

### SYSTEMATIC REVIEW

## Quality of food-frequency questionnaire validation studies in the dietary assessment of children aged 12 to 36 months: a systematic literature review

Amy Lovell<sup>1\*</sup>, Rhodi Bulloch<sup>2</sup>, Clare R. Wall<sup>3</sup> and Cameron C. Grant<sup>4,5,6</sup>

<sup>1</sup>*Discipline of Nutrition, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand*

<sup>2</sup>*Discipline of Nutrition, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand*

<sup>3</sup>*Discipline of Nutrition, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand*

<sup>4</sup>*Department of Paediatrics: Child and Youth Health, University of Auckland, Auckland, New Zealand*

<sup>5</sup>*Centre for Longitudinal Research He Ara ki Mua, University of Auckland, Auckland, New Zealand*

<sup>6</sup>*Starship Children's Hospital, Auckland District Health Board, Auckland, New Zealand*

(Received 27 September 2016 – Final revision received 30 November 2016 – Accepted 3 March 2017)

*Journal of Nutritional Science* (2017), vol. 6, e16, page 1 of 12

doi:10.1017/jns.2017.12

#### Abstract

A child's diet is an important determinant of growth and development. Because of this, the accurate assessment of dietary intake in young children remains a challenge. A systematic search of studies validating FFQ methodologies in children 12 to 36 months of age was completed. English-language articles published until March 2016 were searched using three electronic databases (MEDLINE, EMBASE and CINAHL). Quality assessment of the identified studies was carried out using The Reduced Summary Score and EUROpean micronutrient RECommendations Aligned (EURRECA) scoring system. Seventeen studies were included and categorised according to whether they reflected long-term ( $\geq 7$  d) or short-term ( $< 7$  d) intake, or used a biomarker. A total score for each micronutrient was calculated from the mean of the correlation coefficients weighted by the study quality score. At least three validation studies per micronutrient were required for inclusion. Fifteen studies (83 %) that considered validity of the FFQ in assessing nutrient intakes had quality scores from 2.5 to 6.0. Of those, ten (67 %) studies found FFQ to have good correlations in assessing dietary intake ( $> 0.4$ ). Of the nutrients with three or more studies available, FFQ validated using a reference method reflecting short-term intake had a good weighted correlation for Ca (0.51), and acceptable weighted correlations for vitamin C (0.31) and Fe (0.33). Semi-quantitative FFQ were shown to be valid and reproducible when estimating dietary intakes at a group level, and are an acceptable instruments for estimating intakes of Ca, vitamin C and Fe in children 12 to 36 months of age.

**Key words:** Food-frequency questionnaires; Infants; Validity; Dietary assessment methods

The accurate description and measurement of dietary intake is a necessary step in determining the nutritional adequacy of diets in individuals or a population<sup>(1)</sup>. Having valid and reliable assessment tools is essential to increase our understanding of the relationship between dietary intake and health outcomes, and our understanding of the dietary determinants of nutritional status<sup>(2)</sup>.

Food and nutrient intakes are estimated via dietary assessment methods that differ according to a study's aims and

objectives, skills of the study population, accuracy of the required dietary data, study resources and study design<sup>(3)</sup>. Most epidemiological studies use variations of the FFQ, which can be validated using biomarkers or tools that measure daily dietary intake<sup>(5)</sup>. The FFQ has an advantage of being an inexpensive method of obtaining data from a large number of participants, with a relatively low respondent burden and can be used to estimate an individual's average consumption over an extended period of time<sup>(3,6)</sup>.

**Abbreviations:** 24-HR, 24-h recall; EURRECA, EUROpean Micronutrient RECommendations Aligned; WFR, weighed food record.

\* **Corresponding author:** A. Lovell, email [a.lovell@auckland.ac.nz](mailto:a.lovell@auckland.ac.nz)



There is no definitive 'gold standard' in dietary assessment, nor is there a 'gold standard' for assessing the validity of FFQ<sup>(7)</sup>. Therefore estimation of a tool's relative validity relies upon a comparison with a superior and preferably independent technique, known as comparative validation<sup>(3)</sup>. Here, weighed food records (WFR) and 24-h recalls (24-HR) are commonly used due to their greater precision in the quantification of intake<sup>(3)</sup>. Factors that may affect the validity of a diet questionnaire have been reviewed<sup>(5,8)</sup>.

Early childhood is a life phase where the assessment of dietary intake is particularly challenging. Measurement of energy and nutrient intakes in young children is affected by unique respondent and observer considerations, making the collection of accurate and reliable dietary intakes difficult<sup>(1)</sup>. Young children aged 12 to 36 months, have highly variable diets that are characterised by rapidly changing food habits and transitions in dietary patterns, and often not all food served to an infant is consumed in its entirety<sup>(9–12)</sup>. The acquisition of dietary intake information for children less than 7 years of age is dependent upon surrogate reporters, e.g. parents, caregivers and external caretakers<sup>(1,13)</sup>. Therefore, the accuracy of dietary assessment in this age group depends on an adult's ability to reliably report on their intake, with previous evidence suggesting that parents can provide a more reliable report on foods consumed in the home setting, rather than away from home<sup>(1,13)</sup>.

As a consequence of these methodological challenges, the number and type of validated tools available to assess the dietary intake of young children, particularly children 12 to 36 months of age, are limited. The aim of this systematic literature review was to describe and assess the quality of studies reporting on the validity of FFQ as a method for assessing food and nutrient intakes or dietary patterns in 12- to 36-month-old children.

## Methods

### Protocol registration

The inclusion and exclusion criteria, and analysis methods were specified in advance in a documented protocol. This protocol was not registered with PROSPERO<sup>(14)</sup> as it is an assessment of the quality of validation studies and does not report on a health-related outcome.

### Eligibility criteria

Studies that evaluated the validity of FFQ in the assessment of dietary intake, food(s), and dietary patterns with a reference dietary assessment tool (e.g. 24-HR, diet records, diet histories, WFR and biomarkers) in healthy children aged 12 to 36 months and met all the inclusion criteria (Fig. 1) were included in the review. Randomised controlled trials were not available; therefore analytical study designs were limited to prospective and retrospective cohort studies. Case series, case reports and case-control studies were excluded due to the high potential for bias.

### Information sources

Studies were identified via searching online databases, hand-searching reference lists of original articles, and cited reference

searches. The search focused on relevant studies published before March 2016 and was limited to those published in English, without limits on time frame or country. Grey literature was also considered.

### Search strategy

A literature search was applied to MEDLINE (1946 to present), EMBASE (1980 to present) and CINAHL (1937 to present) electronic databases, and Google Scholar. Medical Subject Headings (MeSH), MeSH major topics, and free text terms were developed under four group headings in MEDLINE and EMBASE databases. The MeSH search terms used in the search were developed under four group headings: (1) infant (12–36 months), e.g. toddler, preschool\*, child, infant, newborn\*, pre-school\*, babies, baby, kindergarten, children under 2, children under 3; (2) diet, e.g. nutrition, dietary pattern, food intake, diet quality, infant nutrition, child nutrition, nutritional assessment, eating pattern, nutritional status, feeding behaviour, food combination, childhood diet, infant food; (3) dietary assessment, e.g. diet surveys, questionnaires, instrument, dietary intake methods, assess\*, evaluat\*, dietary intake methods, nutrition surveys; (4) dietary assessment tool, e.g. food frequency questionnaire, FFQ; (5) instrument validation, e.g. validity, reproducibility, correlation coefficient, reliability, validation studies, replication stud\*, correlation stud\*, repeatability. Key words and combinations were identified in free text, article titles and abstracts, and were used to perform a comprehensive search of the databases. Search terms and strategies were adapted for use in other databases and were peer reviewed. All retrieved articles were sent to Refworks<sup>®</sup> (version 4.4.1237; ProQuest LLC) where duplicates were removed.

