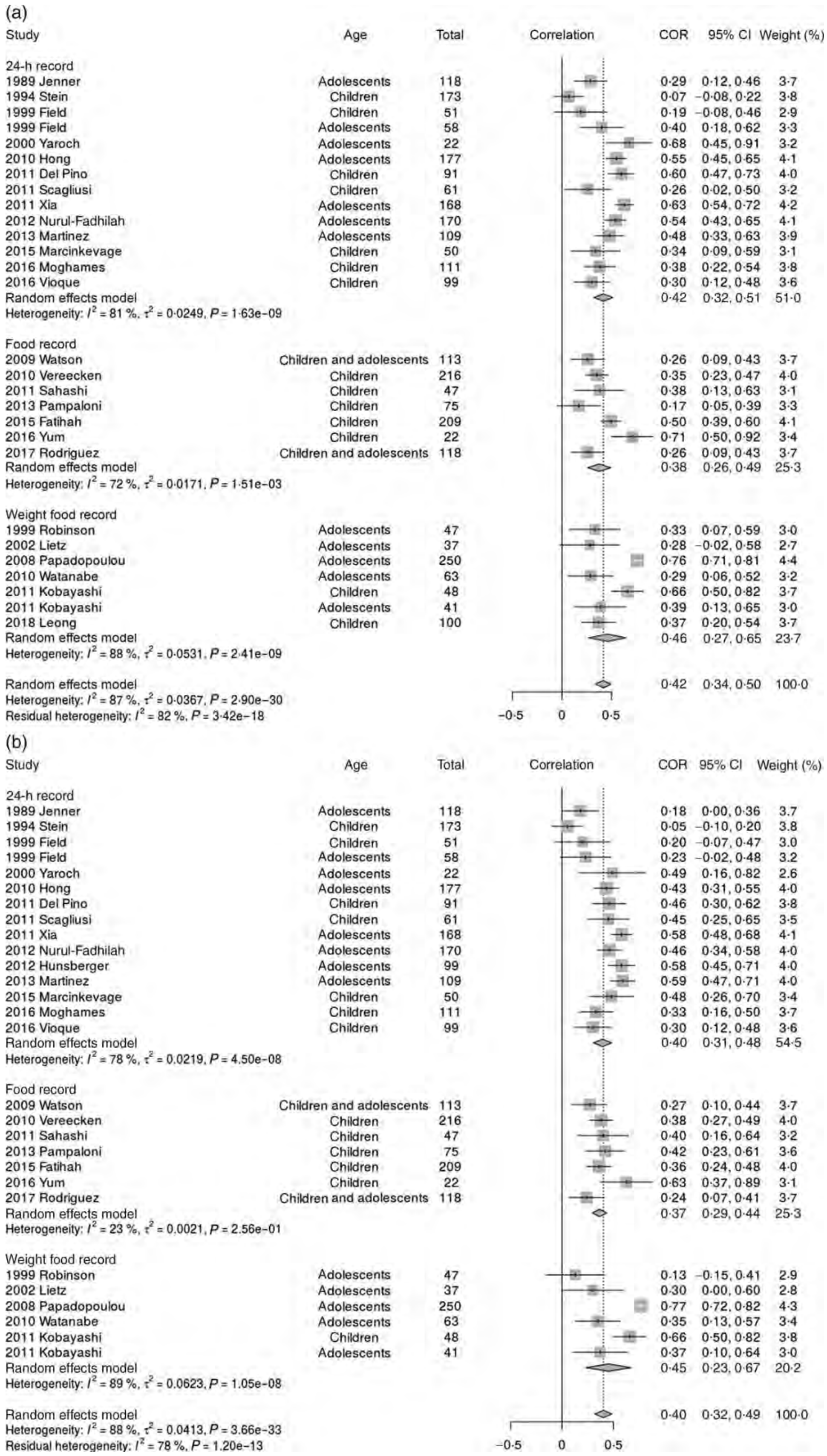

Fig. 4. (a) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate carbohydrate. (b) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate protein. (c) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate fat. (d) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate fibre. COR, correlation.

