

## A short food-group-based dietary questionnaire is reliable and valid for assessing toddlers' dietary risk in relatively advantaged samples

Lucinda K. Bell<sup>1\*</sup>, Rebecca K. Golley<sup>2</sup> and Anthea M. Magarey<sup>1</sup>

<sup>1</sup>Nutrition and Dietetics, Faculty of Medicine, Nursing and Health Sciences, School of Medicine, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia

<sup>2</sup>Division of Health Sciences, Sansom Institute for Health Research, University of South Australia, Adelaide, SA, Australia

(Submitted 6 January 2014 – Final revision received 29 April 2014 – Accepted 2 May 2014 – First published online 2 June 2014)

### Abstract

Identifying toddlers at dietary risk is crucial for determining who requires intervention to improve dietary patterns and reduce health consequences. The objectives of the present study were to develop a simple tool that assesses toddlers' dietary risk and investigate its reliability and validity. The nineteen-item Toddler Dietary Questionnaire (TDQ) is informed by dietary patterns observed in Australian children aged 14 (*n* 552) and 24 (*n* 493) months and the Australian dietary guidelines. It assesses the intake of 'core' food groups (e.g. fruit, vegetables and dairy products) and 'non-core' food groups (e.g. high-fat, high-sugar and/or high-salt foods and sweetened beverages) over the previous 7 d, which is then scored against a dietary risk criterion (0–100; higher score = higher risk). Parents of toddlers aged 12–36 months (Socio-Economic Index for Areas decile range 5–9) were asked to complete the TDQ for their child (*n* 111) on two occasions, 3·2 (SD 1·8) weeks apart, to assess test–retest reliability. They were also asked to complete a validated FFQ from which the risk score was calculated and compared with the TDQ-derived risk score (relative validity). Mean scores were highly correlated and not significantly different for reliability (intra-class correlation = 0·90, TDQ1 30·2 (SD 8·6) *v.* TDQ2 30·9 (SD 8·9); *P* = 0·14) and validity (*r* 0·83, average TDQ ((TDQ1 + TDQ2)/2) 30·5 (SD 8·4) *v.* FFQ 31·4 (SD 8·1); *P* = 0·05). All the participants were classified into the same (reliability 75%; validity 79%) or adjacent (reliability 25%; validity 21%) risk category (low (0–24), moderate (25–49), high (50–74) and very high (75–100)). Overall, the TDQ is a valid and reliable screening tool for identifying at-risk toddlers in relatively advantaged samples.

**Key words:** Toddlers; Dietary risk; Questionnaires; Validity; Reliability

'Dietary risk' is a term used to describe 'any inappropriate dietary pattern' that may impair health<sup>(1)</sup>. Toddlers are vulnerable to dietary risk as they begin to exert their independence in food choices and demonstrate fussy eating behaviours<sup>(2,3)</sup>. As dietary risk habits may persist over time<sup>(4,5)</sup> and influence short-term and long-term health<sup>(6,7)</sup>, early risk identification is important.

The current dietary intakes of toddlers are inadequate, suggesting that many are at dietary risk. In general, intakes of nutrient-rich foods are below the national dietary guideline recommendations and consumption of energy-dense, nutrient-poor foods is common. For example, the 2008/09 UK National Diet and Nutrition Survey revealed that about 50% of 1·5- to 3-year-olds consumed energy-dense, nutrient-poor items such as meat products, fried potato products, confectionery and sweetened beverages over the 4 d food diary period<sup>(8)</sup>. Nutrient-rich foods such as fish, raw vegetables and eggs were consumed by less than half the sample<sup>(8)</sup>. Similarly, a recent Australian study

demonstrated that 11–15% of 12- to 36-month-olds consumed no fruit or vegetables, respectively, less than one-quarter consumed eggs (24%), fish (11%) and legumes (17%), and nearly all (89%) consumed energy-dense, nutrient-poor item/s in the previous 24 h<sup>(9)</sup>. Similar trends are observed in other countries including the USA<sup>(10,11)</sup>. These data highlight that toddlers' dietary patterns are not consistent with dietary guidelines and may place them at risk of nutrient (e.g. Fe and folate<sup>(12,13)</sup>) deficiencies and chronic diseases, including excess weight<sup>(4,14)</sup> and CVD<sup>(15)</sup>. Therefore, the need to screen toddlers' dietary intakes against current dietary guidelines to identify those at risk is evident.

Timely, accurate and cost-effective assessment of dietary intake is important. Traditional dietary assessment methods, such as recalls and records, are time intensive, costly and burdensome<sup>(16)</sup>. Furthermore, it can be difficult to easily extract food intake data using these methods for meaningful comparison with food-group-based dietary guidelines<sup>(17)</sup>. Conversely,

**Abbreviations:** ICC, intra-class correlation; SAIDI, South Australian Infant Dietary Intake; SEIFA, Socio-Economic Index for Areas; TDQ, Toddler Dietary Questionnaire; TDQ<sub>ave</sub>, average Toddler Dietary Questionnaire.

\* **Corresponding author:** L. K. Bell, fax +61 8 8204 6406, email lucy.bell@flinders.edu.au

less costly, time-consuming and laborious methods such as FFQ<sup>(18)</sup> quickly measure food or food-group intakes, allowing easy comparison with food-group-based dietary guidelines<sup>(17)</sup>. Nonetheless, increasing questionnaire length is associated with increasing burden, likely to result in reduced cooperation and completion<sup>(19)</sup>. Therefore, an ideal screening questionnaire that identifies toddlers at dietary risk would be short and simple while providing food or food-group data that can easily be compared with dietary guidelines.

Dietary risk identification requires the assessment of whole diets. In comparison with that of individual dietary components, the assessment of whole diets refers to capturing the intake data of all five 'core' food groups (i.e. foods recommended to be consumed every day including fruit, vegetables, grains (e.g. bread, rice, pasta and noodles), meat and alternatives (e.g. fish, eggs and nuts, and dairy products) and 'non-core' food groups (energy-dense, nutrient-poor items)<sup>(20,21)</sup>. However, current short food or food-group-based questionnaires generally aim to measure a specific aspect of diet (e.g. fat intake<sup>(22)</sup>) or a limited number of food groups (e.g. only fruit and vegetables<sup>(23)</sup>). Supporting this, our recent review<sup>(24)</sup> highlighted the lack of short tools ( $\leq 50$  items) assessing the whole diets of children aged  $< 5$  years<sup>(24)</sup>.

Due to the lack of population-specific, age-appropriate, short tools that characterise whole diets of Australian toddlers, the present study aimed to develop a short, simple food-group-based dietary risk assessment tool for toddlers aged 12–36 months and determine its reliability and validity.

