British Journal of Nutrition (2016), **116**, 1425–1436

doi:10.1017/S0007114516003317

Diet misreporting can be corrected: confirmation of the association between energy intake and fat-free mass in adolescents

Uku Vainik[1,2]*, Kenn Konstabel[1,3], Evelin Lätt[4], Jarek Mäestu[4], Priit Purge[4] and Jaak Jürimäe[4]

[1]*Institute of Psychology, University of Tartu, Näituse 2, 50410, Tartu, Estonia*

[2]*Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, 3801 University St., Montréal, QC, Canada, H3A 2B4*

[3]*Chronic Diseases Department, National Institute for Health Development, Hiiu 42, 11619, Tallinn, Estonia*

[4]*Faculty of Exercise and Sport Sciences, University of Tartu, Jakobi 5, 51014, Tartu, Estonia*

(Submitted 30 November 2015 – Final revision received 7 August 2016 – Accepted 18 August 2016 – First published online 11 October 2016)

Abstract

Subjective energy intake (sEI) is often misreported, providing unreliable estimates of energy consumed. Therefore, relating sEI data to health outcomes is difficult. Recently, Börnhorst *et al.* compared various methods to correct sEI-based energy intake estimates. They criticised approaches that categorise participants as under-reporters, plausible reporters and over-reporters based on the sEI:total energy expenditure (TEE) ratio, and thereafter use these categories as statistical covariates or exclusion criteria. Instead, they recommended using external predictors of sEI misreporting as statistical covariates. We sought to confirm and extend these findings. Using a sample of 190 adolescent boys (mean age = 14), we demonstrated that dual-energy X-ray absorptiometry-measured fat-free mass is strongly associated with objective energy intake data (onsite weighted breakfast), but the association with sEI (previous 3-d dietary interview) is weak. Comparing sEI with TEE revealed that sEI was mostly under-reported (74 %). Interestingly, statistically controlling for dietary reporting groups or restricting samples to plausible reporters created a stronger-than-expected association between fat-free mass and sEI. However, the association was an artifact caused by selection bias – that is, data re-sampling and simulations showed that these methods overestimated the effect size because fat-free mass was related to sEI both directly and indirectly via TEE. A more realistic association between sEI and fat-free mass was obtained when the model included common predictors of misreporting (e.g. BMI, restraint). To conclude, restricting sEI data only to plausible reporters can cause selection bias and inflated associations in later analyses. Therefore, we further support statistically correcting sEI data in nutritional analyses. The script for running simulations is provided.

Key words: **Under-reporting: Plausible reporting: Subjective energy intake: Objective energy intake: Dietary interviews: Selection bias**

Given the global increase in obesity[1], considerable effort has gone into determining the predictors of energy intake. These predictors include fat-free mass[2], psychological self-control and food drive[3,4], socio-economic status[5,6] and various environmental features[7,8]. All these research fields depend on the crucial assumption that energy intake is correctly measured. For accuracy reasons, these studies often expend effort to objectively measure energy intake in the laboratory or find other methods of indirect energy expenditure assessment[5]. Although collecting such subjective energy intake (sEI) with questionnaires is considerably easier and cheaper, such data tend to be misreported, and are therefore often considered unreliable.

Misreporting can be observed when calculating the energy balance percentage (EB%) – that is, how well does energy intake match with total energy expenditure (TEE)? When calculating the EB% for sEI (sEI/TEE × 100), many findings show that sEI tends to be under-reported in adults[9] and also in children[10]. This phenomenon has been clearly established using very large data sets[11,12]. Under-reporting is particularly prevalent in adolescents, with 14–52 % of sample under-reporting[10,13]. The EB% can be predicted from several external variables, ranging from simple BMI[10] and demographic factors[12] to brain activation to food stimuli[14]. Studies in adults add additional factors such as dietary restraint and social desirability ([15], reviewed by Macdiarmid & Blundell[16]). However, very few studies have focused on providing practical advice on how to handle inaccurate sEI data[17–19], and only one previous study has focused on this question in children[20].

Abbreviations: EB%, energy balance percentage; oEI, objective energy intake; PR, plausible reporters; sEI, subjective energy intake; TEE, total energy expenditure; UR, under-reporters.

* **Corresponding author:** U. Vainik, +372 737 6549, email uku.vainik@gmail.com

The study by Börnhorst *et al.*[20] explored various approaches to recover a missing association between obesity and sEI (dietary recall) data in children. They first divided subjects into three groups of diet reporting accuracy – under-reporters (UR), plausible reporters (PR) and over-reporters (OR) – on the basis of discrepancy between energy expenditure and energy intake. Next, they tested various recovery approaches such as restricting analysis only to PR groups, stratifying analysis by reporting group or controlling for co-predictors of misreporting. They concluded that the best approach for recovering an association between BMI and energy intake is to control for predictors of misreporting, rather than excluding misreporting groups[20].