### Study selection

Two reviewers (A. L. and R. B.) determined a study's eligibility in an independent, unblinded and standardised manner. Systematic literature reviews were not included in the analysis. Titles and abstracts were reviewed to assess whether they met the inclusion criteria for full-text review (Fig. 1). Disagreements between reviewers were resolved by consensus, or if the decision on study inclusion or exclusion were unclear, the full text was obtained. In studies where the age range of participants was included, but was much wider than 12 to 36 months, e.g. 2 to 9 years, the reviewers attempted to obtain results from authors specific to the age range of interest. Full-text articles that fulfilled all criteria for inclusion were reviewed in a second screening process as the definitive step for inclusion.

### Data collection process

A data extraction sheet based on examples found in the selected literature was developed. One review author (A. L.) extracted key data into a prepared table, which was checked by a co-author (R. B.). Any disagreements were resolved through discussion between the review authors (A. L. and



Inclusion criteria: studies with all of the following characteristics were included:

- Validation studies in human subjects
- Studies involving healthy children (aged 12 to 36 months)
- Studies where dietary intake is reported by parent/caregiver or is interviewer-administered
- Studies where both FFQ and reference dietary assessment method are conducted within the age range of interest
- Studies where dietary intake is assessed using an assessment tool, such as an FFQ that measures food(s), nutrient intakes and/or food groups, and the dietary assessment tool was validated
- Studies that have measured dietary intake, dietary patterns, and nutrient intake and/or specific components of dietary intake, for example specific micronutrients in children (aged 12 to 36 months)

Exclusion criteria: studies with any one of these characteristics were excluded:

- Studies focused on diseased, unhealthy or non-free-living participants
- Studies that do not include the target age range of 12 to 36 months
- Studies that investigate dietary causes of specific diseases, for example food allergy
- Studies relating diseases to nutrient intake or food consumption
- Studies with no validated intake methods or with only portion size estimations
- Studies that present reference values for nutrient intake, food consumption, biological variables, biochemical markers and anthropometry
- Intervention/therapeutic studies of nutrients or drugs related to the metabolism of these nutrients
- Studies that assess malnutrition risk
- Studies that assess the physiological effects of nutrients and food, particularly in relation to genetic determinants
- Studies of eating behaviours

**Fig. 1.** Inclusion and exclusion criteria used to select studies for inclusion in the systematic review.

R. B.), and if no agreement could be reached a prearranged third reviewer was asked to arbitrate (C. W.). Direct contact via email was made with four authors to obtain information in addition to that which could be abstracted from the published paper. In all four cases this request was for information within the age range of interest (12 to 36 months) from a study that reported data over a wider age range. One follow-up email was sent if no response was received. No authors responded with data from their studies specific to the age range of interest.

### Data items

A concise overview of the seventeen included studies is shown in Table 1<sup>(11,15–24,26–31)</sup>. The areas of interest included: population characteristics (size, age, location, ethnicity), FFQ characteristics (food groups, food items, consumption interval, administration method, portion estimation, number of FFQ administered, and FFQ re-test interval), reference method used, outcome measures (validity, reproducibility) and the statistics employed to assess validity between two methods or reproducibility of the FFQ.

### Synthesis of results

Studies were classified into three categories based on the reference method applied to the validation study. This method has been previously reported and consisted of:

- (1) Long-term intake – the reference method covered  $\geq 7$  d.
- (2) Short-term intake – the reference method covered  $< 7$  d.
- (3) Biomarker – the reference method was a biomarker.

### Quality assessment

Following classification, the two reviewers (A. L. and R. B.) independently completed quality assessment of the included validation studies using the reduced summary score by Dennis *et al.*<sup>(32)</sup> which assessed the quality of the nutrition information from the FFQ, and an additional scoring system developed by the EUROpean Micronutrient RECommendations Aligned (EURRECA) network used in studies assessing nutrient intakes with the aim of including, excluding and weighting studies<sup>(5,12)</sup>. These scoring tools evaluated methodological quality of the identified studies and determined the extent to which a study addressed the possibility of bias in their design, conduct and analysis. This dual scoring system approach was used in a previous review of FFQ for assessing dietary intake in adolescents<sup>(33)</sup>.

Because of the heterogeneity between the dietary assessment methods used as the reference, study designs, populations, and duration of the study, only a narrative review of the literature was performed. A meta-analysis could not be conducted due to a lack of randomised controlled trials.

The summary score by Dennis *et al.*<sup>(32)</sup> scores studies based on objective measures of quality dietary assessment. The

**Table 1.** Characteristics of included studies evaluating long-term or short-term nutrient intake, or biomarker, food or food group

Reference (year), country	Population number (girls/boys), age range	FFQ food groups	FFQ food items	Consumption interval	FFQ type	Administration method	Portion estimation	No. of FFQ administered	FFQ interval (retest)	Reference method (characteristics)	Statistics used for validation	Correlation coefficient
<b>Long-term intake</b>												
Andersen <i>et al.</i> <sup>(11)</sup> (2003), Norway	64 (26/37), 12 months	15	140	Previous 14 d	Semi-quantitative	Self-administered (parent)	Household units Photographs: 16 photographs (small, medium, large)	1	NA	7 d WFR 1–2 weeks following FFQ 1-week interval between FFQ and WFR Two periods: 4 d (rec. 1), 3 d (rec. 2)	Spearman CC, energy-adjusted CC, cross-classification, Bland–Altman	Absolute intake: 0.47*, 0.50† Nutrient density: 0.51*, 0.50† Food intake: 0.62†
Andersen <i>et al.</i> <sup>(15)</sup> (2004), Norway	187 (88/99), 24 months	15	125	Previous 14 d	Semi-quantitative	Self-administered (parent)	Household units Photographs: 16 photographs (small, medium, large)	1	NA	7 d WFR 1–2 weeks following FFQ	Spearman's CC, energy-adjusted CC, Bland–Altman, cross-classification	Absolute intake: 0.36*, 0.38† Nutrient density: 0.53*, 0.52† Food intake: 0.48†
<b>Short-term intake</b>												
Iannotti <i>et al.</i> <sup>(16)</sup> (1994), USA	17 (8/9), 2–4 years	NR	111	Previous 7 d	Willett (semi-quantitative)	Interviewer-administered	Free text for portion sizes	1	NA	3 d WFR Parent-completed Food and drink not consumed Recorded data checked	Pearson's CC	0.37 *
Blum <i>et al.</i> <sup>(17)</sup> (1999), USA	233, 1–5 years	NR	84	4 weeks	Adapted Willett (semi-quantitative)	Self-administered (parent)	Age-appropriate	2	1 month	Three 24 h recalls Over 4 weeks (2 weekdays, 1 weekend) Month between HFFQ1 and HFFQ2	Pearson's CC, adjusted energy intake,	0.52 (all) 0.51 (1–2 years)
Parrish <i>et al.</i> <sup>(18)</sup> (2003), Canada	68, 1–3 years	NR	111	Previous 12 months	Willett (semi-quantitative)	Self-administered (parent)	NR	2	NA	Interviewer-administered 24HR (x3 or 4) Completed prior to FFQ 1 × 24HR every 3 months Interviewer-administered Biomarkers (n 38) Total lipids calculated (vitamins C, D and E, retinol, β-carotene)	Energy-adjusted, Pearson CC, Spearman's CC	0.33*: 1 year 0.30*: 2 years 0.27*: 3 years 0.32*: all years
Marshall <i>et al.</i> <sup>(19)</sup> (2003), USA	240 (119/121), 6 months–5 years	7	NR	Previous week	Quantified beverage frequency questionnaire	Self-administered (parent)	Beverage consumption in servings per week	18	NA	3 d food and beverage diary 1 weekend + 2 weekdays Parent-completed	Spearman's CC, weighted κ	Ca 0.62* all 0.64* drinkers Vitamin D 0.63* all 0.70* drinkers 0.63*
Williams & Innis <sup>(20)</sup> (2005), Canada	148, 8–26 months	21	191	Previous 2 weeks	Quantitative	Interviewer-administered	NR	1	NA	3 d FR (24HR completed prior) Parent-completed 2 weekdays, 1 weekend Biomarkers Serum markers of Fe Food pictures, 3D models, household measures	Spearman's CC, energy-adjusted CC	0.71* (energy adjusted)
Marriott <i>et al.</i> <sup>(21)</sup> (2009), UK	50 (23/27), 12 months	NR	78	Previous 28 d	Semi-quantitative	Interviewer-administered	Household measures, food models	1	NA	4 d WFR following FFQ	Spearman CC, energy-adjusted CC, Bland–Altman, LOA	0.49* 0.48* (energy adjusted)