## Methods

The Toddler Dietary Questionnaire (TDQ) is a nineteen-item, parent-completed, semi-quantitative tool that assesses food-group intake over the previous 7 d. The intake of 'core' food groups (e.g. fruit, vegetables and dairy products) and 'non-core' food groups (e.g. high-fat, high-sugar and/or high-salt foods and sweetened beverages) is then evaluated against a dietary risk criterion. The TDQ risk scores range from 0 to 100, with a higher score representing a higher dietary risk (i.e. poorer dietary intake).

### Development of the Toddler Dietary Questionnaire

The development of the TDQ was informed by dietary patterns observed in the recent dietary intake data of Australian toddlers<sup>(25)</sup>, the Australian Dietary Guidelines Modelling System<sup>(26)</sup> and the Australian Dietary Guidelines<sup>(27,28)</sup>. Questionnaire drafts were pilot tested for readability, understanding and timing with three parent–toddler dyads (university researchers,  $n = 2$ , and a family member of the researchers,  $n = 1$ ) and changes made to the questionnaire format.

The TDQ items were primarily informed by the dietary patterns of Australian children<sup>(25)</sup>, derived using principal component analysis. Principal component analysis is a common type of factor analysis<sup>(29)</sup> that identifies the underlying 'patterns' of intake from a large number of variables by

grouping foods commonly consumed together. Principal component analysis was applied to the average of 24 h recall data collected over 3 d from 14-month-old ( $n = 552$ ) and 24-month-old ( $n = 493$ ) children. Data were derived from two Australian studies, the control arm of NOURISH<sup>(30)</sup>, an obesity prevention randomised controlled trial, and the South Australian Infant Dietary Intake (SAIDI) study, a longitudinal study of infants' and toddlers' dietary intake. The foods that represent extracted patterns account for the greatest variation in diet between individuals<sup>(31)</sup>. At both ages, two patterns were identified representing (1) 'core' intake (e.g. fruit, vegetables, grains, dairy products, meat and water) and (2) 'non-core' intake (e.g. high-fat, high-sugar and/or high-salt products and sweetened beverages)<sup>(25)</sup>. Based on these patterns and the Australian Dietary Guidelines<sup>(26,27)</sup>, a nineteen-item questionnaire comprising three sections was developed. Section 1 assesses 'core' intake (eight items: fruit, vegetables (green, orange and other), dairy products, grains, lean red meat and fish), section 2 'non-core' intake (eight items: spreadable fats, vegemite-type spreads, snack products, hot potato products, meat products, sweet biscuits and cakes, chocolates and ice creams) and section 3 'usual' intake (three items: bread type, milk beverages and non-milk beverages, e.g. fruit juice, soft drink, and cordial (a fruit-flavoured concentrate that is usually mixed with water)).

Sections 1 and 2 comprise questions asking the respondents to report how often and how much their child ate of each food group over the previous week. Based on the appropriateness of categories for a 1-week period of intake, four consumption frequency categories (nil, once, 2–4 times and  $\geq 5$  times) were developed. In addition, three consumption quantity categories (representing 'small' (e.g.  $< 50$  g), 'medium' (e.g. 50–100 g) and 'large' (e.g.  $> 100$  g) portions) were developed. For section 1, portion-size categories were informed by the average serving sizes and weekly number of servings recommended for 13- to 23-month-olds and 2- to 3-year-olds outlined in the Australian Dietary Guidelines Modelling System<sup>(26)</sup>. For TDQ food groups not directly comparable to those in the modelling system, a proportion of the recommended intake was used. For example, for the TDQ food group 'yogurt/custard', portion sizes were informed by applying 25% to the recommended intake of 'dairy foods (milks, yogurts and cheese)'. For section 2, portion-size categories were informed by the tertiles of consumption of 24-month-old NOURISH and SAIDI children ( $n = 742$ ). Food labels that reflect each portion-size category ('small', 'medium' and 'large') were added for each food-group item. For example, a 'small' portion of 'other vegetables' was labelled 'less than one cup of raw salad vegetables or less than half a cup of cooked vegetables', representing  $< 75$  g of vegetables.

Section 3 comprises the following three questions: (1) What proportion of white:non-white bread (e.g. some white:mostly non-white) does your child usually consume? (2) What milk drinks (breast, plain, flavoured or formula) does your child usually consume? (3) What non-milk drinks (water, diluted juice, juice, or cordial/soft drink) does your child usually consume? The final questionnaire is given in Table S1 (available online).



### Scoring of the Toddler Dietary Questionnaire

The dietary risk score is derived by evaluating food-group intake against a scoring criterion (Table 1). For sections 1 and 2 of the TDQ, food-group intake per week in grams is calculated by multiplying the frequency response (zero (nil), one (once), three (2–4) and seven ( $\geq 5$ ) times per week) with the median quantity response (e.g. small =  $< 50$ , 25 g; medium = 50–100, 75 g). For example, if the median of the ‘small category’ is 25 g, then a response of ‘2–4 times’ and ‘small’ amount is 75 g ( $3 \times 25$  g). As the median of the ‘large’ (e.g.  $> 100$  g) category could not be established based on the TDQ categories, an upper limit of consumption of 24-month-old NOURISH and SAIDI children was used (e.g. 300 g) and the median determined (e.g. 200 g). Intake is then compared against recommendations<sup>(26)</sup>. That is, a scale of 0 (lowest score = lowest risk) to 18 (highest score = highest risk) is applied per question, with ‘0’ reflecting intake closest in line with the recommendations and ‘18’ reflecting intake furthest from the recommendations (Table 1). For section 1, a response of ‘2–4 times’ and ‘medium’ amount reflects intake most closely in line with the recommendations and is therefore scored a ‘0’. Lower and higher intakes are scored between 2 and 18 according to the percentage of deviation from the recommendations. Underconsumption is scored slightly more severely than overconsumption due to greater severity of health risks. For example, underconsumption of ‘core’ foods may result in nutrient deficiencies leading to suboptimal growth and development and/or chronic diseases such as CVD and cancer<sup>(20)</sup>. Furthermore, insufficient ‘core’ intake may lead to the overconsumption of ‘non-core’ items and thus an increased risk of overweight and obesity<sup>(20)</sup>. Alternatively, overconsumption of ‘core’ foods may also contribute to overweight through the establishment of a positive energy balance<sup>(32)</sup> and may displace the intake of other core foods from the diet, thus decreasing variety<sup>(20)</sup>. Conversely, for section 2, scores increase proportionally from 0 with increasing consumption frequency and quantity, as the consumption of ‘non-core’ foods should be limited<sup>(20,21,27)</sup> and increasing exposure and familiarity increase the preference for these foods<sup>(33)</sup>.