What was not evident in the study by Börnhorst *et al.*[20] is that excluding misreporting groups can generate artificial positive bias in later analyses. For instance, Mendez *et al.*[19] compared various methods for restricting adult sEI data to PR and then related that restricted sEI data to BMI. They concluded that some methods generate a higher effect size between sEI and BMI than others, recommending the ones with higher effect size. Rhee *et al.*[21] recently re-analysed that data and suggested that the effect size increase likely occurs because of selection bias. Selection bias is well known in the field of epidemiology but can be hard to detect[22–24]. Selection bias happens when the dependent variable is conditioned on a variable that partly relates to the independent variable. Using the case of Mendez *et al.*[19] and Rhee *et al.*[21] as an example, the dependent variable (sEI) is restricted to PR using a formula that depends on body mass, and then sEI is related to BMI, which also depends on body mass. Rhee *et al.*[21] demonstrated that when there was no selection bias, focusing on PR was reasonable. An example would be associating sEI with variables that do not relate to body mass, such as prospective BMI change or various biomarkers[21]. In the follow-up discussion, Mendez[25] acknowledged that not all methods for correcting sEI work as expected, and both Mendez[25] and Rhee & Willet[26] underlined that a better understanding of misreporting correction methods is needed.

The present study seeked to integrate evidence from adults[21] and children[20] to study the optimal method of handling sEI data in adolescents. The main goal was to replicate the exploratory results of Börnhorst *et al.*[20] and extend them in several ways. First, as the effects between BMI and calories consumed are small, the current study focused on a relatively new finding that the amount of calories consumed is determined by fat-free mass[27–29]. This association has been tested previously, as it is known to have medium effect size (β= 0·28–0·42), with objective measures of energy intake[27,29]. The other extensions compared with Börnhorst *et al.*[20] include the additional measure of objective energy intake (oEI) that enables to independently verify the association between fat-free mass and energy intake, a different type of source for sEI data (3-d dietary interview) that tests for generalisibility of the findings, more accurate estimate of TEE by including 7-d accelerometry in the model and extending the set of predictors for misreporting to several psychological predictors. In summary, the current study aimed to test the feasibility, conceptually replicate and extend the statistical approach suggested by Börnhorst *et al.*[20]. As studies on adults have suggested the emergence of selection bias[21], we scrutinised the results from that perspective.

Methods

Study population

The present study analysed data from the fourth wave of a larger project 'Risk factors for metabolic syndrome in boys during pubertal development: a longitudinal study with special attention to physical activity and fitness'[30–39]. The study was originally started in 2009, where all boys from Grades 3 and 4 from twenty-seven elementary schools in the city and the surroundings of Tartu, Estonia, were invited to participate. All schools were in an urban environment. A total of 313 boys, approximately 84 %, agreed to participate. All participants had no disease that prevented them from taking part in different parts of the study and were allowed to take part in the obligatory physical education classes at school (they had no health-related problems, injuries, etc.)[34]. The measurement period of the currently analysed wave was from November 2012 until April 2013.

A total of 190 participants provided accelerometer data and dietary interview (sEI) and body anthropometry data (age: $\bar{x}$ = 13·99 (SD 0·69), zBMI: $\bar{x}$ = 0·37 (SD 1·29), BMI: $\bar{x}$ = 20·91 (SD 4·66)). Responses from psychological questionnaires were available from 128 participants and oEI data from thirty-nine participants. Participants in the subsamples did not differ from participants not within a certain subsample (questionnaire *v.* no questionnaire; oEI *v.* no oEI) in terms of age (t(146·4) = −0·19, P = 0·851; t(54·4) = −1·28, P = 0·205) or zBMI (t(106·8) = 0·97, P = 0·334; t(66·8) = 0·21, P = 0·832), suggesting that the subsamples were a random part of the bigger sample.

Participants recorded their 72-h sEI for 3 or more days before onsite testing. Participants were asked to abstain from breakfast before coming to the test site at approximately 08.00 hours. As participants had to be picked up from school, sometimes kilometres away from the testing site, they arrived in groups of four to eight. The overall study design is presented in Table 1.

All participants completed various questionnaires about their health habits. A subset of the sample also completed various questionnaires related to personality and eating behaviours. All participants had to eat breakfast at the spot; in a subset of participants, their intake was also weighted.

This study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human subjects/patients were approved by the Medical Ethics Committee of the University of Tartu. All children and parents were thoroughly informed of the purposes and contents of the study, written informed consent was obtained from the parents before participation, and the children provided their verbal assent.