Vereecken <i>et al.</i> <sup>(22)</sup> (2010), Belgium	216 (112/104), 3-5 years (mean age)	77	NR	Previous 3 months	Quantitative	Self-administered (parent)	4-8 portion sizes for each item Open category	1	NA	Online dietary A <sub>x</sub> tool Parent-completed online 3 non-consecutive days (1 weekend, 2 weekdays) over 2-week period	Spearman's CC, Bland-Altman	0.42*
Rankin <i>et al.</i> <sup>(23)</sup> (2011), USA	753 (388/365), 6-60 months	NR	16 (6-20 months), 11 (20-60 months)	Previous 7 d	Semi-quantitative	Self-administered (parent)	NR	1	NA	3 d FD Over 72 h (1 weekend, 2 weekdays) After the FFQ	Spearman's CC	0.73*
D'Ambrosio <i>et al.</i> <sup>(24)</sup> (2012), Canada	22, 12-35 months	10	97	Previous 4 weeks	HFFQ (semi-quantitative)	Interviewer-administered	Household measures, food models	2	1 month (n 14)	24HR Parent-completed over 2 weeks Completed 2 and 4 weeks following FFQ (same day as FFQ) Reliability in subset: 4 weeks following FFQ	Pearson's CC, Bland-Altman	0.41* (CC)
Watson <i>et al.</i> <sup>(26)</sup> (2015), New Zealand	160, 12-24 months	16	91	Previous 4 weeks	Southampton Women's Survey FFQ (Quantitative)	Interviewer-administered	'Palm' sizes	2	4 weeks	5 d WFR Non-consecutive days randomly assigned Completed over 5 weeks	Pearson's CC, cross-classification, ICC, Bland-Altman	0.50 (unadjusted) 0.52 (adjusted)
Sochacka-Tatara & Pac <sup>(27)</sup> (2013), Poland	143 (68/75)	95	95	Previous 4 weeks	Semi-quantitative	Self-administered (parent)	Household units	1	NA	24HR (x3) 3 consecutive days (average taken) Weekdays + weekends Interval between FFQ and 24HR ≤ 1 month Parent-completed	Spearman's CC, Bland-Altman	0.46*
Biomarker Orton <i>et al.</i> <sup>(28)</sup> (2008), USA	404 (107/209), 1-11 years	NR	111	Previous year	Semi-quantitative	Self-administered (parent)	Commonly used portion sizes, e. g. slice of bread	1	NA	Biomarker Fatty acid composition measured in erythrocytes from blood sample	CC, linear model	0.10* (energy adjusted) 0.08* (% total fat from FFQ)
Food or food group intake Klohe <i>et al.</i> <sup>(29)</sup> (2005), USA	77, 1-3 years	91	191	Previous month	Semi-quantitative	Self-administered (parent)	Serving sizes: small, medium, large, extra-large (as multiples of a medium serving size)	2	2 weeks	3 d DR (24HR + 2 d DR) 24HR: interviewer-administered Household measures, food models 2 d DR: weekend + week day	Spearman's CC, cross-classification	0.41*
Bel-Serrat <i>et al.</i> <sup>(30)</sup> (2011), Europe	2508 (1244/1264), 2-9 years	43	NR	Previous month	Children's Eating Habits Questionnaire (non-quantitative)	Self-administered (parent)	NR	1	NA	24HR (x2) Computerised Non-consecutive days Interviewer-administered	Pearson's CC (de-attenuated), Spearman's CC, cross-classification, Kruskal-Wallis, Bland-Altman, LOA, weighted κ	0.25* (age 2-6 years)
Mills <i>et al.</i> <sup>(31)</sup> (2015), New Zealand	153 (51/78), 12-24 months	16	91	Previous 4 weeks	Adapted Southampton Women's Survey FFQ (quantitative)	Interviewer-administered	'Palm' sizes	2	4 weeks	5 d WFR Non-consecutive days randomly assigned Completed over 5 weeks	Pearson's CC, cross-classification, ICC	0.61* 0.58†

3D, three-dimensional; 24HR, 24 h recall; CC, correlation coefficient; DR, diet record; FD, food diary; FR, food record; HFFQ, Harvard Service Food Frequency Questionnaire; ICC, intra-class correlation; LOA, limits of agreement; NA, not applicable; NR, not reported; rec., record; WFR, weighed food record.

\* Mean.

† Median.





reduced summary score with a maximum score of 8 was utilised for simplified quality assessment of the FFQ as seen in Tabacchi *et al.*<sup>(33)</sup> Validation studies that had a reduced summary score of  $\geq 5$  were classified as being 'high quality' and scores  $< 5$  as 'low quality'. This scoring tool was used for all included studies. The EURRECA<sup>(5)</sup> scoring system was only applied to studies that assessed nutrient intakes. Summary scores range from 0 (poorest quality) to 7 (highest possible score) and are ranked as 'very good/excellent' score  $\geq 5$ ; 'good' score  $3.5 \leq$  and  $< 5$ ; 'acceptable' score  $2.5 \leq$  and  $< 3.5$ ; and 'poor' score  $< 2.5$ <sup>(5)</sup>. In order to estimate a mean correlation per micronutrient for the included studies, the correlation coefficient from each study was initially multiplied by its quality score. Next, the sum of the weighted correlations was divided by the sum of the quality scores to provide a correlation coefficient that was adjusted for the study's methodological quality. Mean weighted correlation coefficients were only calculated for micronutrients with correlations available from three or more studies<sup>(34)</sup>. This allows for concurrent analysis of multiple validation studies and gives an estimate of a mean correlation coefficient per micronutrient for a given dietary assessment method<sup>(5)</sup>. The intake method was rated as poor when the correlation was  $< 0.30$ , acceptable between 0.30 and 0.50, good between 0.51 and 0.70, and correlations  $> 0.70$  were very good<sup>(5)</sup>.

## Results

### Study selection

A total of 373 articles were identified (Fig. 2). Following removal of duplicates, 236 articles unique by title and abstract remained for review. Application of inclusion and exclusion criteria resulted in fifty-nine articles being selected for full-text review. Thirty-nine studies were included for quality appraisal. All studies were cross-sectional in their design, and thus classified as level IV evidence<sup>(35)</sup>. Following quality appraisal twenty-two studies were excluded, leaving seventeen articles<sup>(11,15-24,26,27-29,31,37)</sup> identified as assessing the validity of an FFQ against a dietary reference instrument in children 12 to 36 months of age.