Each question in section 3 is scored on a scale of 0 (ideal intake, e.g. none white:all non-white, breast milk or plain milk, and water) to 12 (non-ideal intake, e.g. all white:none non-white, no milk drinks, and soft drink or cordial) (Table 1). For questions 2 and 3, a proportionally increasing scale of 0, 4, 8, 12, is applied, with multiple responses being accepted. However, for question 1, a scale of 0, 3, 9, 12 is applied, as the proportions 25%:75% and 75%:25% were used to represent the responses some white:mostly non-white and mostly white:some non-white, respectively.

Dietary risk scores are created for each section, tallied to give a score out of 336, which is converted to a total dietary risk score (range 0–100; higher score = higher risk). Total risk scores are categorised into four levels of dietary risk: (1) low (0–24); (2) moderate (25–49); (3) high (50–74); (4) very high (75–100).

### Reliability and validity of the Toddler Dietary Questionnaire

#### Study design

A validation study was conducted between October 2012 and February 2013 to determine the reliability and relative validity of the TDQ. Ethics approval was granted by the Flinders University Social and Behavioural Research Ethics Committee (SBREC).

#### Study sample

The participants were primary carers of toddlers aged 12–36 months recruited via (1) flyers distributed at South Australian private child care centres, (2) advertisements in Flinders University newsletters and on notice boards, (3) a study-specific Facebook page, and (4) parents enrolled in the SAIDI study who had another eligible child. Children with a food allergy or intolerance or a diagnosed medical condition affecting their dietary intake were excluded. Parents with two eligible children chose one child to participate in the study to prevent a clustering effect. Parental consent was obtained.

#### Data collection

Data collection occurred in two stages. In stage 1, the participants completed a demographic questionnaire and the TDQ (i.e. TDQ1). In stage 2, the participants were mailed a second TDQ (i.e. TDQ2) and a validated semi-quantitative FFQ<sup>(34,35)</sup> to be completed on the same day approximately 2–4 weeks after the completion of TDQ1.

**Demographic questionnaire.** Child (age, sex, country of birth, and parent-reported weight and height), parent (age, country of birth, marital status, education level and employment status) and family (postal code and household numbers) demographic characteristics were assessed via a questionnaire. As a measure of socio-economic status, the Index of Relative Socio-Economic Advantage and Disadvantage, one of the four Socio-Economic Index for Areas (SEIFA) indices that rank geographical areas across Australia on a continuum of disadvantage (lowest score = 1) to advantage (highest score = 10), was applied to the postal code<sup>(36)</sup>.

**FFQ.** To determine the validity of the TDQ, a dietary assessment tool that allowed collected data to be translated into the TDQ and dietary risk calculated was necessary. A recently developed seventeen-item FFQ for Australian 2- to 5-year olds<sup>(37)</sup> was not suitable as the validation tool due to the lack of assessment of dairy product and grain food intakes, preventing the calculation of a dietary risk score. Furthermore, alternative measures, such as 24 h recalls and 2 or 3 d records, do not provide data collected over sufficient number of days to cover that of the TDQ, while 7 d records are associated with high participant burden<sup>(16,38)</sup>. Therefore, a FFQ developed and validated in Belgian 2.5- to 6.5-year-olds<sup>(34,35)</sup> was chosen as the validation reference tool. This FFQ was identified in a recent review as the only short dietary assessment tool for children aged 0–5 years tested for reliability and validity<sup>(24)</sup> from which a TDQ score could be calculated. Food-group items are mostly compatible with those in the TDQ and the

**Table 1.** Scoring template for the Toddler Dietary Questionnaire

Sections	Question	Response	Score		Maximum score per question	Possible section score range	
			Section 1	Section 2			
1 and 2	Each question (sixteen items) scored according to the combination of frequency and quantity categorical responses	Frequency	Quantity	Section 1	Section 2	18	0–144
		Nil	Nil	18	0		
		Once	Small	14	2		
		Once	Medium	11	4		
		Once	Large	8	6		
		2–4 times	Small	6	8		
		2–4 times	Medium	0	10		
		2–4 times	Large	4	12		
		≥ 5 times	Small	2	14		
		≥ 5 times	Medium	6	16		
3	What proportion of white:non-white bread does your child usually* consume? (tick one only)	None white:all non-white		0		12	0–48
		Some white:mostly non-white		3			
		Mostly white:some non-white		9			
		All white:none non-white		12			
	What milk drinks does your child usually* consume? (tick all that apply)	Breast milk or plain milk (dairy or non-dairy)		0		12†	
		Formula		4			
		Flavoured milk (dairy or non-dairy)		8			
		None of the above, i.e. no milk drinks		12			
	What non-milk drinks does your child usually* consume? (tick all that apply)	Water		0		24	
		Diluted juice (fruit and/or vegetable)		4			
Undiluted juice (fruit and/or vegetable)			8				
Cordial or soft drink			12				
Total						0–336 (converted to out of 100)	

\* Usually = on most days.

† Despite the option to tick all that apply, if a response of 'none of the above, i.e. no milk drinks', is provided, no other responses are possible; therefore, any combination of the first three responses, providing a maximum score of 12, or the response 'none of the above, i.e. no milk drinks' only (score = 12) is possible.

**Table 2.** Section and total dietary risk scores for each administration of the Toddler Dietary Questionnaire (TDQ; TDQ1 and TDQ2), average TDQ (TDQ<sub>ave</sub>) and FFQ and classification into dietary risk categories (*n* 111) (Mean values and standard deviations; number of participants and percentages)

Dietary risk measures	Possible score range	Test–retest reliability				Relative validity			
		TDQ1		TDQ2		TDQ <sub>ave</sub> *		FFQ	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
<b>Dietary risk score</b>									
Section 1	0–144	56.0	18.0	56.7	19.5	56.3	17.7	61.0	18.1
Section 2	0–144	39.3	18.6	40.1	19.1	39.7	17.8	39.3	19.3
Section 3	0–48	6.2	6.4	7.1	7.3	6.6	6.6	5.3	5.9
Total	0–100	30.2	8.6	30.9	8.9	30.5	8.4	31.4	8.1
<b>Dietary risk score category</b>									
Low	0.0–24.9								
<i>n</i>		31		33		33		21	
%		27.9		29.7		29.7		18.9	
Moderate	25.0–49.9								
<i>n</i>		76		74		76		87	
%		68.5		66.7		68.5		78.4	
High	50.0–74.9								
<i>n</i>		4		4		2		3	
%		3.6		3.6		1.8		2.7	
Very high	75.0–99.9								
<i>n</i>		0		0		0		0	
%		0.0		0.0		0.0		0.0	

\*TDQ<sub>ave</sub> = ((TDQ1 risk scores + TDQ2 risk scores)/2).