Anthropometry

The participants' body height and mass were obtained on the 1st day of the measurements. Body height was measured in

https://doi.org/10.1017/S0007114516003317 Published online by Cambridge University Press

Table 1. Summary timeline of the study

Time	Activity
3 + days before onsite testing	• Participants recorded the subjective energy intake with help of parents by filling the 72-h food diary*
08.00 hours on the day of onsite testing	• 4–8 participants arriving in a group. Providing participants with instructions and questionnaires
	• Participants were asked about the morning fast
08.15–09.30 hours	• Measurement of perceived hunger on a visual analogue scale
	• Venous blood sampling
	• After blood sampling, subjects were given the opportunity to snack (oEI)*
	• Completion of physical activity questionnaires
	• After snacking, participants started to fill in questionnaires on personality and eating disorders* The questionnaire set included a measurement of perceived hunger on a visual analogue scale
	• Participants completed measurements in this order
09.30–12.00 hours	• Measurement of DXA, weight and height*
	• Dietary interview for measuring subjective energy intake*
	• Tanner estimation*
	• Measurement of VO_{2max}
	• Participants completed the anthropometry and dietary interview in no particular order, but VO_{2max} test was always the last one
7 d after onsite testing	• Wearing the accelerometer*

* For measurements included in the current study, see the 'Anthropometry', 'Questionnaires' and 'Dietary data' sections for details. No higher timing precision was possible, as boys participated in onsite testing session in groups because of the way they were transported onsite.

standing position to the nearest 0·1 cm using a Martin metal anthropometer. Body mass was measured with minimal clothing using a medical balance scale (A&D Instruments) to the nearest 0·05 kg. BMI (kg/m^2) was calculated as body mass divided by the square of body height. BMI is shown for informative purposes; the analysis was conducted with zBMI that corrects for developmental effects, created using World Health Organization scripts[40]. The boys' weight status was categorised according to zBMI cut-off values. The biological age of the participants was assessed according to a self-assessed illustrative questionnaire of the pubertal stage according to the Tanner classification method by evaluation of pubic hair[41].

Body composition: fat mass and fat-free mass were measured using dual-energy X-ray absorptiometry (DXA; DPX-IQ densitometer, Lunar Corporation) equipped with proprietary software (version 3.6). Boys were scanned in the supine position wearing light clothing. The medium scan mode and the standard subject positioning was used for total body measurements, which were analysed using the extended analysis option. To reduce the impact of the operator variability factor, one qualified observer analysed all scans over the 2-year period. The CV for these body composition measurements were <2 %; this was established in our laboratory using duplicate measures in twenty boys of the same age. Fat mass and fat-free mass correlated at 0·24, $P = 0·001$.

Objectively measured physical activity was assessed using an accelerometer (GT1M ActiGraph) that was worn for 7 d on the right hip. The accelerometer was programmed to record activity counts in 15-s epochs, and non-wearing time was defined as ≥20 consecutive minutes of zero counts and was not included in the analysis. Data from the accelerometer were included for further analysis if the subject had accumulated a minimum 8 h of activity data/d, for at least 1 weekend day and 2 weekdays. In the final sample, the median number of valid weekend days was 2, and the median number of valid weekdays was 5. Other details of the accelerometer procedure and data processing have been described elsewhere[35,36]. In the present study, we used counts per min as an indicator of total physical activity. In a previous study, a similar analytic accelerometer approach together with body mass was able to predict doubly labelled water-based TEE with R^2 of 0·82, SE of estimate 0·49, prediction error 0·28[42] (Table 3).

Questionnaires

We included several questionnaires that we suspected could influence EB%[15].

The Eating Disorders Assessment Scale (EDAS[43]) is a twenty-nine-item, self-report questionnaire with four subscales: restrained eating, binge eating, purging and preoccupation with body image and body weight. These subscales show good internal consistency and discriminant validity. The construct validity of the questionnaire has been confirmed by strong correlations with Eating Disorders Inventory – 2 Estonian version[44]. In the current analysis, we used the binge eating subscale and restraint subscale, as these are the two main eating behaviour dimensions[3,45], and the current instrument assesses these behaviours in a continuous manner. The binge eating subscale is very similar to other known measures of loss of control over food[45,46], and loss of control over food is hypothesised to partly reflect reward sensitivity to food[45,47].

Social desirability was estimated from responses to Estonian Brief Big Five Inventory[48]. This is a brief measure of personality based on the example of 'Common Language' California Child Q-Set[49]. The scale assesses basic personality dimensions (neuroticism, extraversion, openness, agreeableness and conscientiousness) with eight items each on a five-point Likert-type scale, and has been previously validated in an adolescent sample[48]. In the current analysis, we used previously measured social desirability scores of the items (unpublished data) to calculate a general tendency for responding in a socially desirable manner, ranging from −1 to 1. The methodology is described elsewhere[50].