Nine of the publications reported results from North American countries (USA and Canada)<sup>(16-20,23,24,28,29)</sup>, five from the UK and Europe<sup>(11,15,21,27,30)</sup>, and three from New Zealand<sup>(26,31,38)</sup>. The number of participants ranged from seventeen<sup>(16)</sup> to 240<sup>(19)</sup>, with two studies presenting data from large cohorts: The Iowa Fluoride Study<sup>(25)</sup> and The IDEFICS Study (Identification and prevention of Dietary- and lifestyle-induced health EFfects In Children and infantS)<sup>(37)</sup>.

### Characteristics of included studies

Characteristics of each of the seventeen included validation studies are described in Table 1. Fourteen studies considered the validity of the FFQ to assess nutrient intakes<sup>(11,15-24,27,28)</sup>, and three studies considered values on the validity of the FFQ to assess food or food group(s)<sup>(29,31,37)</sup>. Two studies assessing nutrient intakes also used biomarkers as an additional reference method<sup>(18,20)</sup>. Eleven of the included FFQ were semi-quantitative<sup>(11,15-19,21,23,27-29)</sup>,

five were quantitative<sup>(19,20,22,26,31)</sup>, and one recorded frequency of consumption and not portion sizes<sup>(37)</sup>. The number of food items ranged from seventy-eight<sup>(21)</sup> to 191<sup>(20,29)</sup> with an average of 113 food items. Those studies that assessed food and/or food group intakes had between seven<sup>(19)</sup> and seventy-seven<sup>(22)</sup> food groups. Food intake intervals ranged from intake over the previous 7 d<sup>(19,23)</sup> to over the last year<sup>(18,28,28)</sup>, with the majority describing intake over the last month<sup>(17,21,24,27,29,37)</sup>.

Two studies were grouped according to a reference method that reflected long-term intake (7-d WFR)<sup>(11,15)</sup>. Ten studies were grouped according to a reference method that reflected short-term intake where four applied 24-HR<sup>(17,18,24,27)</sup> and five applied WFR<sup>(19-23)</sup>, one of these being online<sup>(22)</sup>. One study utilised biomarkers as a reference method<sup>(28)</sup>. Among the seven studies that used WFR, the number of recorded days varied from 3 to 7 d<sup>(11,15,19,20-22,29)</sup>. The number of repeated 24-HR ranged from 2<sup>(24,37)</sup> or 3<sup>(17,20,27)</sup> days of non-consecutive administration. Eleven studies were self-administered<sup>(11,15,17-19,22,23,27-29,37)</sup>, by a parent or equivalent proxy reporter, and six studies were interviewer administered<sup>(16,20,21,24,26,31)</sup>. Methods of portion size estimation ranged from household measures/standard portion sizes<sup>(11,15,21,24,27,28)</sup> to portion sizes derived from national nutrition survey data<sup>(17,22,29)</sup>. Three studies did not describe portion estimation<sup>(18,20,23)</sup>, and two studies used a unique 'palm' measurement<sup>(26,31)</sup>. Of the thirteen studies that calculated food intakes into nutrient intakes, six reported using national food composition databases (e.g. United States Department of Agriculture)<sup>(11,15,22-24,28)</sup>, and two used other food composition databases (e.g. Harvard Nutrient Database)<sup>(17,18)</sup>. Although not the primary aim of the validation study, two studies<sup>(16,18)</sup> examined whether there were any differences between sex and care status (i.e. in child care or at home) when comparing mean nutrient intake values.

### Statistical analysis

Statistical analyses used in the assessment of FFQ validity, and in some cases reproducibility, are described in Table 1. All included studies calculated differences in means and/or mean comparisons. Pearson or Spearman's correlation coefficients were calculated in all studies. Paired Student's *t* tests were used evaluate whether there was any difference between the mean nutrient and food intakes determined by the two assessment methods<sup>(18)</sup>. Factors that affect the validity of a dietary assessment instrument included: population characteristics, acceptability of the reference method data, FFQ design/quantification, quality control and data management<sup>(5,33)</sup>.

The calculation of weighted correlation coefficients allowed comparison with the other included studies. Here, correlation coefficients between 0.51 and 0.7 are considered good<sup>(5,7)</sup>. Four studies considered crude correlation coefficients<sup>(16,19,22,23)</sup>, whilst seven studies adjusted nutrients using energy-adjusted values<sup>(11,15,17,18,20,21,28)</sup>, and three studies calculated de-attenuated values to account for measurement error<sup>(26,27,31)</sup> or intra-class correlations<sup>(24)</sup>. All six studies that performed cross-classification analysis ranked participants by using the same or adjacent quartile. Three of these studies<sup>(11,15,26)</sup> assessed the classification of participants according to their nutrient intakes and three studies<sup>(29,31,37)</sup>

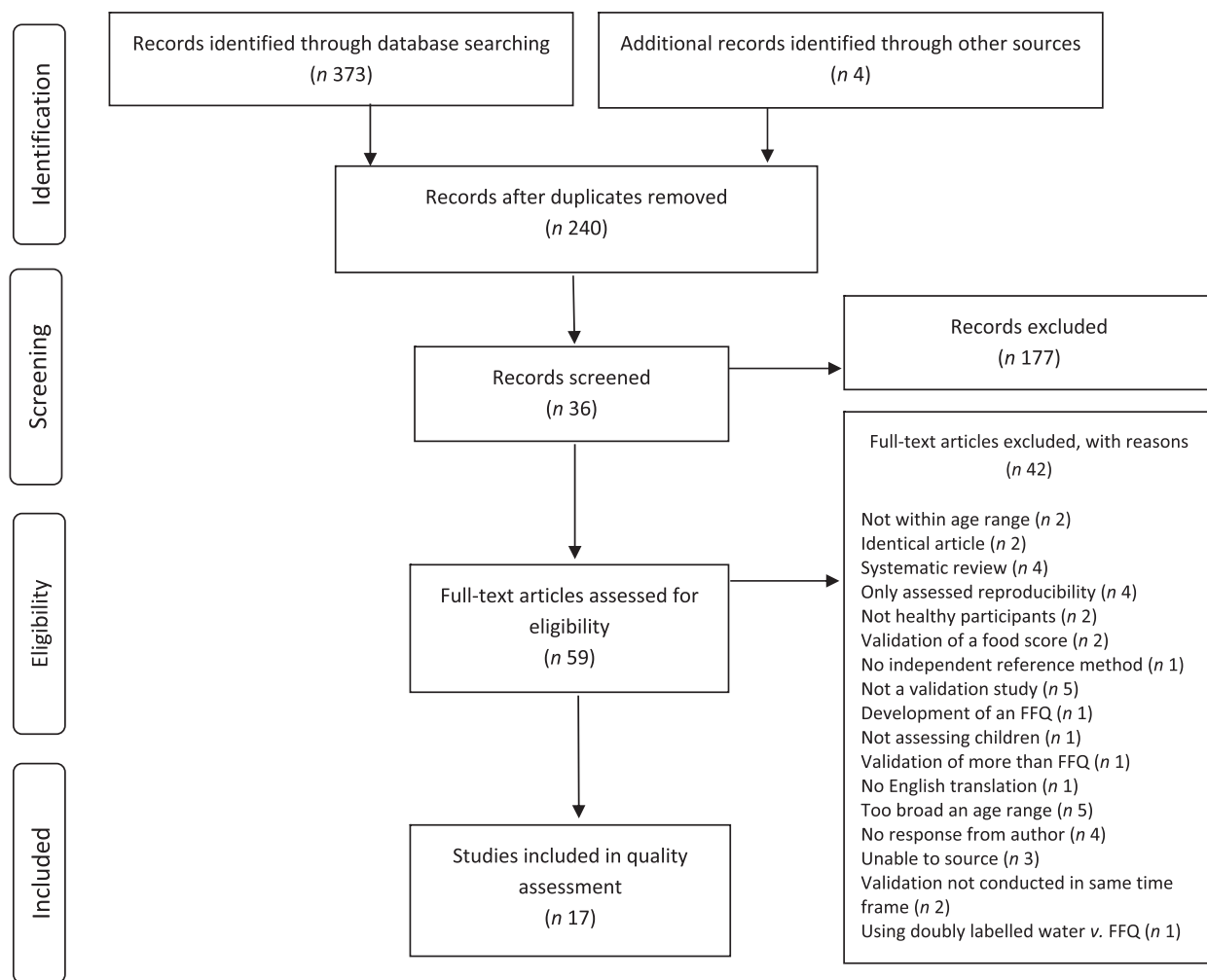


Fig. 2. Selection process flow of articles identified that assess validity of FFQ methods in children aged 12–36 months.