1-month assessment period of the FFQ covers the 1-week assessment period of the TDQ. Small adaptations were made to the FFQ to reflect culturally appropriate foods and terminology (e.g. sugared milk replaced with flavoured milk) and to capture intake over the past month rather than that over the past year.

Comparative validity was assessed to evaluate dietary risk scores determined using the nineteen-item TDQ relative to those determined using the fifty-four-item (forty-seven-food item) FFQ. The final FFQ included six frequency categories (never, 1–3 d/month, 1 d/week, 2–4 d/week, 5–6 d/week and every day) and three quantity categories (representing ‘small’ (e.g. <40 g), ‘medium’ (e.g. 40–120 g) and ‘large’ (e.g. ≥120 g) portions). FFQ data were converted to a third dietary risk score using a standardised format based on comparative quantity and frequency categories, and the risk score was calculated. That is, responses ‘never’ and ‘1–3 d/month’ were translated to ‘nil’ in the TDQ, ‘1 d/week’ to ‘once’, ‘2–4 d/week’ to ‘2–4 times’, and ‘5–6 d/week’ and ‘everyday’ to ‘≥ 5 times’. Quantity responses were translated to the most appropriate TDQ quantity category (‘small’, ‘medium’ or ‘large’) based on gram amount.

### Statistical analysis

Data were analysed using SPSS statistical software package for Windows version 19.0 (SPSS, Inc.). The level of significance was set at *P* < 0.05.

### Individual Toddler Dietary Questionnaire item agreement

The proportion of parents reporting within the same response category (product of frequency and quantity; data not shown)

between each administration of the TDQ (TDQ1 and TDQ2) was determined and the percentage of agreement calculated. The percentage of agreement between the administrations beyond that expected by chance<sup>(39)</sup> was determined by calculating weighted kappa (*K<sub>w</sub>*) (for ordinal data) using MedCalc statistical software version 12.7.7.0 (Microsoft). *K<sub>w</sub>* values were defined as poor (<0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80) and very good (0.81–1.00)<sup>(40,41)</sup>.

### Reliability and validity of dietary risk scores

Risk scores were evaluated for test–retest reliability and relative validity of section and total scores. Reliability was assessed by comparing scores obtained during the first administration (TDQ1) and second administration (TDQ2) of the TDQ and relative validity by evaluating average scores (termed ‘TDQ<sub>ave</sub>’) derived from two administrations of the TDQ ((TDQ1 + TDQ2)/2) against FFQ risk scores. Average risk scores were used in the validity analysis instead of the TDQ1 or TDQ2 scores as these cover a 2-week period of intake, more in line with the 4-week assessment period of the FFQ, and are thus a better representation of ‘usual’ intake and risk. As the majority of responses were normally distributed, parametric tests were used in all analyses for consistency.

To assess reliability and relative validity at the individual level, intra-class correlations (ICC) and Pearson’s correlations, defined as low ≤0.50, moderate 0.51–0.69, and high ≥0.70<sup>(42)</sup>, were used. At the group level, paired *t* tests were used for both analyses. A Bland–Altman plot was constructed to assess the strength of agreement between the two tools by plotting the mean bias, i.e. difference between the TDQ<sub>ave</sub> and FFQ risk scores, against the mean of the tools. The plot was assessed visually and linear regression analysis performed to



**Table 3.** Agreement of Toddler Dietary Questionnaire (TDQ) items (product of frequency and quantity categories, categorical) between each administration among Australian children aged 12–36 months (*n* 111)

Sections	TDQ items	TDQ1 and TDQ2	
		Percentage of agreement*	$K_w$ †
1	Fruit	64	0.48
	Green vegetables	51	0.52
	Orange vegetables	48	0.51
	Other vegetables	52	0.50
	Yogurt or custard	54	0.61
	Grains	40	0.40
	Red meat	55	0.46
	Fish	57	0.55
	2	Spreadable fats	51
Vegemite-type spreads		32	0.51
Snack products		56	0.46
Hot potato products		48	0.53
Meat products		42	0.51
Sweet biscuits or cakes		41	0.46
Chocolates		65	0.60
Ice creams or frozen yogurt		56	0.52
3		Bread type	80
	Milk drinks	89	0.67
	Non-milk drinks	85	0.74

$K_w$ , weighted kappa.

\*Percentage within the same category response, i.e. combination of frequency and quantity categories for sections 1 and 2 (*n* 10). For question 1 on bread type in section 3, one response was allowed and five response options were provided: (1) none white:all non-white; (2) some white:mostly non-white; (3) mostly white:some non-white; (4) all white:none non-white; (5) does not eat bread. For question 2 on milk drinks in section 3, multiple responses were allowed and eight response options were provided: (1) breast milk or plain milk only; (2) formula only; (3) flavoured milk only; (4) no milk drinks; (5) breast milk/plain milk and formula; (6) breast milk/plain milk and flavoured milk; (7) breast milk/plain milk and formula milk and flavoured milk; (8) formula and flavoured milk. For question 3 on non-milk drinks in section 3, multiple responses were allowed and eight response options were provided: (1) water; (2) diluted juice only; (3) water and diluted juice; (4) water and undiluted juice; (5) water and cordial/soft drink; (6) water and diluted juice and cordial/soft drink; (7) water and diluted juice and undiluted juice; (8) water and undiluted juice and cordial/soft drink.

†  $K_w$  was calculated for categorised data as described above.

test for any systematic bias. Agreement at the individual level is defined as the limits of agreement ( $\pm 2$ SD) of the mean bias and that at the group level by the mean bias and slope of the mean bias line<sup>(17)</sup>.

### Cross-classification into dietary risk categories (low–very high)

Classification analysis was conducted to determine whether the participants were classified into the same dietary risk category (low, moderate, high and very high) during each TDQ administration and by TDQ<sub>ave</sub> scores compared with the FFQ scores.

## Results

### Sample characteristics

Of the 117 parents (100% biological mother), 111 completed all the study measures. Mothers (mean age 34 (SD 4) years) were mostly partnered (94%), Australian born (95%), in paid employment (74%) with a university education (67%),

and in the top five SEIFA deciles (range 5–9). Children (54% girls) were, on average, 23.0 (SD 6.9) months of age, primarily Australian born (95%), and lived in a household of 4 (SD 1) members.

### Reliability and validity

The duration between the repeat administrations of the TDQ ranged from 1.0 to 11.9 weeks (average 3.2 (SD 1.8) weeks). The average dietary risk scores ranged from 30.2 (SD 8.6) for TDQ1 to 31.4 (SD 8.1) for the TDQ derived from the FFQ (Table 2). Over two-thirds of children were classified as moderate risk and less than one-third as low risk (Table 2).

