Accepted Manuscript

Reproducibility and relative validity of a semi-quantitative food frequency questionnaire in European preschoolers: the ToyBox-study

T. Mouratidou , M.I. Mesana , I. Huybrechts , E. De Decker , M. De Craemer , O. Androutsos , Y. Manios , S. Galcheva , M. Lateva , B. Gurzkowska , Z. Kułaga , J. Birnbaum , B. Koletzko , L.A. Moreno , on behalf of the ToyBox-study group

PII: S0899-9007(18)30704-4
DOI: https://doi.org/10.1016/j.nut.2019.03.003
Reference: NUT 10472

To appear in: *Nutrition*

Received date: 24 July 2018
Revised date: 9 February 2019
Accepted date: 13 March 2019

Please cite this article as: T. Mouratidou , M.I. Mesana , I. Huybrechts , E. De Decker , M. De Craemer , O. Androutsos , Y. Manios , S. Galcheva , M. Lateva , B. Gurzkowska , Z. Kułaga , J. Birnbaum , B. Koletzko , L.A. Moreno , on behalf of the ToyBox-study group, Reproducibility and relative validity of a semi-quantitative food frequency questionnaire in European preschoolers: the ToyBox-study, *Nutrition* (2019), doi: https://doi.org/10.1016/j.nut.2019.03.003

Highlights

- A food frequency questionnaire for preschoolers offered moderate-good reproducibility.
- A food frequency questionnaire for preschoolers offered low-moderate validity.
- Results were good depending of food and beverage group.

Reproducibility and relative validity of a semi-quantitative food frequency questionnaire in European preschoolers: the ToyBox-study

Running Title: **Reproducibility and validity of a food frequency questionnaire in preschoolers**

Mouratidou T[1], Mesana MI[1,2,3,4], Huybrechts I[5,6], De Decker E[7], De Craemer M[7], Androutsos O[8], Manios Y[8], Galcheva S[9], Lateva M[9], Gurzkowska B[10], Kułaga Z[10], Birnbaum J[11], Koletzko B[11], Moreno LA[1,2,3,12,13] on behalf of the ToyBox-study group

[1]GENUD (Growth, Exercise, NUtrition and Development) Research Group, University of Zaragoza, Zaragoza, Spain

[2]Instituto Agroalimentario de Aragón (IA2),

[3]Instituto de Investigación Sanitaria Aragón (IIS Aragón),

[4]Red de Salud Materno-infantil y del Desarrollo (SAMID)

[5]Department of Public Health, Ghent University, Ghent, Belgium

[6]International Agency for Research on Cancer, Lyon, France

[7]Department of Movement and Sport Sciences, Ghent University, Ghent, Belgium

[8]Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece

[9]Medical University Varna, Varna, Bulgaria

[10]The Children's Memorial Health Institute, Warsaw, Poland

[11]Dr. von Hauner Children's Hospital, University of Munich Medical Centre, Munich, Germany

[12]Faculty of Health Sciences, University of Zaragoza, Zaragoza, Spain

[13]Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición (CIBERObn)

Mouratidou T, conception and design of the study, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript

Mesana MI, conception and design of the study, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript

Huybrechts I, conception and design of the study, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

De Decker E, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

De Craemer M, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

Androutsos O, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

Manios Y, drafting or revision of the manuscript, approval of the final version of the manuscript

Galcheva SV, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

Lateva M, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

Gurzkowska B, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

Kułaga Z, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

Birnbaum J, generation, collection, assembly, analysis and/or interpretation of data; drafting or revision of the manuscript, approval of the final version of the manuscript

Koletzko B, drafting or revision of the manuscript, approval of the final version of the manuscript

Moreno LA, drafting or revision of the manuscript, approval of the final version of the manuscript

Word count for the entire manuscript (including figures and tables): 6.402

Number of figures and tables: 4 Tables

Corresponding author: María Isabel Mesana Graffe, GENUD (Growth, Exercise, NUtrition and Development) Research Group, University of Zaragoza,

C/ Pedro Cerbuna 12, 50009 Zaragoza, Spain.

e-mail: mmesana@unizar.es

Telephone number: +34 87 65 5375, Fax: +34 976 761 752

Abstract

Objectives: To examine the reproducibility and relative validity of a semi-quantitative food frequency questionnaire (FFQ) in assessing food group estimates.

Methods: Food group estimates were assessed via a 37-item FFQ and a three-day food record (FR). Pearson´s correlation coefficients for log-transformed values were calculated to assess the reproducibility and Spearman's rank correlation coefficients for log-transformed values were calculated to assess the validity. Kindergartens from six European countries participated in the preparatory substudies of the ToyBox-intervention study; data from preschool children aged 4-6 years (n=196, reproducibility study; n=324, validation study) were obtained.

Results: In the reproducibility study, positive Pearson's correlation coefficients for single and aggregated food groups ranged from 0.14 for pasta and rice to 0.90 for cooked vegetables. In the validation study, the FR gave higher estimates of 40 of the 50 food items (single and aggregated) examined, compared to those obtained from the FFQ. Positive crude Spearman rank correlation coefficients ranged from 0.01 for total beverages (added sugar) and rice to 0.62 for tea. Corrections for the de-attenuation effect did not improve observed correlations. Quartiles/tertiles were calculated for a small number of food groups (n=14) due to zero consumption in the rest of the groups. **Conclusions:** Moderate-good reproducibility and low-moderate relative validity of the FFQ used in preschool children was observed. Relative validity however, varied by food and beverage group; for some of the "key" foods/drinks targeted in the ToyBox-intervention (e.g.

biscuits), the validity was good. The findings should be considered in future epidemiologic and intervention studies in preschool children.