assessed the classification of participants according to their food or food group intakes. Weighted  $\kappa$  was calculated in two studies that considered food intakes<sup>(19,23)</sup>. Here, four categories were used to calculate  $\kappa$  statistics and classify food intake data.

Two studies<sup>(24,26)</sup> assessed the reproducibility of the FFQ for estimating dietary intake patterns and estimation of reproducibility of nutrient intakes was achieved by calculating correlation coefficients and intra-class correlations. Acceptable intra-class correlations ranged from  $>0.4$ <sup>(7,26,39)</sup> to  $0.7$ <sup>(24)</sup> when establishing test–retest reliability of the FFQ. In order to test reproducibility, five<sup>(17,24,26,29,31)</sup> studies administered the FFQ on two occasions. Intervals between test and retest ranged from 2 weeks<sup>(29)</sup> to 1 month<sup>(17,24,26,31)</sup>. One study<sup>(17)</sup> administered the FFQ on two occasions, 1 month apart but did not report on the statistical analysis used for reproducibility.

#### Results of individual studies by validation method used

Included reviews were analysed according to the reference method used (i.e. WFR, 24-HR or biomarker) and whether the tool reflected long-term or short-term intake.

**FFQ v. 24-h recalls.** Five studies<sup>(17,18,24,27,37)</sup> used 24-HR as their reference method to validate an FFQ. In all studies the FFQ overestimated median/mean nutrient intake estimates but could provide reliable estimates of nutrient intakes in young children with good agreement when compared with the 24-HR (Table 1). Nutrient correlations that were energy-adjusted or de-attenuated (to reduce dependency on between-person variation) were found to have higher correlation coefficients compared with crude values. Cross-classification into low, medium and high consumers was moderate ( $>30\%$  classification into the same quartile). One study<sup>(24)</sup> assessed repeatability/reproducibility using a 24-HR as a reference tool. Correlations for most nutrients were  $>0.70$ , indicating low within-person variation.

**FFQ v. food record ( $\pm$ weighing).** Eleven studies used WFR as their reference method to validate an FFQ<sup>(11,15,16,19,20,21,22,23,26,29,31)</sup>. Ten studies that estimated nutrient intakes found that the FFQ tended to overestimate intakes (Table 1) but found good correlations ( $>0.4$ )<sup>(7)</sup> between the FFQ and WFR for most nutrients, energy



intakes and food intakes. The included FFQ mostly indicated a moderate ability to rank infants according to their nutrient intakes, with two studies by Andersen *et al.*<sup>(11,15)</sup> showing that the ability of the questionnaire to rank infants according to their intakes increased when using nutrient density values over absolute values.

**FFQ v. biomarker.** Using biomarkers as the reference method was less frequent. Three studies used biomarkers<sup>(15,24,35)</sup>. Two articles<sup>(18,20)</sup> presented validation of an FFQ using biomarkers and a second dietary assessment instrument (24-HR or WFR) as reference methods. The biomarkers analysed included: total lipids, plasma levels of vitamins C, D and E, retinol and  $\beta$ -carotene<sup>(18)</sup>, serum markers of Fe<sup>(20)</sup> and fatty acid composition measured in erythrocytes<sup>(28)</sup>.

**Evaluation of food or food groups.** Using a semi-quantitative FFQ excellent reliability and adequate validity were seen in assessing food choices of low-income children<sup>(29)</sup>, with low levels of agreement and limited ability to rank children according to intakes of food groups<sup>(37)</sup>. More recently, in Otago, New Zealand, a semi-quantitative FFQ displayed good validity ( $r$  0.52) and high reproducibility in the identification of dietary patterns, and in ranking the diets of toddlers when compared with a 5-d WFR. The FFQ overestimated energy and nutrient intakes and cannot measure absolute intakes, but could be used to identify toddlers at extreme ends of intake distribution<sup>(26,31)</sup>.

#### Additional analysis: quality assessment

A summary of the quality assessment of the seventeen included studies are shown in Table 2. Using the reduced summary score<sup>(32)</sup>, one validation study that assessed nutrient intakes received a low quality ranking<sup>(19)</sup> and one study that assessed food intake received a low quality ranking<sup>(37)</sup>. The remaining fifteen studies received high quality rankings. Criteria that reduced the quality of the study included the number of food items in the FFQ (<70 food items is likely to reduce the quality of the nutrition information), and if the FFQ was self-administered.

Using the EURRECA scoring system<sup>(5)</sup>, fourteen studies assessed nutrient intakes, with quality scores ranging from 2.5 to 6.0 (maximum 7.0). The average quality score was 3.8, with a median of 3.5. Table 2 illustrates the classification of the included studies according to their reference method and methodological quality, with three studies<sup>(20,26,27)</sup> (21 %) rating as very good/excellent, five studies<sup>(11,15,18,21,24)</sup> (36 %) as good quality, five studies<sup>(16,19,22,23,28)</sup> (36 %) having an acceptable quality, and one study<sup>(23)</sup> (7 %) having a poor quality rating. 'Good' quality scores were seen in the validation studies where FFQ were compared with a reference method that was reflective of long-term intakes, and a majority (58 %) of validation studies where the FFQ was compared with a reference method that was reflective of short-term intakes were either 'good' or 'very good'. Factors affecting the EURRECA quality assessment score<sup>(5)</sup> were the statistical

**Table 2.** Quality scores using methods described by Dennis *et al.*<sup>(32)</sup> and the EUROpean Micronutrient RECommendations Aligned (EURRECA) scoring tool<sup>(5)</sup>

Study	Quality level (Dennis <i>et al.</i> ) <sup>(32)*</sup>	Quality score <sup>(5)†</sup>	EURRECA classification
<b>Nutrient intake</b>			
Williams & Innis <sup>(20)</sup>	High	6.0	Excellent
Watson <i>et al.</i> <sup>(26)</sup>	High	5.0	
Sochacka-Tatara & Pac <sup>(27)</sup>	High	5.0	Good
D'Ambrosio <i>et al.</i> <sup>(24)</sup>	High	4.5	
Marriott <i>et al.</i> <sup>(21)</sup>	High	4.0	
Andersen <i>et al.</i> <sup>(11)</sup>	High	3.5	
Andersen <i>et al.</i> <sup>(15)</sup>	High	3.5	
Parrish <i>et al.</i> <sup>(18)</sup>	High	3.5	
Orton <i>et al.</i> <sup>(28)</sup>	High	3.0	
Iannotti <i>et al.</i> <sup>(16)</sup>	High	3.0	
Blum <i>et al.</i> <sup>(17)</sup>	High	3.0	Poor
Marshall <i>et al.</i> <sup>(19)</sup>	Low	3.0	
Vereecken <i>et al.</i> <sup>(22)</sup>	High	3.0	
Rankin <i>et al.</i> <sup>(23)</sup>	High	2.5	
<b>Food or food group intake</b>			
Klohe <i>et al.</i> <sup>(29)</sup>	High	NA	NA
Bel-Serrat <i>et al.</i> <sup>(37)</sup>	Low	NA	NA
Mills <i>et al.</i> <sup>(31)</sup>	High	NA	NA

NA, not available, fewer than three studies found.