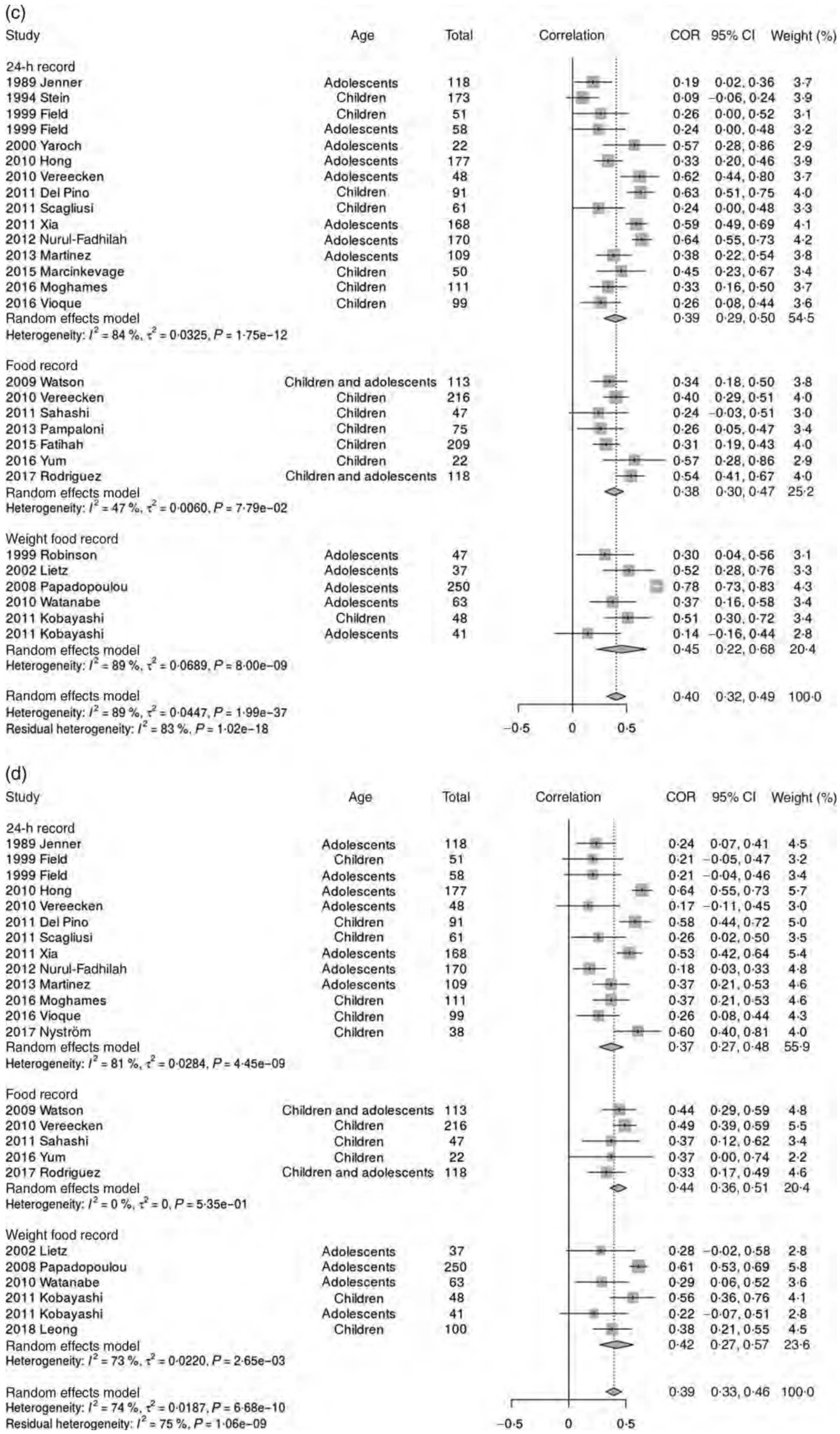

Fig. 4. (continued)

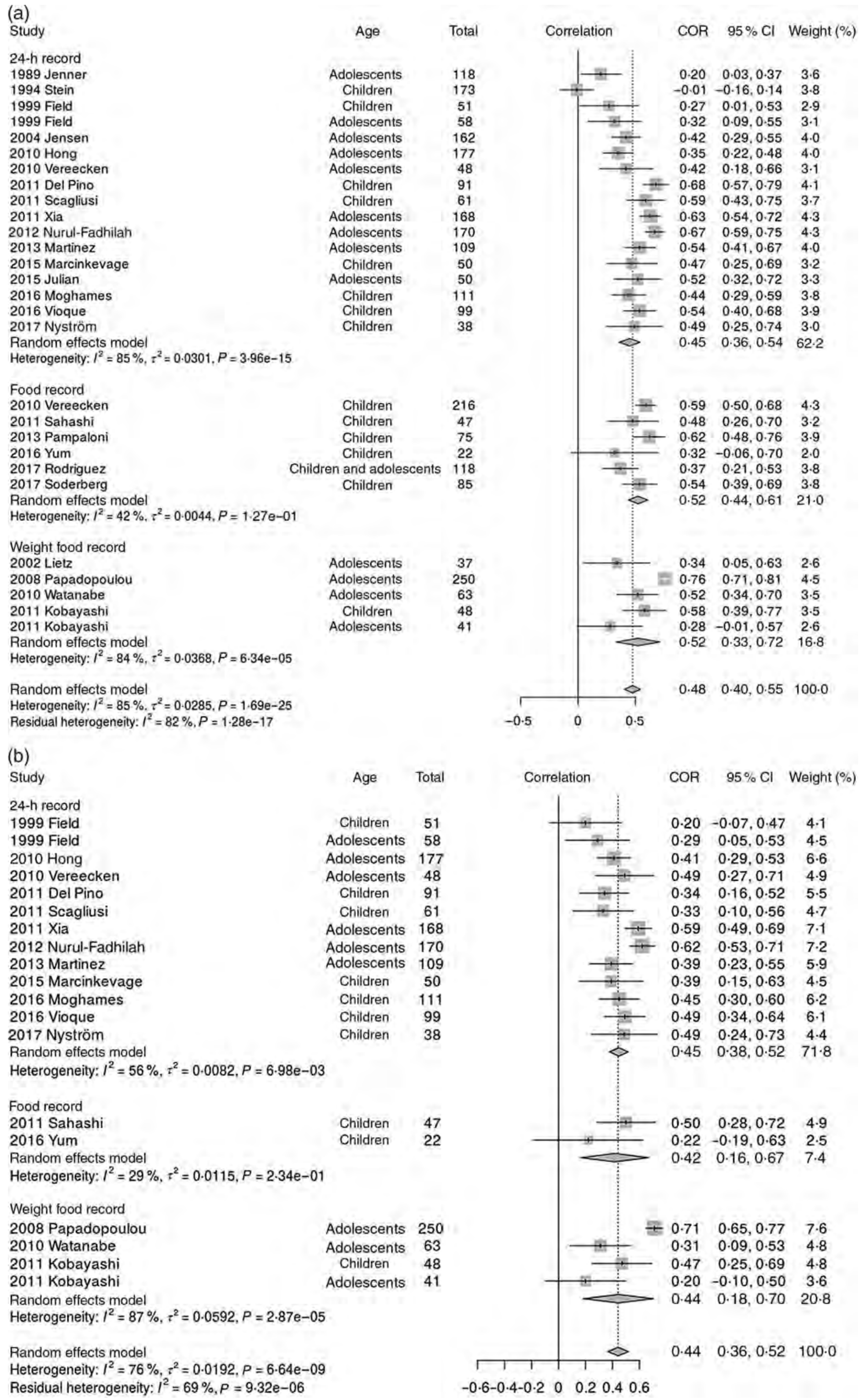

Fig. 5. (a) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate calcium. (b) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate iron. (c) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate zinc. (d) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate vitamin A. (e) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate vitamin C.