**Test–retest reliability.** The percentage of agreement and  $K_w$  for each TDQ item are summarised in Table 3. The percentage of agreement ranged from 32% for vegemite-type spreads to 85% for non-milk drinks.  $K_w$  values ranged from 0.40 to 0.78, indicating fair (grains), moderate (fruit, vegetables (orange, green and other), red meat, fish, vegemite-type spreads, snack products, hot potato products, meat products, sweet biscuits and cakes, chocolates and ice creams) and good (yogurt, spreadable fats, bread, milk drinks and non-milk drinks) agreement.

The results of the test–retest analysis of dietary risk scores are given in Table 4. The total risk scores calculated from each TDQ administration were highly correlated (ICC = 0.90,  $P < 0.001$ ) and not statistically different (30.2 (SD 8.6) *v.* 30.9 (SD 8.9);  $P = 0.14$ ). For section risk scores, all ICC were good (0.88–0.91). Risk scores for section 3 (6.2 (SD 6.4) *v.* 7.1 (SD 7.3);  $P = 0.017$ ), but not for section 1 ( $P = 0.55$ ) or section 2 ( $P = 0.45$ ), were significantly different between each administration. Mean bias ranged from  $-0.88$  for section 3 to  $-0.71$  for section 1 (TDQ1 scores were lower than the TDQ2 scores). All children were classified into the same (*n* 83, 75%) or adjacent (*n* 28, 25%) dietary risk category during each administration (Table 5).

**Relative validity.** The total and section dietary risk scores derived from the TDQ<sub>ave</sub> and those derived from the FFQ were highly correlated (all  $r$  0.71 or greater,  $P < 0.001$ ; Table 4). Risk scores were significantly different for section 1 (TDQ<sub>ave</sub> 56.3 (SD 17.7), FFQ 61.0 (SD 18.1);  $P < 0.001$ ) and section 3 (TDQ<sub>ave</sub> 6.6 (SD 6.6), FFQ 5.3 (SD 5.9);  $P = 0.005$ ), but not for section 2 ( $P = 0.69$ ), and total risk scores were not significantly different (TDQ<sub>ave</sub> 30.5 (SD 8.4), FFQ 31.4 (SD 8.1);  $P = 0.05$ ). Mean bias between the TDQ<sub>ave</sub> and FFQ risk scores ranged from  $-4.68$  (section 1; TDQ<sub>ave</sub> scores were lower than the FFQ scores) to 1.31 (section 3; TDQ<sub>ave</sub> scores were greater than the FFQ scores).

The Bland–Altman plot (Fig. 1) revealed a small negative mean difference between the TDQ<sub>ave</sub> and FFQ risk scores; i.e. the TDQ<sub>ave</sub> tends to provide a lower estimate of risk than the FFQ (mean bias  $-0.89$  ( $-1.79$ ,  $0.02$ )). However, most measurements fell within the 95% limits of agreement and there was no significant linear trend for the fitted regression line ( $\beta = 0.51$ , 95% CI  $-0.08$ ,  $0.15$ ;  $P = 0.60$ ), i.e. no systematic bias between the two tools. Classification analysis between the TDQ<sub>ave</sub> and FFQ revealed that all the participants were classified into the same (*n* 88, 79%) or adjacent (*n* 23, 21%) dietary risk category (Table 5).



**Table 4.** Test–retest reliability of the Toddler Dietary Questionnaire (TDQ) risk scores and relative validity of the average TDQ (TDQ<sub>ave</sub>) and FFQ risk scores for each section and total risk scores (*n* 111) (Correlations, 95 % confidence intervals and 95 % limits of agreement (LOA))

Dietary risk scores	Test–retest reliability (TDQ1 and TDQ2)				Relative validity (TDQ <sub>ave</sub> * and FFQ)				
	ICC†	Mean bias	95 % CI	P‡	Pearson's correlation†	Mean bias	95 % CI	P‡	Slope of the mean bias line§
Section 1	0.88	-0.71	-3.06, 1.64	0.550	0.71	-4.68	-7.24, -2.12	<0.001	-0.03
Section 2	0.89	-0.86	-3.15, 1.42	0.454	0.84	0.40	-1.57, 2.37	0.691	-0.08
Section 3	0.91	-0.88	-1.61, -0.16	0.017	0.71	1.31	0.40, 2.22	0.005	0.12
Total	0.90	-0.73	-1.72, 0.25	0.143	0.83	-0.89	-1.79, 0.02	0.054	0.03

ICC, intra-class correlation.

\*TDQ<sub>ave</sub> = (TDQ1 risk scores + TDQ2 risk scores)/2.

†For all correlations, *P* < 0.001.

‡Paired *t* test was used to compare differences in risk scores.

§Linear regression analysis of difference in risk scores (TDQ<sub>ave</sub> - FFQ) and the mean of difference of risk scores ((TDQ<sub>ave</sub> - FFQ)/2). Agreement at the individual level is defined as the LOA (±2 sd) of the mean bias and that at the group level by the mean bias and slope of the mean bias line.

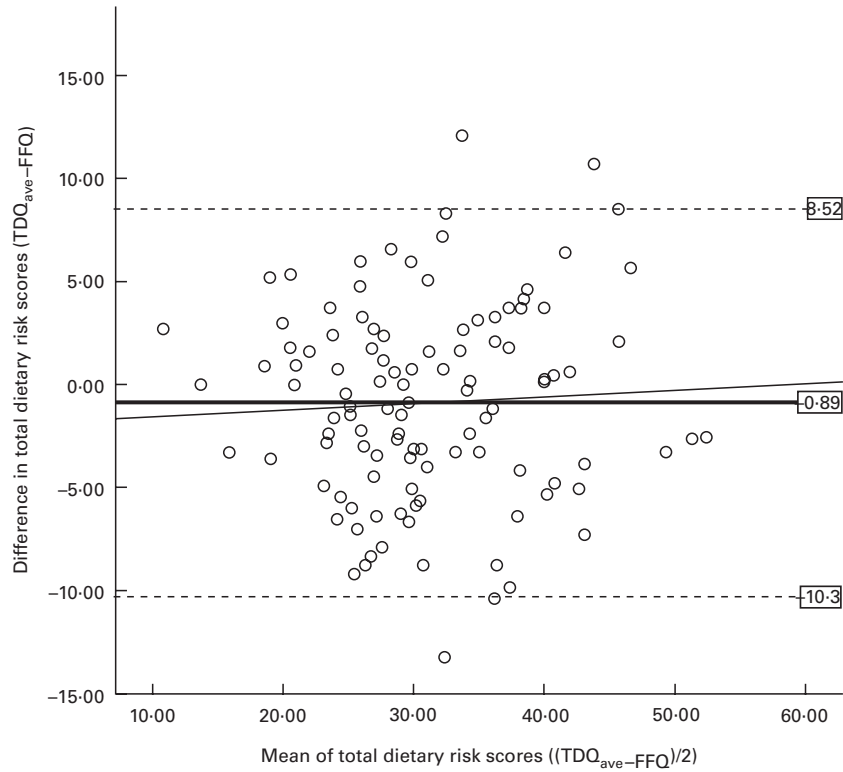
## Discussion

In the present article, the development and testing of a nineteen-item TDQ that assesses the dietary risk of children aged 12–36 months are described. Our findings revealed that the TDQ-derived dietary risk scores of toddlers in the study sample were highly correlated and not significantly different between the two administrations or on comparison with scores derived from a fifty-four-item FFQ. The TDQ is a reliable and valid screening tool for assessing the dietary risk of Australian toddlers from relatively advantaged backgrounds and categorising them into dietary risk categories. The reliability and validity of the TDQ in samples that include Australian toddlers from the lower five SEIFA deciles are yet to be assessed.