\* Dennis *et al.*<sup>(32)</sup> quality level: high ( $\geq 5$ ); low ( $< 5$ ).

† EURRECA quality score: very good/excellent ( $\geq 5$ ); good (3.5 to  $< 5$ ); acceptable/reasonable (2.5 to  $< 3.5$ ); poor ( $< 2.5$ ).

analyses used and data collection via interviewer-administration. Calculation of energy-adjusted<sup>(11,15,17,18,21,24,28)</sup>, de-attenuated (to reduce the dependency on between-person variation)<sup>(26,27)</sup>, or intra-class correlation coefficients increased quality scores<sup>(24)</sup>.

**Concurrent validation analysis.** Table 3 displays concurrent analysis of the included studies where a mean correlation coefficient per nutrient for each dietary assessment method was calculated by multiplying the correlation coefficient by their quality assessment score. This was completed for the EURRECA priority micronutrients and those studies that met the criteria of having nutrient correlations from at least three studies<sup>(40)</sup>. Micronutrients with a sufficient number of studies to be included ( $\geq 3$  studies), and where the validation reference method reflected short-term intakes (<7 d), were vitamin B<sub>12</sub><sup>(17,21,24,26,27)</sup>, vitamin C<sup>(17,20,21,24,26,27)</sup>, vitamin D<sup>(19,21,27)</sup>, Ca<sup>(17,19–22,24,26,27)</sup>, Fe<sup>(17,20,21,24,26,27)</sup> and Zn<sup>(17,21,26,27)</sup>. Fibre<sup>(17,20,22,24,26,27)</sup> and vitamin E<sup>(17,18,21,27)</sup> also had sufficient studies to allow for concurrent analysis. There were insufficient data available for the analysis of two (20 %) micronutrients: folate and Cu. Using the EURRECA scoring tool classifications, correlations were acceptable for vitamin B<sub>12</sub> (0.30), vitamin A (0.34) and Ca (0.49) using FFQ v. 24-HR whilst Fe showed a poor correlation (0.29) on validation. Acceptable correlations were seen for vitamin C (0.32) and Fe (0.39), and Ca presented a good correlation (0.51) using FFQ v. WFR. The intake method was rated as 'good' when the mean correlation coefficient weighted by the quality criteria score was at least 0.5. The number of studies that used a validation reference method that reflected long-term intakes





**Table 3.** Classification of dietary assessment methods for infants aged 12–36 months according to the weighted mean of the correlations of micronutrients with three or more studies available (separate comparisons of those studies reflecting long-term and short-term intakes or comparison of FFQ with a reference method)

Micronutrient	Correlation*				
	Long-term intake (>7 d)	Short-term intake (<7 d)	FFQ v. WFR	FFQ v. 24HR	FFQ v. BM
Vitamin B <sub>12</sub>	NA	P – 0.20 (5 studies)	NA	A – 0.30 (3 studies)	NA
Folate	NA	NA	NA	NA	NA
Vitamin C	NA	A – 0.31 (7 studies)	A – 0.32 (3 studies)	A – 0.34 (4 studies)	NA
Vitamin D	NA	P – 0.13 (3 studies)	NA	NA	NA
Ca	NA	G – 0.51 (8 studies)	G – 0.51 (7 studies)	A – 0.49 (3 studies)	NA
Fe	NA	A – 0.33 (6 studies)	A – 0.39 (5 studies)	P – 0.29 (3 studies)	NA
Zn	NA	P – 0.15 (4 studies)	NA	NA	NA
Cu	NA	NA	NA	NA	NA
Fibre	NA	P – 0.29 (6 studies)	A – 0.31 (3 studies)	P – 0.24 (3 studies)	NA
Vitamin E	NA	P – 0.11 (4 studies)	NA	P – 0.23 (3 studies)	NA

WFR, weighed food record; 24HR, 24 h recall; BM, biomarker; NA, not available, fewer than three studies found.

\* Correlation: G, good (0.51–0.70); A, acceptable (0.30–0.50); P, poor (<0.30).

(>7 d) were insufficient for concurrent analysis (<3 studies per micronutrient).

## Discussion

In this review, using standardised quality assessment methods, we evaluated seventeen studies reporting on the validity of FFQ as a method for assessing food and nutrient intakes or dietary patterns in 12- to 36-month-old children. From the identified studies<sup>(11,15–20,21,23,24,26,27–31,38)</sup>, semi-quantitative FFQ were shown to be valid and reproducible instruments in children as young as 1 year of age, generating adequate estimates specifically for Ca, vitamin C and Fe, with results similar to those seen in older children and adolescents<sup>(18,22)</sup>.

FFQ are used to assess dietary intake due to their practicality, relative ease of administration, low participant burden, ability to assess intake over a prolonged period of time, and lower associated costs<sup>(41,42)</sup>. However, there are limited FFQ that have been specifically validated in 12- to 36-month-old children. In the present review, the methodological qualities of FFQ were considered in conjunction with analysis of weighted correlation coefficients where higher weights were given to studies that employed higher quality methodologies<sup>(5,34)</sup>. Qualities included data collection methods, administration, seasonality, sample size, supplement use and statistics.

It is estimated that at approximately 7 to 8 years of age children become aware of their own food intake. Prior to this age the cognition and attention span required to perceive time frames, have knowledge of foods, recall food intake, and self-report are not sufficiently developed<sup>(1)</sup>. Other explicit issues that arise in this age group of interest relate to the change in dietary practices seen across the age range and the variability in information provided by parent or proxy reporter, on foods that are eaten outside of their supervision, especially when the child is in day care.

The ability of FFQ to rank nutrient and energy intakes is improved through providing detailed quality information which can be achieved through interviewer administration<sup>(21)</sup>. The majority (71 %) of the included FFQ were self-administered by a parent or proxy reporter, similar to that seen in reviews

conducted in wider age groups<sup>(34,43)</sup>. Cade *et al.*<sup>(7)</sup> reported an increase in correlation coefficients when the FFQ was interviewer-administered, with the exception of vitamin C, in comparison with those that were self-administered. This is especially relevant in the age group in question, where all information is obtained from a parent or proxy-reporter. There is a need for further studies designed to evaluate the accuracy of parental-reported intakes in larger, ethnically diverse populations, using different dietary assessment methods<sup>(44)</sup>.

Estimation of portion size appears to have some advantage over using average or specified portion sizes, with higher measures of agreement between FFQ and reference method ( $r$  0.5–0.6) and higher correlation coefficients when assessing repeatability<sup>(33)</sup>. FFQ are seen to commonly overestimate energy intake, which is especially apparent in this population of interest<sup>(11,15,17,18,21,24)</sup>. This could be attributed to the fact that parents/caregivers may not adequately take into account the small portion sizes consumed by their children and that young children often ‘taste’ many foods without consuming full portions, leading to the inclusion of too large a portion size for some foods<sup>(11,18)</sup>. Many of the included studies assessed wider age ranges, i.e. beyond 12 to 36 months, which, as identified in a recent validation study performed in New Zealand, may act to improve validity of the FFQ as older children are more likely to eat meals that are similar to that of the family member or adult completing the FFQ<sup>(26)</sup>. Improvements in validity and bias could be seen through reducing the number of food items in the FFQ, shortening the reporting period, or adjusting portion sizes to more closely reflect those consumed by a young child<sup>(44)</sup>. This unique method has been explored in a study performed in 12- to 24-month-old New Zealand children where the amount of food offered and the amount eaten were recorded separately to encourage parents to differentiate between the two, and portion sizes were described according to the child’s ‘palm volume’. This FFQ showed acceptable to good validity and high reproducibility in the assessment of dietary patterns and ranking nutrient intakes<sup>(26,31)</sup>.