British Journal of Nutrition

British Journal of Nutrition

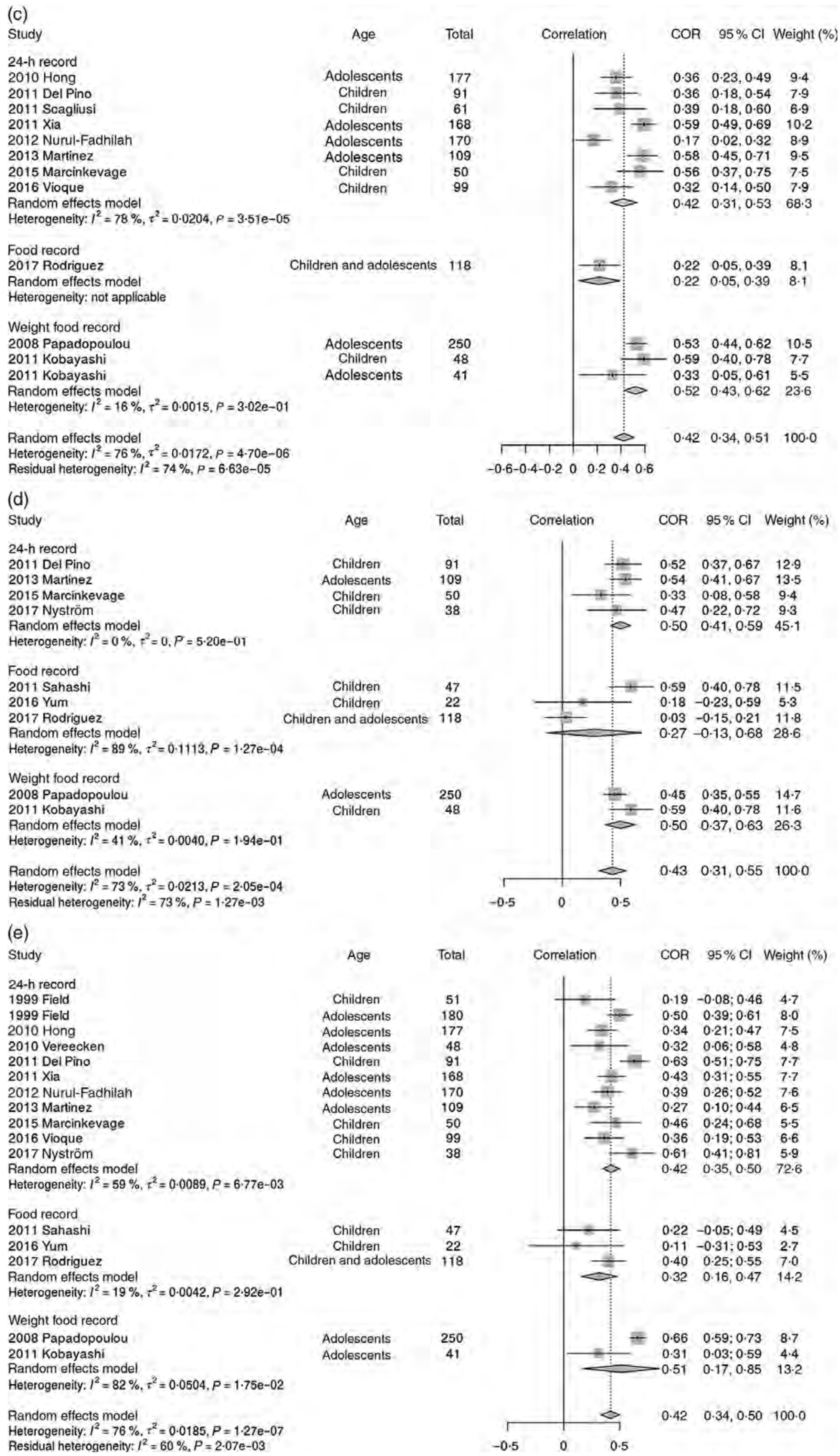

Fig. 5. (continued)

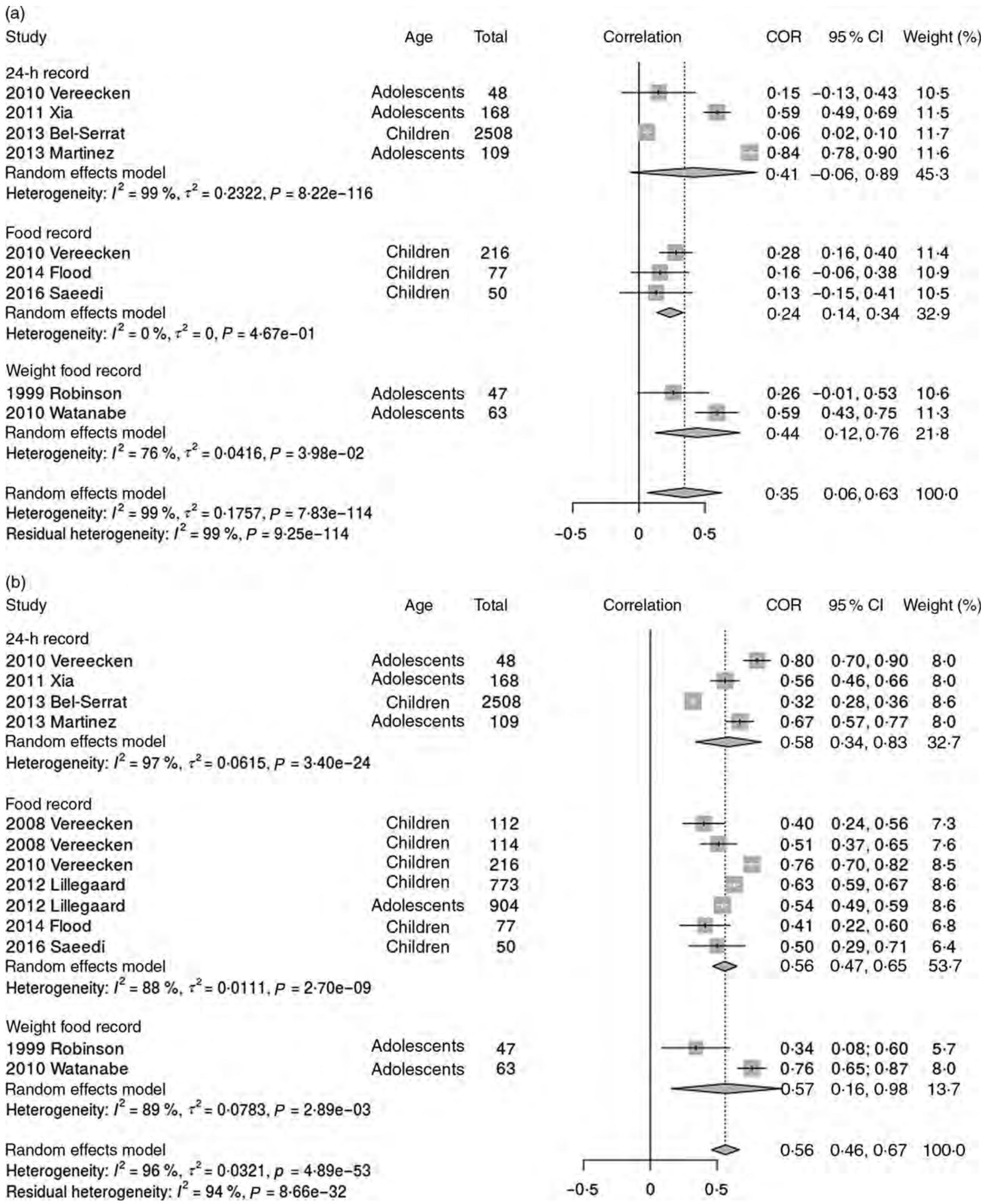

Fig. 6. (a) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate meat. (b) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate milk. (c) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate fruits. (d) Comparison of the FFQ with the 24-h record, the food record and the weight food record to estimate vegetables.