The TDQ performed well in terms of reliability. Repeatability analysis of individual questionnaire items revealed predominately moderate agreement. The percentage of agreement (32–86%; *n* 19 items) was slightly lower than that reported for a FFQ tested in Australian 2- to 5-year-olds (53–97%; *n* 16 items)<sup>(37)</sup>. Yet *K<sub>w</sub>* values derived from the FFQ (0.37 (red meat)–0.85 (take-away foods)) and those derived from the TDQ used in the present study (0.40 (grains)–0.78 (bread)) were similar; the reproducibility of the TDQ was predominately ‘moderate’ (*n* 13/19 items) or ‘good’ (*n* 5/19 items). Test–retest analysis of dietary risk scores revealed that the TDQ is reliable for assessing individuals’ dietary risk. At the group level, total risk scores were not significantly different, with less than one risk score point being observed between the mean scores during each administration. The mean bias was greatest for section 3, with risk scores being statistically, but not meaningfully different (0.9 points out of 48; 1.9%), between the administrations. Classification analysis revealed three-quarters of the children to be in the same dietary risk category during each TDQ administration. Overall, these results suggest that the TDQ is reliable for assessing dietary risk in this population, an important finding considering that the validity of a tool requires reliability<sup>(43)</sup>.

The TDQ performed well in terms of validity. The nineteen-item TDQ accurately derives dietary risk scores and assigns toddlers to risk categories in comparison with a longer fifty-four-item FFQ. The total dietary risk scores derived from the TDQ<sub>ave</sub> and those derived from the FFQ were highly correlated and not significantly different. The Bland–Altman plot for total dietary risk scores revealed narrow limits of agreement, indicating that the TDQ can accurately distinguish dietary risk at the individual level<sup>(44)</sup>. As the slope of the mean bias line indicated no overall bias, the TDQ is acceptable for measuring the dietary risk of toddlers at the group level<sup>(17)</sup>. Classification analysis revealed promising results with the majority of children (approximately three-quarters) being classified into the same dietary risk category by the TDQ<sub>ave</sub> and FFQ. Thus, the TDQ is a valid toddler dietary risk assessment tool suitable for this population in a clinical (individual) or community (group) setting.

The fifty-four-item FFQ developed by Huybrechts *et al.*<sup>(34,35)</sup> was chosen as the reference tool to assess validity. In the absence of a gold standard to measure dietary intake, this FFQ was determined to be the best available validation tool.



**Fig. 1.** Bland–Altman plot assessing the validity of total dietary risk scores derived from the average Toddler Dietary Questionnaire (TDQ<sub>ave</sub>) v. those derived from the FFQ among Australian children (*n* 111) aged 12–36 months. The plot shows the mean difference (—), the 95% limits of agreement (---) and the fitted regression line (—) for total dietary risk scores (*P* for linear trend=0.595). *R*<sup>2</sup> linear = 0.003.

It has been shown to be reliable and valid in terms of food<sup>(34)</sup> and nutrient<sup>(35)</sup> intake assessment compared with the estimated diet records and provides a reasonable measure when compared with the TDQ, capturing the intake data of key foods of interest over a similar time period. Despite this, minor changes were made to the FFQ primarily to reflect cultural differences, possibly altering the reliability and validity of the tool. Ideally, the tool would have been retested in the Australian population; however, this was not feasible within the study constraints. Additionally, translation of items from the FFQ into the TDQ was challenged by incompatible portion-size categories for some items (e.g. fish, snack products, chocolates and ice creams/frozen yogurt). That is, a ‘small’

response in the FFQ was translated to ‘medium’ in the TDQ, while both ‘medium’ and ‘large’ responses in the FFQ were translated to ‘large’ in the TDQ. Nonetheless, this FFQ was the most compatible tool that allowed derivation of dietary risk scores, could be completed in the participants’ own time and was considered least burdensome for the participants.

The novelty of the TDQ is demonstrated by the innovative approach to the selection of food items, through the use of principal component analysis-derived dietary patterns and the formation of portion-size categories, based on toddlers’ intakes. Due to its novel nature, there are few similar tools evaluating an overall score of diet quality in young children with which it can be compared. In a Canadian study of 3- to

**Table 5.** Cross-classification of participants into dietary risk categories (low, moderate, high and very high) between the administrations of the Toddler Dietary Questionnaire (TDQ) and average TDQ (TDQ<sub>ave</sub>) and FFQ (*n* 111)\* (Number of participants and percentages)

	Test–retest reliability (TDQ1 and TDQ2)						Relative validity (TDQ <sub>ave</sub> † and FFQ)						
	TDQ2						FFQ						
	Low		Moderate		High		Low		Moderate		High		
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
TDQ1													
Low	20	18	11	10	–	–	16	14	17	15	–	–	
Moderate	13	12	61	55	2	2	5	5	70	63	1	1	
High	–	–	2	2	2	2	–	–	–	–	2	2	

\* No subject was classified as ‘very high risk’ by the TDQ1, TDQ2, TDQ<sub>ave</sub> or FFQ.

† TDQ<sub>ave</sub> = ((TDQ1 risk scores + TDQ2 risk scores)/2).





5-year-old preschool children, a seventeen-item Nutritional Screening Tool for Every Preschooler questionnaire, which derives a nutrition risk score from five food-group questions and twelve questions on other nutrition risk constructs, was reliable between the administrations (ICC = 0.89) and valid ( $r$  0.48) on comparison with a dietitian rating<sup>(45)</sup>. The total dietary risk scores (reliability, ICC = 0.90; validity,  $r$  0.83) obtained in the present study were comparatively better.

Besides the reliability and validity results of the TDQ, the present study provides information on the dietary risk of Australian toddlers. Scores derived from the TDQ categorised approximately one-third of the study sample as 'low' risk and two-thirds as 'moderate' risk. Few toddlers were categorised as 'high' risk and none as 'very high' risk. This is probably explained by our homogeneous sample, whereby the majority were highly educated, in paid employment and of a relatively high socio-economic status. Additionally, enrolment in the present study was voluntary and thus the participants were likely to be highly motivated parents. The assessment of dietary risk in a more representative sample of toddlers may yield higher proportions at 'high' or 'very high' risk.