In a systematic review by Henríquez-Sánchez *et al.*<sup>(43)</sup>, an improvement in correlation coefficients ( $r$  0.52) was seen when the number of food items included in the FFQ was



greater than 100 ( $r$  0.47). The average number of food items used in the present review was 113. Estimation of supplement use should be considered when evaluating nutrient intake. Information on supplements should be included in dietary assessment with emphasis on the type and dose used. Data from FFQ and reference methods correlated better when supplement intake was captured<sup>(43)</sup>. Supplement use was acknowledged in one study<sup>(20)</sup> and seasonality in another<sup>(24)</sup>, but were not considered in the statistical analysis.

All studies calculated Pearson or Spearman's correlation coefficients (Table 1). Calculation of correlation coefficients does not measure agreement between the two methods of dietary assessment, only the degree in which the two methods are related<sup>(45)</sup>. Their usefulness increases if used in conjunction with an alternative method such as Bland–Altman which provides an analysis of how well the FFQ and reference method agree on average<sup>(45)</sup>. Other methods such as limits of agreement can be used to provide information on reliability and the direction and consistency of bias and the magnitude of errors between the two assessment methods<sup>(7,33)</sup>. It is difficult to summarise the correlation coefficients, agreement of validity and reproducibility of the included FFQ; therefore the present review should be used as a description of included FFQ, with potential for further meta-analyses.

Using 24-HR as the dietary reference method, FFQ were found to be a suitable tool for ranking children according to nutrient intakes ( $r$  0.46), with stronger correlations in foods consumed more frequently<sup>(27,37)</sup>. This highlights the difficulties with episodically consumed food items, as seen in the high day-to-day variability of a young child's diet<sup>(18,37)</sup>. Unadjusted FFQ nutrient estimates were larger than unadjusted nutrient estimates from multiple 24-HR and additional analysis of children that regularly received meals and snacks from other caregivers alongside parents revealed no apparent compromise or differences in correlations<sup>(18)</sup>.

Using WFR as the reference method to assess long-term intakes, correlations were found to increase using nutrient density values over absolute intakes, but the FFQ had a low to moderate ability to rank children according to intakes of nutrients and foods<sup>(11)</sup>. WFR are not affected by the same errors, such as portion size estimation, and memory lapses, as the FFQ<sup>(39)</sup>. The FFQ was found to be a useful tool for estimating short-term energy and nutrient intakes in healthy infants (at a group level)<sup>(21,22)</sup>. Marriott *et al.*<sup>(21)</sup> found that differences in micronutrient intakes were partly explained by changes in the consumption of milk between the two dietary assessments and by the different nutrient compositions of cows' milk and formula<sup>(21)</sup>. This underestimation of Ca intake by the FFQ has been reported in three studies within this age group<sup>(19,21,46)</sup>.

The use of FFQ to provide estimates of beverage intake has not been widely investigated. Marshall and Rankin concluded that a quantitative FFQ could be used to provide relative estimates of beverage, Ca, vitamin D and fluoride intakes in this age group<sup>(19,23)</sup> and higher correlations were seen at younger ages when the diet was more limited ( $r$  0.85 at 6 months *v.*  $r$  0.65 at 60 months)<sup>(23)</sup>.

The present review included correlations from three studies using a biomarker for validation<sup>(18,20,28)</sup>. In the assessment of

specific nutritional status, Williams & Innis<sup>(20)</sup> showed that a semi-quantitative FFQ could be a useful tool in assessing Fe status in infants at a group level (energy adjusted  $r$  0.71), but could result in underestimation of infants deemed to be at high risk of poor Fe status<sup>(18,20)</sup>.

### Evaluating quality assessment

Where correlations for a given nutrient were available from three or more studies, quality-adjusted correlations were calculated. Higher weighted mean correlations were seen in studies that used WFR as the reference method for Ca, Fe and fibre when compared with other methods. This may be a reflection of the fact that a greater number of studies (60 %) used WFR as a reference method. The highest correlation coefficient weighted by quality was 0.51. There were not sufficient data to conduct the analysis for the remaining micronutrients, and only six out of the ten EURRECA priority nutrients could be assessed. This continues to remain a concern in this age group, where valid nutrient intake estimates could not be calculated. FFQ validation studies that assessed long-term intakes or used biomarkers as the reference tool were based on one or two studies, making them insufficient to reach any conclusion (Table 3).

### Limitations

There was a lack of data available to assess the ability of the FFQ in providing adequate estimates for several of the micronutrients highlighted in the present review (Table 3). The heterogeneity in the study designs, methods, outcomes and assessment tools made comparisons difficult, therefore the data were narratively synthesised and described. Due to natural variation, biomarkers may not always be a suitable option for comparison<sup>(18)</sup> and few studies validating FFQ using biomarkers were available for inclusion in the present review which would act to reduce correlated errors associated when the reference method is based on self-reporting<sup>(47)</sup>. Studies that assessed the validity of energy intake measurements using doubly labelled water did not meet our inclusion criteria. Due to the specific range of interest, several studies that reported over a wider age range were excluded as reviewers were unable to extract these data. Correlation coefficients of the included studies were used for analysis and quality assessment in the present review; this limits the interpretation of the review as correlation coefficients only measure the degree to which the two assessment methods are related in a validation study, and not the agreement between the methods<sup>(48)</sup>. De-attenuation and energy adjustment have strong implications for correlation coefficients and make it difficult to compare and draw conclusions. Only validation studies written in English were included for analysis. This may have led to the exclusion of reliable validation studies from other countries.

### Conclusion

This systematic literature review presents a summary of the quality of FFQ validation studies in children aged 12 to 36



months. The included studies and quality assessment have provided information on aspects of FFQ design that increase validity, such as the number of items included, portion size estimations, appropriate food choices, administration method, validation and reproducibility methods, pre-testing, supplement use, seasonality and the statistical analyses. Semi-quantitative FFQ were shown to be valid and reproducible when estimating dietary intakes at a group level, and are an acceptable instrument for estimating intakes of Ca, vitamin C and Fe in children 12 to 36 months of age. There is insufficient evidence for the evaluation of the validity of micronutrients such as folate, vitamin D, Zn and Cu in this population. Using the results of the included studies; meticulously designed and validated FFQ may be acceptable in estimating intakes of a number of important micronutrients in this age group.

Children aged 12 to 36 months would benefit from further validation studies using appropriate population-specific tools addressing areas highlighted in this review that are unique to dietary assessment in young children. Such areas include further development on portion size estimation, capturing irregular eating patterns, overcoming administration errors with the implementation of computer-assisted methods or the development of novel tools to provide evidence for further validation studies of appropriate population-specific tools, alongside the identification, management and primary prevention of diet-related disease processes.

### Acknowledgements

There was no financial support.

A. L. and R. B. completed the literature search, screening process and quality assessment. R. B. was the second independent reviewer. A. L. extracted all data, completed the critical appraisal and completed the first draft of the manuscript and contributed to manuscript revision. C. R. W. and C. C. G. helped develop the review protocol and edited the manuscript. All authors approved the submitted version.

The authors would also like to acknowledge Frances Clements, University of Auckland Faculty of Medical and Health Sciences Subject Librarian for her expert assistance in developing the search strategy for this review.