British Journal of Nutrition

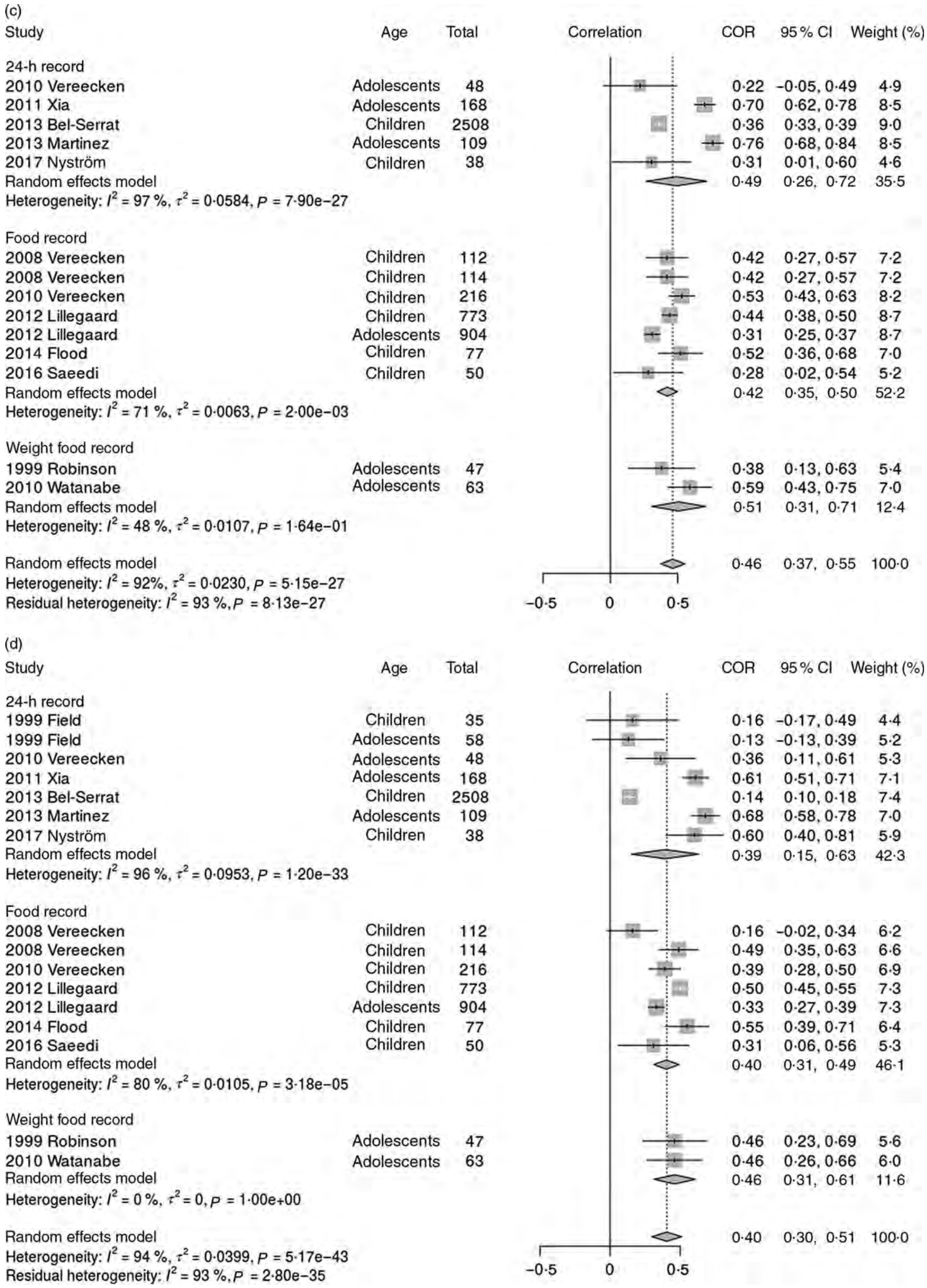

Fig. 6. (continued)

British Journal of Nutrition

findings. Kaskoun *et al.* report that total EI assessed by a FFQ was significantly higher than total energy expenditure assessed with DLW ($P > 0{\cdot}001$)[31], and Dutman *et al.* indicate that mean total EI, assessed with FFQ, did not differ significantly from total energy expenditure assessed with DLW ($P > 0{\cdot}15$)[81].

The result of this meta-analysis showed a weak overall relative validity of energy, macronutrients, certain micronutrients and some food category intake estimation, using the FFQ (correlation coefficients between 0·35 and 0·56).

Strengths

In the search for evidence to answer the research questions, it is preferable to seek a systematic review, especially one that includes a meta-analysis. Single studies are responsible for being unrepresentative of the total evidence and might not be true. Systematic reviews include a wider range of subjects than any single study, potentially increasing confidence in the implementation of the outcome for the case in question. The meta-analysis of a set of tests includes a larger sample than a single study, leading to greater accuracy of estimates, which facilitates confident decision-making. This is especially pertinent to relative validation studies, as the sample sizes are often small. For this systematic review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol[45] and the methodology suggested by the Academy of Nutrition and Dietetics, for the evidence analysis process, 2016[46]. Sixty-seven studies were found to assess relative validity of FFQ with other dietary assessment methods in children and adolescents and thirty-seven studies were included in the meta-analysis. Energy, but also macronutrients (CHO, protein, fat, and fibre), micronutrients (Ca, Fe, Zn, vitamin A and vitamin C) and some food items (meat, milk, fruits and vegetables) were analysed. Because of the high heterogeneity value (I^2) obtained in the meta-analysis, a meta-regression analysis was performed to assess the relationship between the FFQ and the other reference methods for energy, macronutrient (CHO, protein, fat and fibre), micronutrients (Ca, Fe, 11Zn, vitamin A and vitamin C) and some food categories (meat, milk, vegetables and fruits).

Despite of the weak performance of the FFQ, they are still recommended for epidemiological studies because of their low cost, they may be used in a self-administered format, they show the usual dietary intake over long periods of time, they can be used for many participants and they can compare dietary intake between different populations[1,12,32,108–110].