There are several potential uses of the TDQ. In the clinical setting, it could be used by health professionals to rapidly screen the dietary intakes of toddlers from relatively advantaged backgrounds, accurately identify those at risk, and facilitate referral to a dietitian for detailed assessment and intervention to improve dietary patterns. Once tested in a more generalisable sample, the TDQ could be applied in this manner to low-socio-economic status populations. This is important considering that diet quality is socially patterned, whereby consumption of a less-healthy diet is observed in socio-economically disadvantaged populations<sup>(46,47)</sup>. Furthermore, it could potentially be useful in the research setting, for population health monitoring of toddlers' dietary risk, for exploring the sociodemographic predictors of dietary risk, and for furthering our understanding of the relationship between dietary risk and health outcomes. Additionally, as contemporary interventions commonly focus on food-based dietary guidelines, the food-group-based TDQ is particularly useful for developing relevant interventions that aim to improve toddlers' dietary patterns and for determining the effectiveness of these interventions. Thus, further testing of the TDQ is warranted to ensure wider applicability.

The findings of the present study should be interpreted within the context of the strengths and limitations. The TDQ is a novel tool developed based on population-specific evidence and age-appropriate public health dietary recommendations. It is easy and inexpensive to administer and calculates an overall dietary risk score. It does not rely heavily on memory, particularly, in comparison with other short tools<sup>(34,35,48,49)</sup>. Additionally, the high participation rate in the present study suggests that completion of the TDQ is not burdensome for the respondents. Reliability and validity testing was undertaken in a sample size consistent with that recommended for validation studies ( $>100$ <sup>(44,50)</sup>) and the size was comparatively larger than that used in similar studies<sup>(37,51,52)</sup>. Furthermore, we investigated the repeatability

of individual questionnaire items in addition to the reliability and validity of dietary risk scores. Nonetheless, our findings may not be representative of those in the general population due to the highly educated and motivated study sample, although social desirability bias is possible given the self-reporting nature of dietary intake<sup>(53)</sup>. Moreover, while attempts were made to ensure that stage 2 questionnaires were completed approximately 2–4 weeks after the completion of stage 1 questionnaires, this could not be standardised. Consequently, participants completing each stage within 1–2 weeks ( $n$  15) may have remembered their previous responses, while true changes in diet may have occurred for those completing each stage over 5 weeks apart ( $n$  7). To overcome this, however, average risk scores from each TDQ administration were used in the validity analysis. Lastly, despite its limitations<sup>(54,55)</sup>, we used  $K_w$  as a measure of agreement as it is frequently used for ordinal food frequency data<sup>(37,48)</sup> and chose linear analysis over quadratic analysis due to its lower sensitivity to increasing number of categories<sup>(56)</sup>.

In conclusion, the TDQ is a short assessment tool that provides information on toddlers' dietary risk, allowing identification of those requiring intervention. The present study showed that the TDQ is reliable and valid and accurately categorises toddlers from relatively advantaged backgrounds into dietary risk categories. The TDQ may be useful in the clinical setting, enabling screening of toddlers to identify those at risk requiring intervention, and potentially in the research setting for the development and evaluation of interventions. Overall, the TDQ is a multi-purpose tool ideal for preventative nutrition promotion efforts.

### Supplementary material

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S0007114514001184>

### Acknowledgements

The authors thank the families that took part in the present study and all child care centre directors for their help with recruitment. The present study was funded by the National Health and Medical Research Council of Australia and Health. R. K. G. is supported by a National Heart Foundation Fellowship and L. K. B. by an Australian Postgraduate Award.

The authors' contributions are as follows: A. M. M. and R. K. G. designed the research; L. K. B., A. M. M. and R. K. G. contributed to the development of the TDQ; L. K. B. was responsible for the conduct of the study and performed the statistical analysis; L. K. B. wrote the manuscript. All authors reviewed, edited and approved the final manuscript.

None of the authors has any conflicts of interest to declare.

### References

1. Stallings VA, Baranowski T, Briefel RR, *et al.* (2002) *Dietary Risk Assessment in the WIC Program*. Food and Nutrition Board, Institute of Medicine. Washington, DC: The National Academies Press.