There were no conflicts of interest.

### References

- Livingstone M & Robson P (2000) Measurement of dietary intake in children. *Proc Nutr Soc* **59**, 279–293.
- 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.
- Willett W (2013) *Nutritional Epidemiology: Monographs in Epidemiology and Biostatistics*, 3rd ed. New York: Oxford University Press.
- Number not used.
- Serra-Majem L, Frost Andersen L, Henríque-Sánchez P, *et al.* (2009) Evaluating the quality of dietary intake validation studies. *Br J Nutr* **102**, S3–S9.
- Vereecken CA (2010) A longitudinal study on dietary habits and the primary socialization of these habits in young children. *Verh K Acad Geneesk Belg* **72**, 295–308.
- 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.
- Block G & Hartman AM (1989) Issues in reproducibility and validity of dietary studies. *Am J Clin Nutr* **50**, 1133–1138; discussion 1231–1235.
- Coulston AM & Boushey C (2008) *Nutrition in the Prevention and Treatment of Disease*. Amsterdam: Academic Press.
- Serdula MK, Alexander MP, Scanlon KS, *et al.* (2001) What are preschool children eating? A review of dietary assessment 1. *Annu Rev Nutr* **21**, 475–498.
- Andersen LF, Lande B, Arsky GH, *et al.* (2003) Validation of a semi-quantitative food-frequency questionnaire used among 12-month-old Norwegian infants. *Eur J Clin Nutr* **57**, 881–888.
- Ortiz-Andrellucchi A, Henríquez-Sánchez P, Sánchez-Villegas A, *et al.* (2009) Dietary assessment methods for micronutrient intake in infants, children and adolescents: a systematic review. *Br J Nutr* **102**, S87–S117.
- Livingstone M, Robson P & Wallace J (2004) Issues in dietary intake assessment of children and adolescents. *Br J Nutr* **92**, S213–S222.
- University of York, Centre for Reviews and Dissemination (2015) International Prospective Register of Systematic Reviews. <http://www.crd.york.ac.uk/PROSPERO/> (accessed May 2016).
- Andersen L, Lande B, Trygg K, *et al.* (2004) Validation of a semi-quantitative food-frequency questionnaire used among 2-year-old Norwegian children. *Public Health Nutr* **7**, 757–764.
- Iannotti RJ, Zuckerman AE, Blyer EM, *et al.* (1994) Comparison of dietary intake methods with young children. *Psychol Rep* **74**, 883–889.
- Blum RE, Wei EK, Rockett HR, *et al.* (1999) Validation of a food frequency questionnaire in Native American and Caucasian children 1 to 5 years of age. *Matern Child Health J* **3**, 167–172.
- Parrish LA, Marshall JA, Krebs NF, *et al.* (2003) Validation of a food frequency questionnaire in preschool children. *Epidemiology* **14**, 213–217.
- Marshall TA, Gilmore JME, Broffitt B, *et al.* (2003) Relative validation of a beverage frequency questionnaire in children ages 6 months through 5 years using 3-day food and beverage diaries. *J Am Diet Assoc* **103**, 714–720.
- Williams PL & Innis SM (2005) Food frequency questionnaire for assessing infant iron nutrition. *Can J Diet Pract Res* **66**, 176–182.
- Marriott LD, Inskip HM, Borland SE, *et al.* (2009) What do babies eat? Evaluation of a food frequency questionnaire to assess the diets of infants aged 12 months. *Public Health Nutr* **12**, 967–972.
- 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.
- Rankin SJ, Levy SM, Warren JJ, *et al.* (2011) Relative validity of an FFQ for assessing dietary fluoride intakes of infants and young children living in Iowa. *Public Health Nutr* **14**, 1229–1236.
- D'Ambrosio A, Tiessen A & Simpson JR (2012) Development of a food frequency questionnaire for toddlers of Low-German-speaking Mennonites from Mexico. *Can J Diet Pract Res Spring* **73**, 40–44.
- Number not used.
- Watson EO, Heath AM, Taylor RW, *et al.* (2015) Relative validity and reproducibility of an FFQ to determine nutrient intakes of New Zealand toddlers aged 12–24 months. *Public Health Nutr* **18**, 3265–3271.
- Sochacka-Tatara E & Pac A (2014) Relative validity of a semi-quantitative FFQ in 3-year-old Polish children. *Public Health Nutr* **17**, 1738–1744.
- Orton HD, Szabo NJ, Clare-Salzler M, *et al.* (2008) Comparison between omega-3 and omega-6 polyunsaturated fatty acid intakes as assessed by a food frequency questionnaire and erythrocyte membrane fatty acid composition in young children. *Eur J Clin Nutr* **62**, 733–738.
- Klohe DM, Clarke KK, George GC, *et al.* (2005) Relative validity and reliability of a food frequency questionnaire for a triethnic population of 1-year-old to 3-year-old children from low-income families. *J Am Diet Assoc* **105**, 727–734.
- Bel-Serrat S, Fernandez Alvira JM, Pala V, *et al.* (2011) Relative validation of two dietary assessment methods: SACINA (24-h recall) and food frequency questionnaire. *Int J Obes* **35**, S152.



31. Mills VC, Skidmore PM, Watson EO, *et al.* (2015) Relative validity and reproducibility of a food frequency questionnaire for identifying the dietary patterns of toddlers in New Zealand. *J Acad Nutr Diet* **115**, 551–558.
32. Dennis LK, Snetselaar LG, Nothwehr FK, *et al.* (2003) Developing a scoring method for evaluating dietary methodology in reviews of epidemiologic studies. *J Am Diet Assoc* **103**, 483–487.
33. 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.
34. Roman-Viñas B, Ortiz-Andrellucchi A, Mendez M, *et al.* (2010) Is the food frequency questionnaire suitable to assess micronutrient intake adequacy for infants, children and adolescents? *Matern Child Nutr* **6**, 112–121.
35. National Health and Medical Research Council (NHMRC). (2000) *How to Use the Evidence: Assessment and Application of Scientific Evidence*. Canberra: Biotex.
36. Number not used.
37. 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.
38. Metcalf PA, Scragg RKR, Sharpe S, *et al.* (2003) Short-term repeatability of a food frequency questionnaire in New Zealand children aged 1–14 y. *Eur J Clin Nutr* **57**, 1498–1503.
39. Gibson RS (2005) *Principles of Nutritional Assessment*, 2nd ed. New York: Oxford University Press.
40. Treadwell JR, Tregear SJ, Reston JT, *et al.* (2006) A system for rating the stability and strength of medical evidence. *BMC Med Res Methodol* **6**, 52.
41. Subar AF (2004) Developing dietary assessment tools. *J Am Diet Assoc* **104**, 769–770.
42. Schatzkin A, Kipnis V, Carroll RJ, *et al.* (2003) A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. *Int J Epidemiol* **32**, 1054–1062.
43. Henríquez-Sánchez P, Sánchez-Villegas A, Doreste-Alonso J, *et al.* (2009) Dietary assessment methods for micronutrient intake: a systematic review on vitamins. *Br J Nutr* **102**, S10–S37.
44. Collins CE, Burrows TL, Truby H, *et al.* (2013) Comparison of energy intake in toddlers assessed by food frequency questionnaire and total energy expenditure measured by the doubly labeled water method. *J Acad Nutr Diet* **113**, 459–463.
45. Bland JM & Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* **8**, 135–160.
46. 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.
47. Day NE, Wong MY, Bingham S, *et al.* (2004) Correlated measurement error – implications for nutritional epidemiology. *Int J Epidemiol* **33**, 1373–1381.
48. Altman DG (1990) *Practical Statistics for Medical Research*. Boca Raton, FL: CRC Press.