Dietary data

sEI data were self-recorded. Before study start, participants were asked to record everything they ate during 2 weekdays and 1 weekend with the help of their parents. Participants were asked to observe their food intake as closely as possible before the testing day. During the testing day, participants brought their written summary of the 3 d, based on which they were interviewed by a trained nutritionist. The nutritionist helped in recalling possible forgotten energy items and entered the energy items into an energy database that automatically calculated relevant energy[51]. From that we estimated their average energy intake (MJ) per day as an indicator of sEI.

oEI data were measured once in a subset of the sample during morning snacking on the day of testing. The main goal of the snacking was to provide participants with an opportunity to recover from morning fast before various other measurements were obtained. In the current analysis, the oEI data provided an opportunity to verify independently that energy intake is related to fat-free mass. Clearly, the oEI meal was not the same as the previous meals, based on which sEI data were reported. At the same time, previous evidence has shown that the association between fat-free mass and energy intake is robust – it is present both for individual meals[29] and for energy intake aggregated across a full day[27]. Therefore, this single meal data were used to verify the association between fat-free mass and energy intake in this sample. Because of the study design, offering wide range of foods was not feasible; the participants were provided the following easy-to-handle foods: a Mars bar (1·89 MJ, 451 kcal) or Snickers bar (2·13 MJ, 509 kcal), a pack of cookies (1·81 MJ, 432 kcal) and a 0·5-litre bottle of juice (0·17 MJ, 41 kcal). After the participants had stopped eating, they were asked to leave the remaining food on the table. The remaining food was weighted using a Soehnle Attraction kitchen scale (Leifheit AG), with 1 g precision, and weight was converted to energy on the basis of nutritional information on the packaging. Number of total MJ consumed was the indicator of oEI.

Statistical methods

TEE (MJ) was estimated from weight and accelerometer data. The estimators were obtained from a validation analysis where body mass and accelerometer data could explain 81–82 % of TEE expenditure in children, estimated with doubly labelled water[42]. Although the validation sample was younger than the current sample, we are unaware of other validation studies that would provide estimations more suitable for the current sample. The formula is provided by the second author; it was inadvertently not published in the original article[42]:

$$\mathrm{TEE} = 0{\cdot}722 + \mathrm{weight} \times 0{\cdot}160 + 0{\cdot}003 \times \mathrm{counts/min}.$$

We compared the formula-based TEE with TEE derived from equations developed by Brooks et al.[52] that were based on age, weight, height and physical activity level. Physical activity level was converted from estimates of moderate-to-vigorous physical activity[13] based on accelerometry data. The two TEE estimates correlated very highly (r 0·97).

To detect UR and OR, energy balance percentage (EB%) was derived from the formula sEI/TEE × 100. A common method to detect misreporting is to classify participants as UR or OR if they deviate more than ±1 SD from 100 %. The particular SD values are derived from a formula that accounted for intra-individual variation in energy intake (CV adjusted for age), day-to-day variation (here 3 d) and energy requirement predictor errors[53]. We used the approach of Noel et al.[13] who provided updated CV for boys <14 and ≥14. As a result, the cut-off values for younger and older age groups for under-reporting were 85·675 and 85·798 % and the values for over-reporting were 114·325 and 114·202 %, respectively.

We first tested whether fat-free mass would predict oEI, in attempt to replicate previous findings[27,29] using linear regression, correcting for age. Next, we used a similar regression model to test whether fat-free mass would predict sEI. Thereafter, we tried various correction methods such as excluding UR and OR, controlling for dietary group status in the regression analysis and adding predictors of EB% to the regression. Predictors of EB% were chosen among variables suggested by previous studies (see first paragraph for an overview). These predictors included BMI, psychological traits such as restraint, binge eating and responding to a personality questionnaire in a socially desirable manner.

In the last model, we used multiple imputation to overcome the issue that anthropological data were available for the full sample (n 190) but psychological predictors were available only for a subset of the sample (n 128). When these predictors are used together in a regular multiple regression model, the models would use list-wise data deletion, which would have considerably reduced the statistical power of anthropological measures. Multiple imputation[54], in turn, creates multiple versions of the data set. In each data set, missing values are drawn from a plausible distribution. Each of the imputed data sets was analysed separately, and then the results were aggregated. In this case, we created 100 imputed versions of the data set using Amelia package[55]. These data sets were analysed and aggregated with the mice package[56], relying on small-sample method to calculate aggregate df[57].

As can be seen in the results, we recovered an unexpectedly strong association between fat-free mass and sEI when focusing only on PR (e.g. $\beta = 0{\cdot}77$, model 2 in Table 4). This suggests the emergence of selection bias. To test for selection bias, we re-sampled sEI data – every participant randomly received another participant's sEI value. Different reporting groups (UR, PR, OR) were re-identified using the same method as mentioned above. Thereafter, we re-ran previously tested regression analyses. As re-sampled data are equivalent of random noise, no variable should be able to predict re-sampled data. However, if some variables are able to predict the re-sampled data, the prediction can be considered to be an artifact arising from correction methods or selection bias.

To explore how selection bias can influence the results, we simulated the study data 10 000 times to demonstrate a robust replication of the artifact. We further explored the extent of selection bias by varying the association strength between variables in the simulation. The code used to simulate the data is provided in the online Supplementary Material.

All analyses were conducted in R environment 3.2.3[58], occasionally relying on 'plyr', 'plotrix', 'truncnorm', 'MASS', 'Amelia' and 'mice' packages[55,56,59–62], as well as online resources[63]. Variables that displayed non-normality based on the Shapiro–Wilk test and observing histograms were transformed to log scale. To avoid values taking log of 0, +1 was added to all EDAS scores when represented in the log form.