Keywords: reproducibility; validity; food-frequency questionnaire; preschool children

List of abbreviations:

FFQ, Food frequency questionnaires; FR, Food records

Introduction

The preschool period plays an important role in the management of the weight as many of the energy balance-related behaviours including eating habits are developed and adopted early in life [1]. At a European level, studies have already reported low compliance of children with nutrition recommendations [2-5]. Food frequency questionnaires (FFQ) are often used to assess food consumption/nutrient intakes in a wide variety of settings and populations because of their user-friendliness. Food records (FR) on the other hand are often used as reference method but impose high participant burden and require a relatively high level of literacy [6-7]. All retrospective or prospective self-reporting dietary methodologies are prone to measurement error (random and systematic) leading to bias in estimates, which may not necessarily represent the “true” usual intake [6-7]. For this reason, this kind of evaluation studies are necessary to assess the effect of measurement error and prevent incorrect estimations, in order to strengthen evidence-based public health recommendations [7-8].

The validity of an FFQ in assessing nutrient intakes in various population groups and settings is well documented, but not its ability to capture food group estimates especially in groups of preschool children [9]. The latter is particularly important within the context of the importance of a whole food approach rather than that of individual nutrients when it comes to public health prevention strategies [10]. Recently, in a study with children aged 2-9

years from eight European countries, Bel-Serrat et al. [11] examined the agreement of proxy-reported food group estimates from an FFQ and two non-consecutive 24h dietary recalls highlighting the importance of validation studies in young population groups.

The FFQ used in the ToyBox-study was adapted from the one developed by Huybrechts et al. [12] among preschool children (2.5-6.5 years old) in Flanders. The aim of this study was to evaluate the reproducibility and relative validity of a proxy self-administered, semi-quantitative FFQ in a European population of preschool children participating in the ToyBox-study (more specifically, in the preparatory substudies of the ToyBox-intervention study).

Methods and materials

Study design and population

The ToyBox-study aimed to develop, implement and evaluate the effectiveness of a kindergarten-based, family-involved intervention (assessment of children's snacking behaviour including salty and sweet snacks, and beverage consumption) in preschool children applied in six European countries (Belgium (Flanders), Bulgaria, Greece, Germany, Poland and Spain). The design and methodology of the ToyBox-study is described elsewhere [13-15]. Data for the current analysis were obtained from the ToyBox reproducibility and relative validity studies, prior to the ToyBox-intervention. These studies were performed among pre-schoolers aged 4-6 years from different socioeconomic levels randomly selected. The studies took place between September and October 2011 in each of the six participating centres. The characteristics of the sample in the test-retest and validity of the FFQ reflected those of the main study. The ratio of males and females was balanced. Different schools and classes were recruited for the reproducibility and the validity study, respectively; nevertheless, the participants were representative to the whole Toy-Box study. [15,16].

In the recruitment, municipalities within 50 km vicinity in all participating countries were sorted by common (i.e. available in all countries) SES variables. Tertiles of municipalities were created and a convenient sample of municipalities was selected from each tertile. Within each tertile, a list of the kindergartens in the selected municipalities was created; kindergartens were randomly selected from each tertile (1/3 of kindergartens came from the selected municipalities in the first tertile, 1/3 from the selected municipalities in the second one and 1/3 from the selected municipalities in the third one). In addition, no differences between compliers and non-compliers were found, in terms of socioeconomic status (data not showed). Teachers were informed about the project and asked to distribute and collect envelopes containing information sheet (instructions for children´s proxies, mainly parents) and the appropriate questionnaires. Children took home the envelopes. Support and explanation were provided individually if required.

In the reproducibility study participants provided two FFQs within a two-week interval (FFQ1 and FFQ2) and participants in the validity study provided one FFQ followed by a three-day FR, with an interval of at least 7 days separating the FFQ administration to the food diary. Reproducibility addressed the question of how consistent the answers were from one occasion to the next in the same subject in terms of food/beverage group estimates. Validity was determined by measuring the agreement of the FFQ food group estimates with FR estimates averaged across three consecutive predefined food records/days obtained for each child participating (or two in some occasions). One of the 3 days was a weekend day and the rest two week days. Therefore, two possible combinations for data collection existed: 1. Sunday, Monday and Tuesday or 2. Thursday, Friday and Saturday. The food diaries were distributed on the day of the FFQ collection. Finally, the teacher collected the envelopes with the completed FR the agreed day (arranged between investigator, teacher and proxies within a period of 7 or 14 days after the administration) and gave them to the investigator.

In the reproducibility study, 30 to 50 participants should be recruited per country, and finally in total, 196 participants were included in the analysis, providing one complete FFQ1 and one complete FFQ2, of 326 recruited participants. Participants who do not provide one complete FFQ1 and one complete FFQ2 were excluded from the analysis (130).

In the validity study, 60 to 100 participants should be recruited per country, and finally in total, 324 participants were included in the analysis, 279 participants providing one complete FFQ and a three-day FR and 45 participants providing a complete FFQ and a two-day FR, of 331 recruited participants. Participants who provided a single day FR (7) were excluded of the analysis.

Food frequency questionnaire

The proxy-reported semi-quantitative FFQ [16] covered a wide-range of food items to address children's food and beverage consumption relevant to the ToyBox-intervention objectives.