Limitations

When performing the systematic review, a lack of detailed information in the original articles included was identified. This jeopardised the inclusion of certain studies in the meta-analysis, as thirty out of sixty-seven were not statistically comparable.

In this systematic review in children and adolescents, only two studies validating EI with DLW were found and they obtained discrepant results. For biomarkers, seven studies validated the FFQ, but none of the studies could be compared with each other because they validated different micronutrients with differently measured biomarkers. Therefore, there is an urgent need to perform validation studies of EI using DLW as the reference standard and biomarkers (25-hydroxy vitamin D in plasma for vitamin D, blood samples for vitamin C, retinol and carotenoids, urine samples for Ca and K, among others) for macronutrients, micronutrients and foods. There is also a need to develop, validate and use modern tools (such as smartphones, mobile devices, applications or interactive software) to assess dietary intake.

Most of the studies only provided correlation coefficients as estimates of the relative validity of results obtained with the FFQ. However, correlation coefficients do not provide information on the potential misreporting of the FFQ. Other methods, such as the Bland–Altman plots, which assess the agreement between quantitative measurements across the range of intakes[109], providing information on the agreement/bias of the results, should be used. This is also important in terms of identifying the relative validity of the method at individual or group level[111].

In meta-analysis, heterogeneity in results is expected because data from studies that are diverse usually encounter this limitation. It was suggested that there is not much sense in simply assessing heterogeneity, when what matters is the degree to which it affects the findings of the meta-analysis[112]. As the heterogeneity in the performed meta-analysis was high, a meta-regression was performed.

Conclusion

The relative validity of the dietary assessment methods is a topic of current interest. FFQ are the preferred dietary assessment method in most epidemiological studies in children and adolescents mainly due to their low cost, ease of administration and the fact that they allow for people to be classified, considering a long period of food intake. However, all self-reporting methods of food intake are subject to errors and, therefore, validation studies, with the appropriate method, are required to assess the effect of measurement error and to avoid incorrect estimations wherever possible. From this systematic review, it can be concluded that, in children and adolescents, the FFQ has fair relative validity to assess dietary intake. The meta-analysis performed showed that overall relative validity of energy, macronutrients (CHO, protein, fat and fibre), some micronutrients (Ca, Fe, Zn, vitamin A and vitamin C) and some food item (milk, fruits and vegetables) intake estimation, using the FFQ, may be considered as weak.

Acknowledgements

L. S. was supported by a doctoral scholarship from the Carolina Foundation. I. I. is supported by RETICS founded by the PN I+D+I 2013-2016, ISCIII – Subdirección General de Evaluación y Fomento de la Investigación y el Fondo Europeo de Desarrollo Regional (FEDER), ref. RD16/0022. M. V. N.-F. received a PhD Student Internships Abroad scholarship from the National Council of Technological and Scientific Development (CNPq; proc. 200340/2015-8) and a Brazilian PhD Student scholarship from the São Paulo Research Foundation – FAPESP (proc. 2016/18436-8 and 2017/11732-3).

L. S. and L. A. M. were involved in the design and conducting of the study, the writing of the first draft paper, and critically

reviewed the manuscript. L. S., M. L. M.-B. and I. I. were involved in the systematic review of the literature. G. P. and M. V. N.-F. were involved in the meta-analysis of the data. G. P. was involved in the meta-regression of the data. I. B. and B. S. critically reviewed the manuscript. All authors contributed to the writing and finally approved the submitted and published versions.