Regression diagnostics were first conducted by scrutinising the residuals for normality, homoscedasticity and linearity. No visual violations were found. Thereafter, we analysed whether any model would have standardised residuals higher than values usually expected based on typically used criteria. For instance, <5 % of observations should have standardised residuals above 1·95. Similarly, <1 % of observations should have standardised residuals >2·58, and <0·1 % of observations should have standardised residuals >3·29[64]. Occasionally, some models were borderline (e.g. 5·3 % of observations had standardised residuals above 1·96). These borderline models were inspected further with visual analysis[65]. Visual analysis was based on Cook's distance plots that were inspected for potential outliers – that is, we looked for data points that would have significantly higher Cook's distance than other variables. Only in one analysis, such an outlier was found (see the 'Associations between fat-free mass and energy intake' section). However, as removing that outlier did not change the general model, all data points were retained. In the multiple imputation analysis, five randomly drawn regression analyses from the 100 analyses conducted were inspected for outliers.

Results

Descriptive variables

Plotting of EB% data revealed that under-reporting was widespread – 74·2 % of the participants under-report their sEI (Fig. 1). Table 2 summarises various descriptive statistics for the whole sample, as well as for each subgroup. Expectedly, the reporting groups differed in sEI. Compared with the median intake of PR, the median intake of under-reporters was 67 % and the median intake of OR was 127 %. At the same time, the groups had no

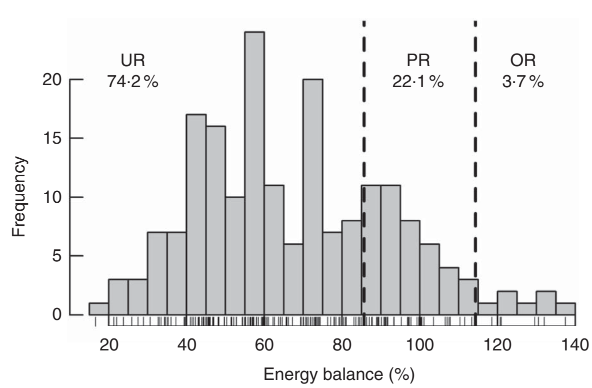

Fig. 1. Histogram of different energy balance percentages. PR, plausible report; OR, over-report; UR, under-report. - - - -, Cut-off values of the younger group (see the 'Statistical methods' section for details). Tick marks (|) represent actual values, jittered with a factor of 1. When TEE was estimated with the Brooks *et al.* method[52], the diet group prevalence percentages were as follows: UR = 80·5 %, PR = 14·7 % and OR = 4·7 %.

Table 2. Descriptive analyses of variables stratified by reporting group and differences between the reporting groups tested with ANOVA or the Kruskal–Wallis rank sum test* (Means and standard deviations)

	TOT (*n* 190)		UR (*n* 141)		PR (*n* 42)		OR (*n* 7)		Difference
	$\bar{x}$ (Md)	SD (min–max)	$\bar{x}$ (Md)	SD (min–max)	$\bar{x}$ (Md)	SD (min–max)	$\bar{x}$ (Md)	SD (min–max)	
sEI (MJ)†	7·55	2·6–13·81	6·62	2·6–12·33	9·91	6·53–13·81	12·61	10·52–13·37	$F_{2,2}$ = 64·95, P < 0·001
oEI (MJ)‡	2·9	1·21	3·06	1·25	2·51	1·2	2·78	0·52	$F_{2,2}$ = 0·74, P = 0·483
Age (years)	13·99	0·69	14·01	0·67	13·92	0·81	14·02	0·53	$F_{2,2}$ = 0·27, P = 0·76
BMI†	19·61	13·77–45·51	20·55	14·6–45·51	18·13	13·77–26·37	16·44	15·8–19·2	$F_{2,2}$ = 14·11, P < 0·001
zBMI	0·37	1·29	0·66	1·25	−0·37	1·06	−0·99	0·76	$F_{2,2}$ = 16·64, P < 0·001
Percentage underweight/normal weight/overweight/obese (%)	3/70/12/14		7/86/7/0		2/64/14/19		0/100/0/0		$H(2)$ = 7·74, P = 0·021
Fat-free mass (kg)†	43·65	26·95–74·28	44·74	27·27–74·28	41·09	26·95–57·59	40·74	36·81–43·17	$F_{2,2}$ = 2·91, P = 0·057
Fat-free mass index (kg/m²)	15·34	1·96	15·58	2	14·74	1·77	14·05	0·7	$F_{2,2}$ = 4·75, P = 0·01
Fat mass (kg)†	9·13	2·68–73·46	10·49	3·53–73·46	6·97	2·68–23·58	6·25	3·91–10·33	$F_{2,2}$ = 11·77, P < 0·001
Fat mass index (kg/m²)†	3·34	1–21·93	3·82	1·19–21·93	2·63	1–9·44	2·1	1·24–3·94	$F_{2,2}$ = 11·18, P < 0·001
Percentage in Tanner stages 1/2/3/4/5 (%)	4/22/40/33/1		2/31/29/38/0		4/20/43/33/1		14/29/43/14/0		$H(2)$ = 2·04, P = 0·361
Physical activity (counts/min)†	388·4	113·3–1010·3	388·29	134·91–1010·3	397·8	167·66–779·19	330·44	113·3–525·2	$F_{2,2}$ = 2·35, P = 0·098
EDAS: restraint§†	9	0–25	10	0–25	6	0–17	7	0–21	$F_{2,2}$ = 4·47, P = 0·013
EDAS: binge eating§†	6	0–33	6	0–33	7	0–21	5	3–14	$F_{2,2}$ = 0·26, P = 0·774
Social desirability§	0·17	0·08	0·17	0·08	0·17	0·09	0·17	0·07	$F_{2,2}$ = 0·03, P = 0·966