The FFQ consisted of 37 food and beverage items in total and portion size estimates were obtained for each of the food items. Some food groups were aggregated or considered in the same group according to their nutritional profile at a higher level and these results are also presented in the tables (13 aggregated food groups): (1) Water, (2) Soft drinks/beverages with added sugar, (3) Soft drinks/ beverages light, (4) Fizzy drinks (all), (5) Fruit juice, homemade, freshly squeezed, (6) Fruit juice, pre-packed, bottled, (7) Juices (all), (8) Tea, (9) Smoothies (all kinds), (10) Total beverages (added sugar), (11) Plain milk, (12) Sugared or chocolate milk, (13) Plain yogurt, (14) Fruit, sugared or aromatized yoghurt, (15) Yogurt (all), (16) Cheese, (17) Dried fruit, (18) Canned fruit, (19) Fresh fruit, (20) Fruits (all), (21) Raw vegetables, (22) Cooked vegetables, (23) Vegetables (all), (24) Chocolate, (25) Milk-based desserts, (26) Cakes, (27) Biscuits, (28) Pastries, (29) Cakes and pastries (all), (30) Sugar-based desserts, (31) Total sweets, (32) Chocolate spreads/

other sweet spreads, (33) Unsweetened breakfast cereals, (34) Sweetened breakfast cereals, (35) Breakfast cereals (all), (36) White bread and other bakery product, (37) Brown, whole grain bread and other bakery products, (38) Breads and bakery products (all), (39) Salty snacks, (40) Meat and poultry, (41) Fish and fish products, (42) Meat products, (43) Meat products (all), (44) Pasta, (45) Rice, (46) Pasta and rice (all), (47) Fried potato products, (48) Potatoes, (49) Potatoes (all), (50) Potatoes (all).

Due to the multi-centre nature of the study, country-specific food items were included. The parents/caregivers of the pre-schoolers filled in the questionnaire at home and reported the number of times the child consumed the food items included in the questionnaire over the last 12 months. The frequency responses consisted of the following categories of consumption: 'never/less than once a month', '1–3 times a month', '1 day a week', '2–4 days a week', '5–6 days a week', 'everyday'. These were converted into food group estimates per times per week ranging from 0 up to 30 and thereafter into times per day. To relate FFQ food consumption estimates to those of the FR, and to enable comparisons, the "number of times per day" as reported in the FFQ was equated to "number of portions per day" [11]. A guide with food portions (no house hold measures have been used) specifically developed for the study to assist the proxy reporters was used for both assessment methods.

Food item-specific information on the average amount per day was also obtained. The FFQ offers different options for the amount of the ingested food items. For example, options for average amount of plain yogurt included: (1) 65 g or less, (2) between 65 and 195 g and (3) 195 g or more.

The questionnaire inserts some questions on selected "key" foods and drinks, due to the relevance for the ToyBox-intervention objectives: beverages (water consumption, sugar sweetened beverages, low-calories beverages, freshly squeezed juices and manufactured juices), savoury snacks (nuts, chips and similar, crackers, pizza), sweet snacks (sweet bakery, chocolate, biscuits, confectionary, milk products) and fruit and vegetables

consumption. The associations between these foods and drinks with sedentary/physical activity behaviours have been reported previously [17].

The questionnaire also included five questions on foods eaten in-between meals and supplement use

.

Estimated food records

In this study, the relative validity of food intake estimates derived from the FFQ is evaluated by comparison with a three-day FR.

The proxies/parents received written instructions for the recording of the foods and drinks consumed by their child over the three days including one weekend day. The schoolteachers received written and oral instructions for the recording of foods and drinks consumed during school days (snacking and lunches). Teachers had to report this information about what the children consumed at school to the proxies/parents so that they could include it in the diaries. In these structured FR, days were subdivided into six eating occasions: breakfast, morning snacks, lunch, afternoon snacks, dinner, and evening snacks. Detailed information on the type (including brand names) and portion size of the foods consumed was collected using an open entry format (use of standard portion sizes was inevitable for some food products for which portions sizes were difficult to describe/estimate by the respondent). After collection, the FR was checked on quality and completeness by trained survey personnel. Only good quality FR, containing sufficiently detailed descriptions of the food products and portion sizes consumed, were included in the analysis. Each reported FR food item was mapped and subsequently matched against one of the 37 food items included in the FFQ.

Statistical methods

Statistical analyses were performed using the Predictive Analytics SoftWare (PASW, version 20; SPSS Inc., Chicago, IL, USA). Means and standard deviations (SD) were

calculated for food consumption estimates obtained from FFQ1, FFQ2 and FR (for single and aggregated food groups). Crude data was log-transformed ($\log_n$) to improve normality for all the 37 food groups. No cases were found and subsequently no cases were excluded from this analysis on the basis of being rarely consumed (< 5%) or due to incompleteness (participants with 25% of missing values in the FFQ1 and FFQ2). Pearson product-moment correlation coefficients for log-transformed values were calculated to assess the reproducibility of FFQ1 compared to FFQ2. Spearman's rank correlation coefficients (Spearman´s rho) for log-transformed values were calculated to assess the validity of the FFQ compared to FR. All validity coefficients were corrected for attenuation due to random error in the FR. De-attenuation of crude correlation coefficients ($r_{adjusted}$) was computed according to the equation from Willet [11]:

$r_{adjusted} = r_{observed} \sqrt{1+ \lambda_x / n_x}$

where λ_x is the ratio of the within- and between-person variances (variance ratio) for x, and n_x is the number of replicates for the x variables (here n = 3).