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Willett WC (2013) *Nutritional Epidemiology*. New York: Oxford University Press.
2. Rankin D & Hanekom S (2010) Dietary assessment methodology for adolescents: a review of reproducibility and validation studies. *South Afr J Clin Nutr* **23**, 65–74.
3. Livingstone MBE & Robson PJ (2000) Measurement of dietary intake in children. *Proc Nutr Soc* **59**, 279–293.
4. Baxter SD, Thompson WO, Davis HC, *et al.* (1997) Impact of gender, ethnicity, meal component, and time interval between eating and reporting on accuracy of fourth-graders' self-reports of school lunch. *J Am Diet Assoc* **97**, 1293–1298.
5. Baxter SD, Hitchcock DB, Royer JA, *et al.* (2017) Fourth-grade children's dietary reporting accuracy by meal component: results from a validation study that manipulated retention interval and prompts. *Appetite* **113**, 106–115.
6. Baxter SD, Hardin JW, Guinn CH, *et al.* (2009) Fourth-grade children's dietary recall accuracy is influenced by retention interval (target period and interview time). *J Am Diet Assoc* **109**, 846–856.
7. Baxter SD, Smith AF, Litaker MS, *et al.* (2004) Recency affects reporting accuracy of children's dietary recalls. *Ann Epidemiol* **14**, 385–390.
8. Sharman SJ, Skouteris H, Powell MB, *et al.* (2016) Factors related to the accuracy of self-reported dietary intake of children aged 6 to 12 years elicited with interviews: a systematic review. *J Acad Nutr Diet* **116**, 76–114.
9. Smith AF (1991) *Cognitive Processes in Long-Term Dietary Recall*. Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics.
10. Smith AF, Jobe JB & Mingay DJ (1991) Retrieval from memory of dietary information. *Appl Cogn Psychol* **5**, 269–296.
11. Livingstone MBE, Robson PJ & Wallace JMW (2004) Issues in dietary intake assessment of children and adolescents. *Br J Nutr* **92**, S213.
12. Saravia L, González-Zapata LILI, Rendo-Urteaga T, *et al.* (2018) Development of a food frequency questionnaire for assessing dietary intake in children and adolescents in South America. *Obesity* **26**, S31–S40.
13. Cade JE, Burley VJ, Warm DL, *et al.* (2004) Food-frequency questionnaires: a review of their design, validation and utilisation. *Nutr Res Rev* **17**, 5–22.
14. Riboli E, Hunt K, Slimani N, *et al.* (2003) European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* **5**, 1113–1124.
15. CDC (2014) *National Health and Nutrition Examination Survey, 2013–2014. Nat*ional Health Statistics Report *2013–2014.*
16. Gupta S (2014) Healthy heart. **5**, 1–8.
17. Colditz GA & Hankinson SE (2005) The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* **5**, 388–396.
18. Millen AE, Midthune D, Thompson FE, *et al.* (2006) The National Cancer Institute Diet History Questionnaire: validation of pyramid food servings. *Am J Epidemiol* **163**, 279–288.
19. Pinket A-SS, De Craemer M, Huybrechts I, *et al.* (2016) Diet quality in European pre-schoolers: evaluation based on diet quality indices and association with gender, socio-economic status and overweight, the ToyBox-study. *Public Health Nutr* **19**, 2441–2450.
20. Moreno LA, Gottrand F, Huybrechts I, *et al.* (2014) Nutrition, lifestyle in European Adolescents: the HELENA (HealthyLiftestyle in Europe by Nutrition in Adolescence) Study. *Adv Nutr* **5**, 615A–623AS.
21. Naska A, Lagiou A & Lagiou P (2017) Dietary assessment methods in epidemiological research: current state of the art and future prospects. *F1000Res* **6**, 1–8.
22. Wrieden W, Peace H, Armstrong J, *et al.* (2003) A short review of dietary assessment methods used in National, Scottish Research Studies. *Working Group on Monitoring Scottish Dietary Targets Work.*
23. Ball SC, Benjamin SE & Ward DS (2007) Development and reliability of an observation method to assess food intake of young children in child care. *J Am Diet Assoc* **107**, 656–661.
24. Byers T, Trieber F, Gunter E, *et al.* (1993) The accuracy of parental reports of their children's intake of fruits and vegetables: validation of a food frequency questionnaire with serum levels of carotenoids and vitamins C, A, and E. *Epidemiology* **4**, 350–355.
25. Lietz G, Barton KL, Longbottom PJ, *et al.* (2002) Can the EPIC food-frequency questionnaire be used in adolescent populations? *Public Health Nutr* **5**, 783–789.
26. Slater B, Enes CC, López RVM, *et al.* (2010) Validation of a food frequency questionnaire to assess the consumption of carotenoids, fruits and vegetables among adolescents: the method of triads. *Cad Saude Publica* **26**, 2090–2100.
27. Aguilar SS, Wengreen HJ, Lefevre M, *et al.* (2014) Skin carotenoids: a biomarker of fruit and vegetable intake in children. *J Acad Nutr Diet* **114**, 1174–1180.
28. Vioque J, Gimenez-Monzo D, Navarrete-Muñoz EM, *et al.* (2016) Reproducibility and validity of a food frequency questionnaire designed to assess diet in children aged 4–5 years. *PLOS ONE* **11**, 1–17.
29. Söderberg L, Lind T, Karlsland Åkeson P, *et al.* (2017) A validation study of an interviewer-administered short food frequency questionnaire in assessing. *Nutrients* **9**, 682.
30. Kuhnle GGC (2012) Nutritional biomarkers for objective dietary assessment. *J Sci Food Agric* **92**, 1145–1149.
31. Kaskoun MC, Johnson RK & Goran MI (1994) Comparison of energy intake by semiquantitative food-frequency questionnaire with total energy expenditure by the doubly labeled water method in young children. *Am J Clin Nutr* **60**, 43–47.
32. Tabacchi G, Amodio E, Di Pasquale M, *et al.* (2014) Validation and reproducibility of dietary assessment methods in adolescents: a systematic literature review. *Public Health Nutr* **17**, 2700–2714.
33. Marcinkevage J, Mayén AL, Zuleta C, *et al.* (2015) Relative validity of three food frequency questionnaires for assessing dietary intakes of Guatemalan schoolchildren. *PLOS ONE* **10**, 1–18.
34. Pérez-Cueto Eulert FJA, Roberfroid D & Kolsteren PW (2006) Desarrollo y evaluación de un cuestionario semi-cuantitativo de frecuencias alimenticias para adolescentes bolivianos (Development and testing of a semi-quantitative food