TOT, total sample; UR, under-report; PR, plausible report; OR, over-report; $\bar{x}$, mean; Md, median; sEI, subjective energy intake; oEI, objective energy intake; EDAS, Eating Disorders Assessment Scale.
* The Kruskal–Wallis rank sum test was applied on counts of participants in body weight category or Tanner stage.
† Non-normal variables. The median and range are reported in parentheses. In addition, these variables were log-transformed during ANOVA testing to obtain a distribution closer to normality.
‡ Reduced sample: TOT (*n* 39), UR (*n* 27), PR (*n* 10), OR (*n* 2).
§ Reduced sample: TOT (*n* 128), UR (*n* 91), PR (*n* 30), OR (*n* 7).

difference in oEI, suggesting that the group differences in sEI were due to the EI measurement method. Regarding physiological variables, the groups differed in terms of BMI, zBMI, fat mass, fat mass index, fat-free mass and fat-free mass index. From psychological measures, the only difference was observed in restraint. However, restraint correlated with zBMI (r 0·42, $P<0{\cdot}001$) and fat mass (r 0·42, $P<0{\cdot}001$). As many variables displayed non-normality, their log-transformation values have been used in all reported correlation and regression analyses.

Associations between fat-free mass and energy intake

The results demonstrated that participants indeed chose the amount of food based on their fat-free mass, as suggested by Blundell[27]. The association was clear for oEI but was considerably weaker for sEI (Table 2, Fig. 2, online Supplementary Fig. S1). Given that sEI was mostly under-reported (Fig. 1), the current results highlighted the need for a method for correcting varying EB% (Table 3).

Methods that adjust for misreporting

In Table 4, the left column summarises the results for different methods. As expected from Börnhorst *et al.*[20], the plain model (model 1) had the poorest explanatory power and R^2, and the model adjusting for predictors of EB% (model 4) was considerably better than model 1 by restoring the beta value closer to what was expected from oEI data and from previous studies. Intriguingly, excluding under- and OR (model 2) or controlling for dietary groups (model 3) seemed to provide even better results. Both models had very high R^2, and model 2 had very high standardised β. Do these results imply that these methods are even better?

To test for potential method artifacts, we used the re-sampling procedure – every participant received a random sEI value of another participant. We expected that all models (Table 4, right column) should have non-significant results, as fat-free mass was predicting essentially noise. Indeed, the simple model (model 1) and the model adjusting for predictors of EB% (model 4) had R^2 <0·01 (Fig. 3, dashed line; Table 4, right column). However, the models based on dietary groups created from re-sampled sEI (models 2–3) still showed a strong effect (Fig. 3, dashed line; Table 4, right column). This suggests that group-based methods are unsuitable for the current purposes; there was an association between sEI and fat-free mass, even though the groups were created based on re-sampled sEI data and both diet reporting groups and sEI itself should not be informative. The scatter plots of models 1 and 2 for actual and simulated data are shown in the online Supplementary Fig. S2. Together, these results suggest that grouping methods are unsuitable for recovering the association between sEI and fat-free mass.

Possible causes of method artifacts

A possible reason as to why the sEI and fat-free mass association appears in models 2 and 3 is selection bias; fat-free mass

Table 3. Regression coefficients of fat-free mass predicting objective energy intake or subjective energy intake, accounting for participant's age

	β	B	SE	t	P
Objective energy intake*					
(Intercept)		−8·26	3·39	−2·43	0·02
Fat-free mass†	0·55	3·1	1·05	2·95	0·006
Age	−0·02	−0·04	0·3	−0·13	0·896
Adjusted R^2	0·25				
df	3, 36				
Subjective energy intake†					
(Intercept)		1·14	0·47	2·44	0·016
Fat-free mass†	0·17	0·23	0·11	2·04	0·043
Age	0·02	0·01	0·03	0·23	0·82
Adjusted R^2	0·02				
df	3, 187				

* Regression diagnostics found one potentially influential outlier in this model (Cook's distance = 0·29, standardised residual = 2·09). Without that outlier, the effect of fat-free mass would be β = 0·64, B = 3·58, SE = 0·91, t = 3·91.
† These variables have been log-transformed to normalise distributions.