Agreement in ranking individuals was examined by the construction of quartiles for each food group. Non-consumers of a food group were considered as one group, and the remaining individuals were grouped into tertiles (adapted food groups) [11]. Cross-classification analyses were conducted only for a limited number of food items, fourteen food groups (11 non-adapted food groups and 3 adapted groups) due to zero consumption observed for more than 25% of the participants in the rest of the food items (in the FFQ and/or FR). Non-adapted food groups are: (1) Fruits (all), (2) Raw vegetables, (3) Cooked vegetables, (4) Vegetables (all), (5) Cakes and pastries (all), (6) White bread and other bakery product, (7) Breads and bakery products (all), (8) Meat products (all), (9) Pasta, (10) Rice, (11) Pasta and rice (all). Adapted food groups are: (1) Fruit juice, homemade, freshly squeezed, (2) Breakfast cereals (all), (3) Brown, whole grain bread and other bakery products.

Results

Reproducibility

Mean food consumption estimates obtained from the FFQ1 and FFQ2 (plus FFQ2 as % of FFQ1) is presented in Table 1 including single and aggregated groups. Mean consumption estimates for 29 food items of the questionnaire were higher for FFQ1 than FFQ2 including water, yogurt and raw vegetables, whereas FFQ2 gave higher estimates for 21 food items including soft drinks/beverages (light) and biscuits. The largest differences between administrations were observed for fruit juice, pre-packed, bottled (145% of FFQ1) and for potatoes, (244% of FFQ1). The rest of the comparisons showed a relatively high consistency between the two administrations (most of comparisons ranged between 88 and 119% of FFQ1). Significant Pearson correlation coefficients ranged from 0.14 for pasta and rice to 0.90 for cooked vegetables, even though for most food consumption estimates correlations ranged from 0.52 to 0.79 showing moderate correlation.

Table 1. Mean food group estimates (equalled to mean daily number of portions) from the FFQ1 and FFQ2 and Pearson correlation coefficients between the two FFQ administrations

Food group (portions/day)	**N**	**FFQ1 (mean)**	**SD**	**FFQ2 (mean)**	**SD**	**% of FFQ1**	**Pearson *r*†**
Water	316	1.15	0.42	1.10	0.43	104	0.768**
Soft drinks/ beverages with added sugar	313	0.15	0.29	0.13	0.25	117	0.698**
Soft drinks/ beverages light	312	0.02	0.08	0.03	0.11	73	0.554**
Fizzy drinks (all)	315	0.52	0.22	0.51	0.20	103	0.689**
Fruit juice, home made, freshly squeezed	315	0.22	0.27	0.22	0.27	100	0.738**
Fruit juice, pre-packed, bottled	316	0.46	0.56	0.32	0.31	145	0.668**
Juices (all)	316	0.69	0.21	0.68	0.21	101	0.693**

Tea	316	0.17	0.30	0.15	0.24	116	0.773**
Smoothies (all kinds)	312	0.06	0.17	0.05	0.13	131	0.582**
Total beverages (added sugar)	316	3.85	0.64	3.82	0.62	99	0.618**
Plain milk	314	0.69	0.41	0.66	0.40	105	0.760**
Sugared or chocolate milk	312	0.13	0.25	0.14	0.26	96	0.857**
Plain yogurt	305	0.27	0.28	0.25	0.27	107	0.752**
Fruit, sugared or aromatized yoghurt	307	0.20	0.23	0.22	0.23	92	0.708**
Yogurt (all)	307	0.67	0.21	0.66	0.21	101	0.659**
Cheese	312	0.28	0.22	0.25	0.22	108	0.672**
Dried fruit	310	0.06	0.17	0.06	0.16	103	0.702**
Canned fruit	314	0.04	0.11	0.03	0.09	114	0.659**
Fresh fruit	315	0.72	0.30	0.69	0.28	104	0.754**
Fruits (all)	315	1.04	0.13	1.02	0.11	102	0.710**
Raw vegetables	313	0.39	0.33	0.34	0.29	114	0.774**
Cooked vegetables	313	0.36	0.26	0.36	0.26	100	0.900**
Vegetables (all)	313	0.80	0.26	0.76	0.23	105	0.822**
Chocolate	313	0.20	0.20	0.20	0.21	100	0.741**
Milk-based desserts	311	0.15	0.17	0.13	0.16	119	0.623**
Cakes	314	0.18	0.20	0.16	0.19	112	0.459**
Biscuits	314	0.44	0.41	0.47	0.38	94	0.745**
Pastries	311	0.19	0.22	0.20	0.22	92	0.687**
Cakes and pastries (all)	314	0.60	0.18	0.59	0.17	101	0.529**
Sugar-based desserts	313	0.39	0.41	0.33	0.35	118	0.574**
Total sweets	314	7.18	1.49	7.05	1.44	98	0.745**
Chocolate spreads/ other sweet spreads	312	0.25	0.26	0.80	0.12	31	0.775**
Unsweetened breakfast cereals	305	0.15	0.24	0.16	0.24	96	0.711**
Sweetened breakfast cereals	299	0.20	0.25	0.19	0.24	109	0.762**
Breakfast cereals (all)	305	0.58	0.20	0.58	0.20	100	0.754**
White bread and other	314	0.59	0.36	0.57	0.36	104	0.586**

bakery product							
Brown, whole grain bread and other bakery products	304	0.33	0.37	0.31	0.35	104	0.696**
Breads and bakery products (all)	314	0.58	0.20	0.92	0.25	63	0.152*
Salty snacks	315	0.17	0.24	0.16	0.27	102	0.477**
Meat and poultry	311	0.50	0.22	0.48	0.21	105	0.610**
Fish and fish products	311	0.19	0.17	0.19	0.17	97	0.791**
Meat products	305	0.43	0.33	0.43	0.30	100	0.701**
Meat products (all)	311	0.87	0.25	0.88	0.19	99	0.098
Pasta	310	0.28	0.18	0.29	0.18	97	0.499**
Rice	312	0.15	0.12	0.16	0.13	96	0.619**
Pasta and rice (all)	312	0.67	0.15	0.66	0.14	102	0.149*
Fried potato products	312	0.10	0.11	0.11	0.11	88	0.644**
Potatoes	311	0.44	0.29	0.46	0.30	97	0.646**
Potatoes (all)	312	0.58	0.20	0.24	0.27	244	0.115
Legumes	311	0.18	0.22	1.10	0.43	104	0.671**