frequency questionnaire for Bolivan adolescents). *Nutr Hosp* **21**, 573–580.
35. Del Pino DL & Friedman R (2011) Adaptation and validation of an FFQ for 6–10-year-old children. *Public Health Nutr* **14**, 826–834.
36. Araujo MC, Yokoo EM & Pereira RA (2010) Validation and calibration of a semiquantitative food frequency questionnaire designed for adolescents. *J Am Diet Assoc* **110**, 1170–1177.
37. Rodriguez CA, Smith ER, Villamor E, *et al.* (2017) Development and validation of a food frequency questionnaire to estimate intake among children and adolescents in Urban Peru. *Nutrients* **9**, 1–10.
38. Fatihah F, Ng BK, Hazwanie H, *et al.* (2015) Development and validation of a food frequency questionnaire for dietary intake assessment among multi-ethnic primary school-aged children. *Singapore Med J* **56**, 687–694.
39. Kolodziejczyk JK, Merchant G & Norman GJ (2012) Reliability and validity of child/adolescent food frequency questionnaires that assess foods and/or food groups. *J Pediatr Gastroenterol Nutr* **55**, 4–13.
40. Moghames P, Hammami N, Hwalla N, *et al.* (2016) Validity and reliability of a food frequency questionnaire to estimate dietary intake among Lebanese children. *Nutr J* **15**, 1–12.
41. Hong TK, Dibley MJ & Sibbritt D (2010) Validity and reliability of an FFQ for use with adolescents in Ho Chi Minh City, Vietnam. *Public Health Nutr* **13**, 368–75.
42. Kobayashi T, Tanaka S, Toji C, *et al.* (2010) Development of a food frequency questionnaire to estimate habitual dietary intake in Japanese children. *Nutr J* **9**, 17.
43. Papadopoulou SK, Barboukis V, Dalkiranis A, *et al.* (2008) Validation of a questionnaire assessing food frequency and nutritional intake in Greek adolescents. *Int J Food Sci Nutr* **59**, 148–154.
44. Kelley GA & Kelley KS (2019) Systematic reviews and meta-analysis in nutrition research. *Br J Nutr* **122**, 1279–1294.
45. Liberati A, Altman DG, Tetzlaff J, *et al.* (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* **62**, e1–34.
46. Handu D, Moloney L, Wolfram T, *et al.* (2016) Academy of nutrition and dietetics methodology for conducting systematic reviews for the evidence analysis library. *J Acad Nutr Diet* **116**, 311–318.
47. Academy of Nutrition and Dietetics (2012) *Evidence Analysis Manual: Steps in the Academy Evidence Analysis Process*. Chicago, IL: Academy of Nutrition and Dietetics.
48. Murad MH, Montori VM, Ioannidis JPA, *et al.* (2014) How to read a systematic review and meta-analysis and apply the results to patient care. *JAMA* **312**, 171.
49. Schwarzer G, Mair P & Hatzinger R (2016) meta: An R package for meta-analysis.
50. Chan YH (2003) 104 - Correlation analysis. *Singapore Med J* **44**, 614–619.
51. Cohen J (1992) Statistical power analysis. *Curr Dir Psychol Sci* **1**, 98–101.
52. Yum J & Lee S (2016) Development and evaluation of a dish-based semiquantitative food frequency questionnaire for Korean adolescents. *Nutr Res Pract* **10**, 433–441.
53. 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**, 1–13.
54. Julián-Almárcegui C, Huybrechts I, Gómez Bruton A, *et al.* (2015) Validity of a food-frequency questionnaire for estimating calcium intake in adolescent swimmers. *Nutr Hosp* **32**, 1773–1779.
55. Ortega FB, Artero EG, Ruiz JR, *et al.* (2008) Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. *Int J Obes* **32**, S49–S57.
56. Martinez MF & Estima C (2013) Validity and reproducibility of a food frequency questionnaire to assess food group intake in adolescents. *Cad Saude Publica* **29**, 1795–1804.
57. Bel-Serrat S, Mouratidou T, Pala V, *et al.* (2014) Relative validity of the Children's Eating Habits Questionnaire – food frequency section among young European children: the IDEFICS Study. *Public Health Nutr* **17**, 266–276.
58. Pampaloni B, Bartolini E, Barbieri M, *et al.* (2013) Validation of a food-frequency questionnaire for the assessment of calcium intake in schoolchildren aged 9–10 years. *Calcif Tissue Int* **93**, 23–38.
59. Flood VM, Wen LM, Hardy LL, *et al.* (2014) Reliability and validity of a short FFQ for assessing the dietary habits of 2–5-year-old children, Sydney, Australia. *Public Health Nutr* **17**, 498–509.
60. Hunsberger M, O'Malley J, Block T, *et al.* (2012) Relative validation of Block Kids Food Screener for dietary assessment in children and adolescents. *Matern Child Nutr* **11**, 260–270.
61. Lillegaard ITL, Overby NC & Andersen LF (2012) Evaluation of a short food frequency questionnaire used among Norwegian children. *Food Nutr Res* **56**, doi:10.3402/fnr.v56i0.6399.
62. Scagliusi FB, Garcia MT, Indiani ALC, *et al.* (2011) Relative validity of a food-frequency questionnaire developed to assess food intake of schoolchildren living in the Brazilian Western Amazon. *Cad Saude Publica* **27**, 2197–2206.
63. Nurul-Fadhilah A, Teo PS & Foo LH (2012) Validity and reproducibility of a food frequency questionnaire (FFQ) for dietary assessment in Malay adolescents in Malaysia. *Asia Pac J Clin Nutr* **21**, 97–103.
64. Xia W, Sun C, Zhang L, *et al.* (2011) Reproducibility and relative validity of a food frequency questionnaire developed for female adolescents in Suihua, North China. *PLoS ONE* **6**, 1–7.
65. Sahashi Y, Tsuji M, Wada K, *et al.* (2011) Validity and reproducibility of food frequency questionnaire in Japanese children aged 6 years. *J Nutr Sci Vitaminol* **57**, 372–376.
66. Kobayashi T, Kamimura M, Imai S, *et al.* (2011) Reproducibility and validity of the food frequency questionnaire for estimating habitual dietary intake in children and adolescents. *Nutr J* **10**, 27.
67. Vereecken CA, De Bourdeaudhuij I & Maes L (2010) The HELENA online food frequency questionnaire: reproducibility and comparison with four 24-h recalls in Belgian–Flemish adolescents. *Eur J Clin Nutr* **64**, 541–548.
68. Vereecken C, Covents M & Maes L (2010) Comparison of a food frequency questionnaire with an online dietary assessment tool for assessing preschool children's dietary intake. *J Hum Nutr Diet* **23**, 502–510.
69. Watanabe M, Yamaoka K, Yokotsuka M, *et al.* (2010) Validity and reproducibility of the FFQ (FFQW82) for dietary assessment in female adolescents. *Public Health Nutr* **14**, 297–305.
70. Watson JF, Collins CE, Sibbritt DW, *et al.* (2009) Reproducibility and comparative validity of a food frequency questionnaire for Australian children and adolescents. *Int J Behav Nutr Phys Act* **6**, 62.
71. 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.
72. Jensen JK, Gustafson D, Boushey CJ, *et al.* (2004) Development of a food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and white youth. *J Am Diet Assoc* **104**, 762–769.