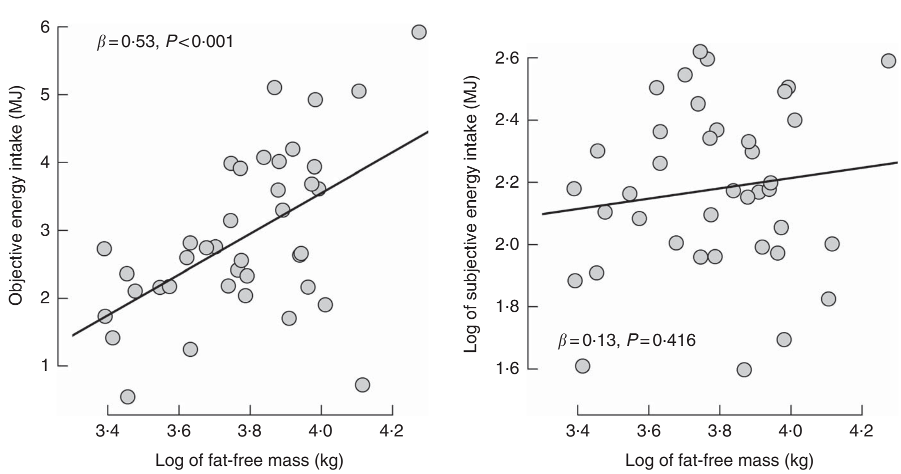

Fig. 2. Fat-free mass associations with objective (left) and subjective (right) energy intake. Objective energy intake was measured on the same day, whereas subjective energy intake was assessed from dietary interview from an earlier period of 3 d. Data not corrected for the effects of age. For illustrative purposes, variables here are not log-transformed. For log-transformed plots, see the online Supplementary Fig. S1.

Table 4. Fat-free mass predicting subjective energy intake across different approaches that adjust for misreporting*

	Subjective energy intake						Re-sampled subjective energy intake					
	β	*B*	SE	*t*	df	*P*	β	*B*	SE	*t*	df	*P*
Model 1: simple model												
(Intercept)	<0·01	1·14	0·47	2·44	187	0·016	<0·01	2·21	0·54	4·06	187	<0·001
Fat-free mass†	0·17	0·23	0·11	2·04	187	0·043	0·03	0·05	0·13	0·37	187	0·715
Age	0·02	0·01	0·03	0·23	187	0·82	−0·06	−0·03	0·04	−0·76	187	0·448
Adjusted R^2	0·02						−0·01					
Model 2: exclusion of other groups												
(Intercept)	<0·01	0·19	0·29	0·65	39	0·518	<0·01	−0·07	0·48	−0·15	35	0·883
Fat-free mass†	0·77	0·58	0·12	4·89	39	<0·001	0·69	0·62	0·12	4·95	35	<0·001
Age	0·02	<0·01	0·03	0·13	39	0·901	0·03	0·01	0·04	0·2	35	0·842
Adjusted R^2	0·59						0·46					
Model 3: adjusting for group												
(Intercept)	1·05	0·92	0·34	2·71	185	0·007	1·65	1·11	0·45	2·47	185	0·014
Fat-free mass†	0·3	0·42	0·08	5·03	185	<0·001	0·26	0·42	0·11	3·82	185	<0·001
Age	−0·02	−0·01	0·02	−0·26	185	0·793	−0·03	−0·01	0·03	−0·42	185	0·673
Group = UR		−0·42	0·04	−11·33	185	<0·001		−0·19	0·11	−1·73	185	0·086
Group = OR		0·2	0·08	2·4	185	0·017		−0·67	0·11	−6·26	185	<0·001
Adjusted R^2	0·48						0·36					
Model 4: adjusting for predictors												
(Intercept)	0·01	0·91	0·49	1·88	168·9	0·062	0	2·44	0·59	4·13	175·4	<0·001
Fat-free mass†	0·35	0·49	0·14	3·47	166·2	0·001	−0·03	−0·05	0·17	−0·31	172	0·759
Age	−0·09	−0·04	0·03	−1·04	173·3	0·299	−0·03	−0·01	0·04	−0·29	177·1	0·774
zBMI	−0·27	−0·06	0·02	−2·7	148·9	0·008	0·14	0·04	0·03	1·34	164·7	0·182
EDAS: restraint†	−0·15	−0·05	0·03	−1·51	88·6	0·135	−0·04	−0·02	0·04	−0·41	123·7	0·682
EDAS: binge eating†	0·07	0·03	0·04	0·73	84·3	0·466	−0·08	−0·04	0·04	−0·81	99·5	0·42
Social desirability	−0·08	−0·26	0·34	−0·76	84·3	0·45	−0·01	−0·05	0·36	−0·13	125·8	0·896
Pooled adjusted R^2	0·13						N/A					

UR, under-report; OR, over-report; EDAS, Eating Disorders Assessment Scale.