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

†log transformed estimates

Validity

Table 2 presents food consumption estimates (converted to mean daily number of portions) obtained from the FFQ and the FR (including single and aggregated groups). The FR gave higher estimates for 40 of the food items out of the 50 (mean Δ of -0.45 portions per day). Significant mean Δ differences between estimates were found for the majority of the food items; non-significant differences were observed for 6 items such as potatoes and unsweetened breakfast cereals. The largest mean Δ difference of 5.49 portions per day between the measurements was observed for total sweets ($p < 0.05$).

Table 2. Food group consumption estimates (equalled to mean daily number of portions) obtained from the FFQ and the FR

		FFQ		FR			
Food group (portions/day)	**N**	**Mean**	**SD**	**Mean**	**SD**	**Mean Δ**	**p value**[∞]
Water	321	2.64	1.31	0.93	0.54	1.71	0.000*
Soft drinks/ beverages with added sugar	324	0.29	0.72	0.27	0.51	0.02	0.661
Soft drinks/ beverages light	319	0.03	0.18	0.09	0.96	-0.06	0.023*
Fizzy drinks (all)	324	1.80	0.72	1.78	0.51	0.02	0.716
Fruit juice, home made, freshly squeezed	320	0.25	0.43	0.16	0.39	0.09	0.004*
Fruit juice, pre-packed, bottled	322	0.60	0.75	0.86	0.83	-0.26	0.000*
Juices (all)	322	0.68	0.23	0.09	0.52	0.59	0.000*
Tea	322	0.27	0.55	0.48	0.64	-0.21	0.000*
Smoothies (all kinds)	319	0.49	0.13	0.10	0.35	0.39	0.014*
Total beverages (added sugar)	324	3.84	0.85	3.87	0.78	-0.03	0.618
Plain milk	321	1.06	0.92	0.90	0.48	0.16	0.003*
Sugared or chocolate milk	318	0.20	0.46	0.35	0.52	-0.15	0.000*
Plain yogurt	317	0.29	0.36	0.42	0.52	-0.13	0.005*
Fruit, sugared or aromatized yoghurt	318	0.28	0.34	0.34	0.49	-0.06	0.049*
Yogurt (all)	318	1.93	0.39	1.60	0.55	0.33	0.000*
Cheese	320	0.32	0.32	0.76	0.65	-0.44	0.000*
Dried fruit	314	0.08	0.24	0.13	0.39	-0.05	0.033*
Canned fruit	321	0.04	0.11	0.13	0.33	-0.09	0.000*
Fresh fruit	323	1.07	0.61	1.29	0.81	-0.22	0.004*
Fruits (all)	323	2.82	0.37	3.03	0.57	-0.21	0.000*
Raw vegetables	321	0.45	0.51	1.34	1.11	-0.89	0.000*
Cooked vegetables	320	0.47	0.45	1.38	0.77	-0.91	0.000*
Vegetables (all)	321	2.19	0.59	3.53	1.21	-1.34	0.000*
Chocolate	323	0.24	0.32	0.52	0.66	-0.28	0.000*
Milk-based desserts	322	0.17	0.24	0.26	0.47	-0.09	0.002*
Cakes	322	0.29	0.44	0.37	0.53	-0.08	0.031*
Biscuits	321	0.78	0.88	0.62	0.61	0.16	0.002*
Pastries	322	0.25	0.43	0.65	0.65	-0.4	0.000*
Cakes and pastries (all)	322	1.91	0.50	2.20	0.57	-0.29	0.000*
Sugar-based desserts	320	0.61	0.91	0.29	0.50	0.32	0.000*
Total sweets	322	1.61	0.91	7.1	1.39	-5.49	0.000*
Chocolate spreads/ other sweet spreads	318	0.35	0.47	0.81	0.47	-0.46	0.000*
Unsweetened breakfast cereals	312	0.17	0.29	0.20	0.44	-0.03	0.259
Sweetened breakfast cereals	317	0.24	0.33	0.52	0.55	-0.28	0.000*
Breakfast cereals (all)	317	1.79	0.35	1.96	0.53	-0.17	0.000*
White bread and other bakery	322	0.94	0.80	1.25	0.61	-0.31	0.000*

product							
Brown, whole grain bread and other bakery products	315	0.50	0.69	0.59	0.70	-0.09	0.036*
Breads and bakery products (all)	322	2.47	0.66	2.93	0.62	-0.46	0.000*
Salty snacks	322	0.16	0.23	0.18	0.28	-0.02	0.000*
Meat and poultry	319	0.64	0.37	0.96	0.56	-0.32	0.000*
Fish and fish products	320	0.21	0.23	0.61	0.60	-0.4	0.000*
Meat products	311	0.58	0.52	0.95	0.63	-0.37	0.000
Meat products (all)	319	2.44	0.51	2.95	0.63	-0.51	0.000*
Pasta	319	0.34	0.23	0.64	0.56	-0.3	0.000*
Rice	320	0.14	0.13	0.48	0.51	-0.34	0.000*
Pasta and rice (all)	320	1.96	0.29	2.39	0.62	-0.43	0.000*
Fried potato products	318	0.09	0.10	0.18	0.40	-0.09	0.012*
Potatoes	320	0.62	0.53	0.68	0.50	-0.06	0.186
Potatoes (all)	320	1.90	0.27	2.09	0.55	-0.19	0.000*
Legumes	319	0.15	0.22	0.24	0.49	-0.09	0.004*

*significant at the 0.05 level (2-tailed)

∞ Paired t-test

Table 3 presents Spearman rank correlation coefficients (Spearman´s rho) of consumption estimates (equalled to daily number of portions) obtained from the FFQ and the FR for single and aggregated groups. Positive coefficients values ranged from 0.01 for total beverages (added sugar) and rice to 0.62 for tea. For nearly all food items absence or low correlations (0.01-0.25) were observed. Corrections for the attenuation effect due to random error observed in the FR did not change the observed crude correlations. De-attenuated correlation coefficients for single food items ranged from 0.01 for rice to 0.68 for tea.