British Journal of Nutrition

73. Yaroch AL, Resnicow K, Davis M, *et al.* (2000) Development of a modified picture-sort food frequency questionnaire administered to low-income, overweight, African-American adolescent girls. *J Am Diet Assoc* **100**, 1050–1056.
74. Field AE, Peterson KE, Gortmaker SL, *et al.* (1999) Reproducibility and validity of a food frequency questionnaire among fourth to seventh grade inner-city school children: implications of age and day-to-day variation in dietary intake. *Public Health Nutr* **2**, 293–300.
75. Robinson S, Skelton R, Barker M, *et al.* (1999) Assessing the diet of adolescent girls in the UK. *Public Health Nutr* **2**, 571–577.
76. Stein AD, Shea S, Basch CE, *et al.* (1994) Assessing changes in nutrient intakes of preschool children : comparison of 24-hour dietary recall and food frequency methods. *Epidemiology* **5**, 109–115.
77. Jenner DA, Neylon K, Croft S, *et al.* (1989) A comparison of methods of dietary assessment in australian children aged 11–12 years. *Eur J Clin Nutr* **43**, 663–673.
78. Leong C, Taylor RW, Haszard JJ, *et al.* (2018) Relative validity and reproducibility of a food frequency questionnaire to assess nutrients and food groups of relevance to the gut microbiota in young children. *Nutrients* **10**, 1627.
79. Nyström CD, Henriksson H, Alexandrou C, *et al.* (2017) Validation of an online food frequency questionnaire against doubly labelled water and 24 h dietary recalls in pre-school children. *Nutrients* **9**, 66.
80. Huybrechts I, Börnhorst C, Pala V, *et al.* (2011) Evaluation of the Children's Eating Habits Questionnaire used in the IDEFICS study by relating urinary calcium and potassium to milk consumption frequencies among European children. *Int J Obes* **35**, S69–S78.
81. Dutman AE, Stafleu A, Kruizinga A, *et al.* (2010) Validation of an FFQ and options for data processing using the doubly labelled water method in children. *Public Health Nutr* **14**, 410–417.
82. Borenstein M, Hedges LV, Higgins JPT, *et al.* (2010) A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* **1**, 97–111.
83. Rahmawaty S, Charlton K, Lyons-Wall P, *et al.* (2017) Development and validation of a food frequency questionnaire to assess omega-3 long chain polyunsaturated fatty acid intake in Australian children aged 9–13 years. *J Hum Nutr Diet* **30**, 429–438.
84. De Cock N, Van Camp J, Kolsteren P, *et al.* (2016) Development and validation of a quantitative snack and beverage food frequency questionnaire for adolescents. *J Hum Nutr Diet* **30**, 141–150.
85. 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**, S22–S28.
86. Appannah G, Pot GK, O'Sullivan TA, *et al.* (2014) The reliability of an adolescent dietary pattern identified using reduced-rank regression: comparison of a FFQ and 3 d food record. *Br J Nutr* **112**, 609–615.
87. Mulasi-Pokhriyal U & Smith C (2013) Comparison of the Block Kid's Food Frequency Questionnaire with a 24 h dietary recall methodology among Hmong-American children, 9–18 years of age. *Br J Nutr* **109**, 346–52.
88. Huybrechts I, Vereecken C, De Bacquer D, *et al.* (2010) Reproducibility and validity of a diet quality index for children assessed using a FFQ. *Br J Nutr* **104**, 135–144.
89. Di Noia J & Contento IR (2009) Use of a brief food frequency questionnaire for estimating daily number of servings of fruits and vegetables in a minority adolescent population. *J Am Diet Assoc* **109**, 1785–1789.
90. Huybrechts I, De Backer G, De Bacquer D, *et al.* (2009) Relative validity and reproducibility of a food-frequency questionnaire for estimating food intakes among Flemish preschoolers. *Int J Environ Res Public Health* **6**, 382–399.
91. Stiegler P, Sausenthaler S, Buyken AE, *et al.* (2009) A new FFQ designed to measure the intake of fatty acids and antioxidants in children. *Public Health Nutr* **13**, 38–46.
92. Zemel BS, Carey LB, Paulhamus DR, *et al.* (2009) Quantifying calcium intake in school age children: development and validation of the Calcium Counts! food frequency questionnaire. *Am J Hum Biol* **22**, 180–186.
93. Harnack LJ, Lytle LA, Story M, *et al.* (2006) Reliability and validity of a brief questionnaire to assess calcium intake of middle-school-aged children. *J Am Diet Assoc* **106**, 1790–1795.
94. Huybrechts I, De Bacquer D, Matthys C, *et al.* (2006) Validity and reproducibility of a semi-quantitative food-frequency questionnaire for estimating calcium intake in Belgian preschool children. *Br J Nutr* **95**, 802–816.
95. Haraldsdóttir J, Thórsdóttir I, Vaz De Almeida MD, *et al.* (2005) Validity and reproducibility of a precoded questionnaire to assess fruit and vegetable intake in European 11- to 12-year-old schoolchildren. *Ann Nutr Metab* **49**, 221–227.
96. Andersen LF, Bere E, Kolbjornsen N, *et al.* (2004) Validity and reproducibility of self-reported intake of fruit and vegetable among 6th graders. *Eur J Clin Nutr* **58**, 771–777.
97. Vereecken C a & Maes L (2003) A Belgian study on the reliability and relative validity of the Health Behaviour in School-Aged Children food-frequency questionnaire. *Public Health Nutr* **6**, 581–588.
98. Lambe J, Kearney J, Leclercrq C, *et al.* (2000) Enhancing the capacity of food consumption surveys of short duration to estimate long term consumer-only intakes by combination with a qualitative food frequency questionnaire. *Food Addit Contam* **17**, 177–187.
99. Taylor RW & Goulding A (1998) Validation of a short food frequency questionnaire to assess calcium intake in children aged 3 to 6 years. *Eur J Clin Nutr* **52**, 464–465.
100. Samuelson G, Bratteby LE, Enghardt H, *et al.* (1996) Food habits and energy and nutrient intake in Swedish adolescents approaching the year 2000. *Acta Paediatr Suppl* **415**, 1–19.
101. Bellú R, Riva E, Ortisi MT, *et al.* (1996) Validity of a food frequency questionnaire to estimate mean nutrient intake of Italian school children. *Nutr Res* **16**, 197–200.
102. Crawford PB, Obarzanek E, Morrison J, *et al.* (1994) Comparative advantage of a 3-day food records over 24-hour recall and 5-day food frequency validated by observation of 9 and 10 year old girls. *J Acad Nutr Diet* **94**, 626–630.
103. Frank GC, Nicklas TA, Webber LS, *et al.* (1992) A food frequency questionnaire for adolescents: defining eating patterns. *J Am Diet Assoc* **92**, 313–318.
104. Kunaratnam K, Halaki M, Ming L, *et al.* (2018) Reliability, comparative validity of a Diet Quality Index for assessing dietary patterns of preschool-aged children in Sydney, Australia. *Eur J Clin Nutr* **72**, 464–468.
105. Vereecken CA & Maes L (2003) A Belgian study on the reliability and relative validity of the Health Behaviour in School-Aged Children food-frequency questionnaire. *Public Health Nutr* **6**, 581–588.
106. Taylor C, Lamparello B, Kruczek K, *et al.* (2009) Validation of a food frequency questionnaire for determining calcium, vitamin D intake by adolescent girls with anorexia nervosa. *J Am Diet Assoc* **109**, 479–485.e3.

107. Araujo MC, Veiga GV da, Sichieri R, *et al.* (2010) Development of a semiquantitative food frequency questionnaire for adolescents from the Rio de Janeiro metropolitan area. *Rev Nutr* **23**, 179–189.
108. Rendo-Urteaga T, Saravia L, Sadalla Collese T, *et al.* (2020) Reliability and validity of an FFQ for South American children and adolescents from the SAYCARE study. *Public Health Nutr* **23**, 13–21.
109. Cade J, Thompson R, Burley V, *et al.* (2002) Development, validation and utilisation of food-frequency questionnaires – a review. *Public Health Nutr* **5**, 567–587.
110. Tabacchi G, Filippi AR, Amodio E, *et al.* (2016) A meta-analysis of the validity of FFQ targeted to adolescents. *Public Health Nutr* **19**, 1168–1183.
111. Börnhorst C, Bel-Serrat S, Pigeot I, *et al.* (2014) Validity of 24-h recalls in (pre-)school aged children: comparison of proxy-reported energy intakes with measured energy expenditure. *Clin Nutr* **33**, 79–84.
112. Hinnig P de F, Mariath AB, Freaza SRM, *et al.* (2014) Development of a food frequency questionnaire for children from 7 to 10 years old. *Rev Bras Epidemiol* **17**, 479–494.

British Journal of Nutrition