2. Cowbrough K (2010) Feeding the toddler: 12 months to 3 years – challenges and opportunities. *J Fam Health Care* **20**, 49–52.
3. Dovey TM, Staples PA & Gibson EL (2008) Food neophobia and ‘picky/fussy’ eating in children: a review. *Appetite* **50**, 181–193.
4. Craigie AM, Lake AA, Kelly SA, *et al.* (2011) Tracking of obesity-related behaviours from childhood to adulthood: a systematic review. *Maturitas* **70**, 266–284.
5. Skinner JD, Carruth BR, Bounds W, *et al.* (2002) Do food-related experiences in the first 2 years of life predict dietary variety in school-aged children? *J Nutr Educ Behav* **34**, 310–315.
6. Owen CG, Martin RM, Whincup PH, *et al.* (2005) Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence. *Pediatrics* **115**, 1367–1377.
7. Wu TC & Chen PH (2009) Health consequences of nutrition in childhood and early infancy. *Pediatr Neonatol* **50**, 135–142.
8. Bates B, Lennox A & Swan G (2010) *National Diet and Nutrition Survey: Headline Results from Year 1 of the Rolling Programme (2008/2009)*. London: Food Standards Agency and Department of Health <http://multimedia.food.gov.uk/multimedia/pdfs/publication/ndnsreport0809.pdf> (accessed March 2014)
9. Chan L, Magarey AM & Daniels LA (2011) Maternal feeding practices and feeding behaviors of Australian children aged 12–36 months. *Matern Child Health J* **15**, 1363–1371.
10. Siega-Riz AM, Kinlaw A, Deming DM, *et al.* (2011) New findings from the Feeding Infants and Toddlers Study 2008. *Nestle Nutr Workshop Ser Pediatr Program* **68**, 83–100.
11. Siega-Riz AM, Deming DM, Reidy KC, *et al.* (2010) Food consumption patterns of infants and toddlers: where are we now? *J Am Diet Assoc* **110**, Suppl. 12, S38–S51.
12. Morgan J (2005) Nutrition for toddlers: the foundation for good health – 1. toddlers’ nutritional needs: what are they and are they being met? *J Fam Health Care* **15**, 56–59.
13. Lioret S, McNaughton SA, Spence AC, *et al.* (2013) Tracking of dietary intakes in early childhood: the Melbourne InFANT Program. *Eur J Clin Nutr* **67**, 275–281.
14. World Health Organisation (WHO) (2002) *Diet, Nutrition and the Prevention of Chronic Diseases: Report of a Joint WHO/FAO Expert Consultation*. Geneva: WHO.
15. Brazionis L, Golley RK, Mittinty MN, *et al.* (2013) Diet spanning infancy and toddlerhood is associated with child blood pressure at age 7.5 y. *Am J Clin Nutr* **97**, 1375–1386.
16. Collins CE, Watson J & Burrows T (2010) Measuring dietary intake in children and adolescents in the context of overweight and obesity. *Int J Obes* **34**, 1103–1115.
17. Magarey A, Golley R, Spurrier N, *et al.* (2009) Reliability and validity of the Children’s Dietary Questionnaire; a new tool to measure children’s dietary patterns. *Int J Pediatr Obes* **4**, 257–265.
18. 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.
19. Sinkowitz-Cochran RL (2013) Survey design: to ask or not to ask? That is the question. *Clin Infect Dis* **56**, 1159–1164.
20. National Health and Medical Research Council (NHMRC) (2013) *Eat for Health: Australian Dietary Guidelines; Providing the Scientific Evidence for Healthier Australian Diets*. Canberra: NHMRC.
21. National Health and Medical Research Council (NHMRC) (2013) *Australian Guide to Healthy Eating*. Canberra: NHMRC.
22. Dennison BA, Jenkins PL & Rockwell HL (2000) Development and validation of an instrument to assess child dietary fat intake. *Prev Med* **31**, 214–224.
23. Bennett CA, de Silva-Sanigorski AM, Nichols M, *et al.* (2009) Assessing the intake of obesity-related foods and beverages in young children: comparison of a simple population survey with 24 hr-recall. *Int J Behav Nutr Phys Act* **6**, 71.
24. Bell LK, Golley RK & Magarey AM (2013) Short tools to assess young children’s dietary intake: a systematic review focusing on application to dietary index research. *J Obes* **2013**, 709626.
25. Bell LK, Golley RK, Daniels L, *et al.* (2013) Dietary patterns of Australian children aged 14 and 24 months, and associations with socio-demographic factors and adiposity. *Eur J Clin Nutr* **67**, 638–645.
26. National Health and Medical Research Council (NHMRC) (2011) *A Modelling System to Inform the Revision of the Australian Guide to Healthy Eating*. Canberra: NHMRC.
27. National Health and Medical Research Council (NHMRC) (2003) *Dietary Guidelines for Children and Adolescents in Australia; Incorporating the Infant Feeding Guidelines for Health Workers*. Canberra: NHMRC.
28. Kellet L, Smith A & Schmerlaib Y (1998) *Australian Guide to Healthy Eating (AGHE)*. Canberra: Commonwealth Department of Health and Family Services.
29. Wajers PM, Feskens EJ & Ocke MC (2007) A critical review of predefined diet quality scores. *Br J Nutr* **97**, 219–231.
30. Daniels LA, Magarey A, Battistutta D, *et al.* (2009) The NOURISH randomised control trial: positive feeding practices and food preferences in early childhood – a primary prevention program for childhood obesity. *BMC Public Health* **14**, 387.
31. Craig LC, McNeill G, Macdiarmid JI, *et al.* (2010) Dietary patterns of school-age children in Scotland: association with socio-economic indicators, physical activity and obesity. *Br J Nutr* **103**, 319–334.
32. World Health Organisation (WHO) (2000) Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser* **894**, 1–253.
33. Cooke L (2007) The importance of exposure for healthy eating in childhood: a review. *J Hum Nutr Diet* **20**, 294–301.
34. 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.
35. 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.
36. Australian Bureau of Statistics (ABS) (2006) *Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA)*. ABS Report no. 2039.0. Canberra: ABS.
37. Flood VM, Wen LM, Hardy LL, *et al.* (2013) 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.
38. Magarey A, Watson J, Golley RK, *et al.* (2010) Assessing dietary intake in children and adolescents: considerations and recommendations for obesity research. *Int J Pediatr Obes* **6**, 2–11.
39. Sim J & Wright CC (2005) The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* **85**, 257–268.
40. Altman DG (1991) *Practical Statistics for Medical Research*, 1st ed. Boca Raton, London, New York, Washington, DC: Chapman and Hall/CRC.





41. Landis JR & Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174.
42. Pallant J (2011) *SPSS Survival Manual: A Step by Step Guide to Data Analysis using SPSS*, 4th ed. Crows Nest: Allen and Unwin.
43. Vogt WP (1999) *Dictionary of Statistics & Methodology. A Nontechnical Guide for the Social Sciences*. 2nd ed. Newbury Park, CA: SAGE Publications Inc.
44. Bland JM & Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* **8**, 135–160.
45. Randall Simpson JA, Keller HH, Rysdale LA, *et al.* (2008) Nutrition Screening Tool for Every Preschooler (NutriSTEP): validation and test–retest reliability of a parent-administered questionnaire assessing nutrition risk of preschoolers. *Eur J Clin Nutr* **62**, 770–780.
46. Turrell G, Bentley R, Thomas LR, *et al.* (2009) A multilevel study of area socio-economic status and food purchasing behaviour. *Public Health Nutr* **12**, 2074–2083.
47. James WP, Nelson M, Ralph A, *et al.* (1997) Socioeconomic determinants of health. The contribution of nutrition to inequalities in health. *BMJ* **314**, 1545–1549.
48. Lanfer A, Hebestreit A, Ahrens W, *et al.* (2011) Reproducibility of food consumption frequencies derived from the Children's Eating Habits Questionnaire used in the IDEFICS study. *Int J Obes* **35**, Suppl. 1, S61–S68.
49. Ebenegger V, Marques-Vidal P, Barral J, *et al.* (2010) Eating habits of preschool children with high migrant status in Switzerland according to a new food frequency questionnaire. *Nutr Res* **30**, 104–109.
50. Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.
51. Marriott LD, Robinson SM, Poole J, *et al.* (2008) What do babies eat? Evaluation of a food frequency questionnaire to assess the diets of infants aged 6 months. *Public Health Nutr* **11**, 751–756.
52. 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.
53. Hebert JR, Clemow L, Pbert L, *et al.* (1995) Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *Int J Epidemiol* **24**, 389–398.
54. Maclure M & Willet WC (1987) Misinterpretation and misuse of the kappa statistic. *Am J Epidemiol* **126**, 161–169.
55. Roberts C & McNamee R (2005) Assessing the reliability of ordered categorical scales using kappa-type statistics. *Stat Methods Med Res* **14**, 493–514.
56. Brenner H & Kliebsch U (1996) Dependence of weighted kappa coefficients on the number of categories. *Epidemiology* **7**, 199–202.