Table 3. Spearman's rho correlation coefficients between food group consumption estimates (equalled to daily number of portions) obtained from the FFQ and the FR

Food group	**Spearman's rho**	**Variance ratio**	**De-attenuated correlation coefficient (Spearman)**
Water	0.233*	0.499	0.252*
Soft drinks/ beverages	0.091	0.857	0.103

with added sugar			
Soft drinks/ beverages light^	-0.037		
Fizzy drinks (all) ^	0.126*		
Fruit juice, home made, freshly squeezed	0.265**	0.434	0.284**
Fruit juice, pre-packed, bottled	0.304**	0.530	0.330**
Juices (all) ^	0.169*		
Tea	0.623**	0.614	0.684**
Total beverages (added sugar) ^	0.113*		
Smoothies (all kinds) ^	-0.020		
Plain milk	0.208**	0.766	0.233**
Sugared or chocolate milk	0.070	1.272	0.084
Plain yogurt	0.167**	1.553	0.206**
Fruit, sugared or aromatized yoghurt	0.082	0.949	0.094
Yogurt (all) ^	0.158**		
Cheese	0.155**	0.729	0.173**
Dried fruit^	0.256**		
Canned fruit^	0.038		
Fresh fruit	0.144*	0.483	0.155*
Fruits (all) ^	0.258**		
Raw vegetables	0.275**	0.619	0.302**
Cooked vegetables	0.014	0.693	0.016
Vegetables (all) ^	0.260**		
Chocolate	0.061	1.550	0.075
Milk-based desserts	0.188**	3.788	0.283**
Cakes	-0.009	0.689	-0.010
Biscuits	0.339**	1.174	0.400**
Pastries	0.088	1.049	0.102
Cakes and pastries (all) ^	0.088		
Sugar-based desserts	0.040	1.592	0.049
Total sweets^	0.023		
Unsweetened breakfast cereals	0.253**	0.935	0.290**
Sweetened breakfast cereals	0.283**	0.907	0.323**
Breakfast cereals (all) ^	0.246**		
White bread and other bakery product	0.148**	0.684	0.164**
Brown, whole grain bread and other bakery products	0.392**	0.543	0.426**
Breads and bakery products (all)	0.025		
Salty snacks	-0.057	1.397	-0.069

Meat and poultry	0.320**	0.831	0.362**
Fish and fish products	0.157**	3.953	0.239**
Meat products	0.198**	0.962	0.228**
Meat products (all) ^	-0.017		
Pasta	0.211*	1.414	0.256*
Rice	0.113*	1.490	0.138*
Pasta and rice (all) ^	-0.063		
Fried potato products	0.080		
Potatoes	0.049	1.163	0.058
Potatoes (all)	-0.052		
Chocolate spreads/ other sweet spreads	0.224**	0.982	0.258**
Legumes^	0.139*		

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

^variance was not calculated

Table 4 presents the results of the cross-classification agreement and weighted kappa (where possible) values describing the ability of the FFQ to classify individuals into the same quartile (tertile in the case of the adapted groups) of food consumption estimates as to those obtained with the FR. Cross-classification analyses were conducted only for a limited number of food items, fourteen food groups (11 non-adapted food groups and 3 adapted groups) due to zero consumption observed for more than 25% of the participants in the rest of the food items (in the FFQ and/or FR). In the non-adapted items (n=11), the proportion of subjects classified in the same quartile ranged from 14% for vegetables (all) to 37% for pasta in preschoolers and gross misclassification ranged from 6% for white bread and other bakery products to 25% for rice. Weighted kappa values for all food items addressed showed low agreement (< 0.20). In the adapted items (n=3), the proportion of subjects classified in the same quartile was 57% for fruit juice (homemade, freshly squeezed), 43% for breakfast cereals (all), and 63% for brown, whole grain bread and other bakery products and gross misclassification of 25%, 15% and 34% respectively

showing a substantial agreement for fruit juice (homemade, freshly squeezed) and fair agreement for whole grain bread and other bakery products.

Table 4. Cross-classification of food group consumption estimates (equalled to daily number of portions) obtained from the FFQ and the FR

	FFQ versus FR		
Food groups	**Correctly classified (%)**	**Grossly misclassified (%)**	**Weighted Kappa**
Fruits (all)	34	11	0.12
Raw vegetables	32	7	0.93
Cooked vegetables	25	7	-0.04
Vegetables (all)	14	7	0.11
Cakes and pastries (all)	28	8	0.04
White bread and other bakery product	27	6	0.03
Breads and bakery products (all)	23	12	-0.02
Meat products (all)	27	14	0.02
Pasta	37	19	0.06
Rice	35	25	0.05
Pasta and rice (all)	22	13	-0.04
Adapted food groups*			
Fruit juice, homemade, freshly squeezed	57	25	0.73
Breakfast cereals (all)	43	15	0.15
Brown, whole grain bread and other bakery products	63	34	0.36

*zero consumers were considered as one group and the rest of participants were classified into tertiles

Discussion

The aim of the present study was to evaluate the ability of the FFQ in estimating proxy-reported food consumption as part of the wider ToyBox-intervention objectives. To the authors' knowledge, this is one of the few studies, at least at a European level, assessing the reproducibility and relative validity of food estimates obtained via an FFQ in a sample of preschool children. Overall, the study findings demonstrated moderate-good reproducibility and low-moderate relative validity of the FFQ; however, one should note the

observed differences in relative validity across different food and beverage groups, and that for some key foods targeted by the ToyBox-intervention (like biscuits) good results were observed. Below, we relate observed findings to those reported elsewhere (studies which included similar age groups and European populations), but one should be aware that comparison of findings as such is often compromised due to the different FFQs used (nature and size of the questionnaire), population sample size and characteristics of the type of reference method used [18]. In the original FFQ [12], study performed in only one country, estimated diet records (3d FR) were used as reference method and reproducibility was measured by repeated FFQ administrations five weeks apart, showing an overall high level of reproducibility (for most foods a moderate correlation (0.5-0.7) was obtained between FFQ1 and FFQ2), and moderate levels of relative validity in estimating food group intakes (for median differences between the 3d FR and the FFQ, six food groups gave a difference of > 20% and the proportion of subjects classified within one quartile (in the same/adjacent category) by FFQ and FR ranged from 67% to 88%), in 2.5-6.5 year-old Belgian children (650 children were included in the validity analyses and 124 in the reproducibility analyses).

Reproducibility study

Mean estimates differed slightly between the two FFQ administrations and were within the range of ± 7%. The study by Huybrechts et al. [12] suggested good reproducibility for almost all food groups examined in a large survey of 2.5-6.5 years-old Flemish children (n=124). One should note however, that their study examined FFQ reproducibility in capturing food intakes expressed as grams per recall period whereas in this study food consumption was examined as daily number of portions. Furthermore, in a similar age group (258 children aged 2-9 years participating in the IDEFICS study), the results of the study conducted by Lanfer et al. [19] also showed moderate FFQ reproducibility. More specifically, Spearman's correlation coefficients ranged from 0.32 to 0.76, with lowest

values observed for ‘diet soft drinks’ and highest for ‘sweetened milk’, in the sample of 258 children aged 2-9 years participating in the IDEFICS study.

Validity study

The FR gave slightly higher estimates as compared to estimates obtained from the FFQ. Overall validity of the FFQ, as mentioned above, was low-moderate and differed by food and beverage groups; the same observations have also been found in other studies of young European population groups. For instance, the study by Bel-Serrat et al. [11] examined the agreement of proxy-reported food group estimates from an FFQ and two non-consecutive 24h dietary recalls and reported that observed associations varied by food group. Correlations slightly improved after correction for within-person variation. In our study, corrections for the de-attenuation effect did not significantly improve the correlations for food consumption estimates, meaning that there was not a large random error within the 3d FR as expected. The study by Huybrechts et al. [12] has also shown large differences by food groups when examining the relative validity of the FFQ. The results of the cross-classification analysis varied by the FFQ food and beverage groups examined; the limited number of groups for which quartiles/tertiles were calculated does not facilitate drawing of clear-cut conclusions. The study by Bel-Serrat et al. [11], in which cross-classification analysis was conducted in 15 out of the 36 food groups, reported limited ability of the FFQ in discriminating between quartiles of food groups.

Strengths and Weaknesses of the Study

Low-moderate validity (expressed as correlation values) observed in our study is not unexpected, as reports suggest that correlations in young population groups are generally lower as compared to those in adult populations [20]. There are a number of factors to which such observations can be attributed including the use of proxies (i.e. parents/caregivers and teachers) and underreporting, especially at an out-of-home setting

[21]. The fact that for a number of participants however, only proxy-reported FR data were available (e.g. school meal data not available) and in other cases only school FR data were available, could explain, almost partly, the observed results.

Another important factor is the nature of the diet of young population groups, which in a way reflects the difficulties and the complex nature of assessing energy balance-related behaviours of younger children [22]. The reference method used to assess the validity of an FFQ and its specific limitations are also very relevant. It is widely known that dietary and food information are subject to day-to-day variability and this fact could explain the observed low-moderate agreement between the methods in our study [20]. This point was addressed in our study by the use of three-day (or two-days in some occasions) FR, which is supposed to be more accurate to other methods of dietary assessment such as a single 24h dietary recalls [20], and because of disadvantages of longer recording periods and in the hope to minimize the refusal rate and/or drop-out within the study. Difficulties in portion size estimations during completion of the three-day FR might however bias the true validity of the FFQ.

Differences might also reflect different recall timeframes covered by the two methods. In this way, a higher number of record days in the FR, distributed throughout the year, would have been better as reference method, since this fact could take into account seasonal variation of intake.

Another limitation to the study could be that to relate FFQ food consumption estimates to those of the FR, the “number of times per day” as reported in the FFQ was equated to “number of portions per day”.

An important strength of the study is that it was conducted using standardized procedures during the fieldwork data collection with a high level of quality control procedures applied throughout the sub-study [16]. Another strength of the study is the fact that the FFQ includes country-specific food items due to the multi-centre nature of the study, and that the reproducibility and validity studies were performed in a heterogeneous sample

comprising the six countries. The high sample size in both the reproducibility and validity assessments is also a strength of the study.

Conclusions

Overall, the findings of this study suggest moderate-good reproducibility and low-moderate validity of the FFQ used in the ToyBox-study. The latter however, varied by food and beverage groups. For some of the "key" foods/drinks targeted in the ToyBox-intervention (e.g. biscuits) the results were good. The results of this analysis should be taken into consideration when interpreting future results of the ToyBox-intervention study and to inform the design and data interpretation of future studies addressing similar objectives.

Ethical standards

All studies have been approved by the appropriate ethics committees and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All persons gave their informed consent prior to their inclusion in the study.

References

[1]. Reilly JJ (2008) Physical activity, sedentary behaviour and energy balance in the preschool child: opportunities for early obesity prevention. Proc Nutr Soc 67:317 –25.

[2]. Coppinger T, Jeanes YM, Dabinett J, Vögele C, Reeves S (2010) Physical activity and dietary intake of children aged 9-11 years and the influence of peers on these behaviours: a 1-year follow-up. Eur J Clin Nutr 64:776-81.

[3]. Kyttälä P, Erkkola M, Kronberg-Kippilä C, Tapanainen H, Veijola R, Simell O, et al (2010) Food consumption and nutrient intake in Finnish 1-6-year-old children. Public Health Nutr 13:947-56.

[4]. Manios Y, Grammatikaki E, Papoutsou S, Liarigkovinos T, Kondaki K, Moschonis G (2010) Nutrient intakes of toddlers and preschoolers in Greece: the GENESIS study. J Am Diet Assoc 108:357-61.

[5]. Huybrechts I, De Henauw S (2007) Energy and nutrient intakes by preschool children in Flanders-Belgium. Br J Nutr 98:600-10.

[6]. Cade J, Thompson R, Burley V, Warm D (2002) Development, validation and utilisation of food-frequency questionnaires - a review. Public Health Nutr 5:567-87.

[7]. Willet WC (1998) Nutritional Epidemiology. 2nd ed. New York: Oxford University Press.

[8]. Saland JM (2007) Update on the metabolic syndrome in children. Curr Opin Pediatr 19:183-91.

[9]. Raitakari OT, Juonala M, Kahonen M Taittonen L, Laitinen T, Mäki-Torkko N, et al (2003) Cardiovascular risk factors in childhood and carotid artery intima-media thickness in adulthood: the Cardiovascular Risk in Young Finns Study. JAMA 290:2277-83.

[10]. Subar AF (2004) Developing dietary assessment tools. J Am Diet Assoc 104:769-70.

[11]. Bel-Serrat S, Mouratidou T, Pala V, Huybrechts I, Börnhorst C, Fernández-Alvira JM, et al (2013) Relative validity of the Children's Eating Habits Questionnaire-food frequency section among young European children: the IDEFICS Study. Public Health Nutr 17:266-76.

[12]. Huybrechts I, De Backer G, De Bacquer D, Maes L, De Henauw S (2009) Relative validity and reproducibility of a Food-Frequency Questionnaire for Estimating Food Intakes among Flemish Preschoolers. Int. J. Environ Res Public Health 6:382-99.

[13]. Manios Y, Grammatikaki E, Androutsos O, Chinapaw MJ, Gibson EL, Buijs G, et al (2012) A systematic approach for the development of a kindergarten-based intervention for the prevention of obesity in preschool age children: the ToyBox-study. Obes Rev 13:S3-S12.

[14]. Androutsos O, Katsarou C, Payr A, Birnbaum J, Geyer C, Wildgruber A, et al (2014) Designing and implementing teachers' training sessions in a kindergarten-based, family-involved intervention to prevent obesity in early childhood. The ToyBox-study. Obes Rev 15:S48-S52.

[15]. Manios Y, Androutsos O, Katsarou C, Iotova V, Socha P, Geyer C, et al (2014) Designing and implementing a kindergarten-based, family-involved intervention to prevent obesity in early childhood: the ToyBox-study. Obes Rev 3:S5-S13.

[16]. Mouratidou T, Miguel ML, Androutsos O, Manios Y, De Bourdeaudhuij I, Cardon G, et al (2014) Tools, harmonization and standardization procedures of the impact and outcome evaluation indices obtained during a kindergarten-based, family involved intervention to prevent obesity in early childhood. The ToyBox-study. Obes Rev 3:S53-S60.

[17]. De Craemer M, De Decker E, De Bourdeaudhuij I, Vereecken C, Deforche B, Manios Y, Cardon G; ToyBox-study group (2012) Correlates of energy balance-related behaviours in preschool children: a systematic review. Obes Rev. 13 Suppl 1:13-28.

[18]. Mouratidou T, Ford FA, Fraser RB (2009) Reproducibility and validity of a food frequency questionnaire in assessing dietary intakes of low-income Caucasian postpartum women living in Sheffield, United Kingdom. Matern Child Nutr 7:128–139.

[19]. Lanfer A, Hebestreit A, Ahrens W, Krogh V, Sieri S, Lissner L, et al (2011) Reproducibility of the food frequency questionnaire section of the Children's Eating Habits Questionnaire used in the IDEFICS study. Int J Obes (Lond) 35:S61–S68.

[20]. Thompson F, Subar A (2008) Dietary assessment methodology. In Nutrition in the Prevention and Treatment of Disease, 2nd ed., pp. 3–39 [A Coulston and C Boushey, editors]. San Diego, CA, Elsevier Academic Press.

[21]. Livingstone MB, Robson PJ (2000) Measurement of dietary intake in children. Proc Nutr Soc 59: 279–293.

[22]. Mouratidou T, Mesana MI, Manios Y, Koletzko B, Chinapaw MJ, De Bourdeaudhuij I, et al (2012) Assessment tools of energy balance related behaviours used in European obesity prevention strategies: review of studies during preschool. Obes Rev 13: S42-S5