* The reference group in the 'adjusting for group' model provided plausible reports. Data were re-sampled by assigning each participant an energy intake value of another participant. Model 4 is based on the multiple imputation procedure (see the 'Statistical methods' section for details).

† These variables have been log-transformed to normalise distribution.

https://doi.org/10.1017/S0007114516003317 Published online by Cambridge University Press

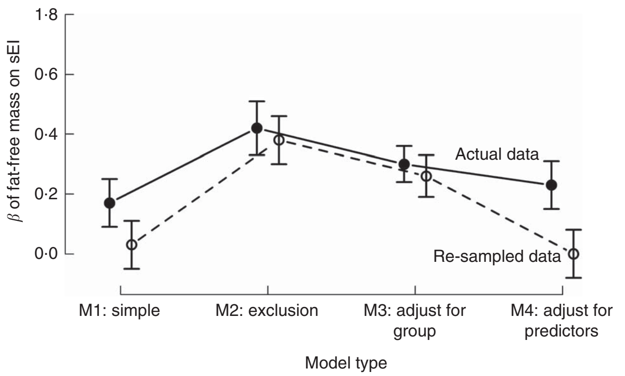

Fig. 3. Graphical comparison of the standardised regression coefficients (β) of actual data (——) and re-sampled data (- - - -). Initial observation of β in the actual data suggests that the association between fat-free mass and subjective energy intake (sEI) can be best recovered with data exclusion or group adjustment strategies (models 2 and 3). However, these strategies also show an effect in re-sampled data, where no effect should be present. Models 1 and 4 correctly show no effect in re-sampled data. Therefore, adjusting for predictors (model 4) is the most viable approach when recovering an association between fat-free mass and sEI. Re-sampled data were obtained by assigning each participant an energy intake value of another participant. Errors bars denote standardised standard errors, obtained by standardising the variables and re-computing the regressions in Table 4. sEI was assessed from a dietary interview from an earlier period of 3 d.

Fig. 4. A summary of the direct and indirect pathways of how fat-free mass is associated with subjective energy intake (sEI). The direct effect between fat-free mass and sEI is of main interest. However, when sEI data are restricted to plausible reporters (PR), then this creates a selection bias – a secondary association between fat-free mass and sEI because fat-free mass is part of body mass, which defines total energy expenditure (TEE). TEE, in turn, defines the PR group. When the analysis does not account for the indirect pathway, the effect estimate of the direct pathway gets amplified.

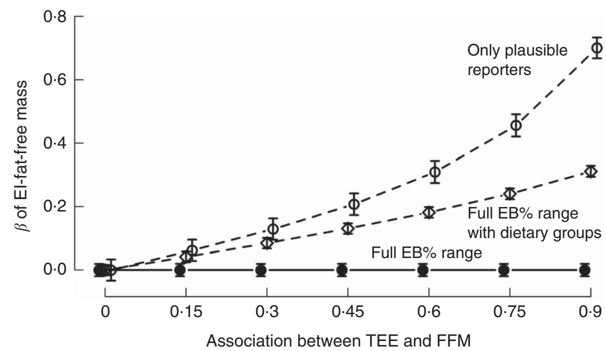

Fig. 5. The effect of restricting variance based on a partly related variable on standardised regression using simulated data. The expected standardised association (β) between fat-free mass (FFM) and energy intake (EI) is zero, and full sample data show this (—— with ●). - - - -, Same analysis in case the analysis focused only on plausible reporters (- - - - with ○, as in model 2) or in case a variable with dietary groups was added as a covariate (- - - - with ◇, as in model 3). In the latter two cases, the artificial association varied as a function of the associated strength between total energy expenditure (TEE) and fat-free mass. Data simulated on 10 000 participants, 10 000 times. Variables have similar properties as actual data in terms of distribution. For precise parameters, see the online Supplementary Material. Error bars denote 95 % CI (standard errors multiplied by 1·96). See the online Supplementary Fig. S3 for non-standardised regressions and the online Supplementary Material for R script used to generate the data. EB%, energy balance percentage.

relates to the variables used to create the diet reporting accuracy groups. Namely, the diet reporting groups are generated based on EB% – that is, the sEI:TEE ratio. In the current sample, TEE was highly dependent on participants' body mass (see formula in the 'Statistical methods' section, r 0·98), and body mass correlates highly with fat-free mass (r 0·72, $P<0·001$). Indeed, fat-free mass correlates with TEE (r 0·71, $P<0·001$). Therefore, restricting re-sampled sEI data to a PR group that has an EB% from 85 to 115 could create an association between re-sampled sEI and fat-free mass (models 2 and 3 in Fig. 3). In contrast, when all re-sampled sEI data were used in the analysis, the re-sampled sEI and fat-free mass were not related (models 1 and 4 in Fig. 3).

The artificial emergence of the association between sEI and fat-free mass is illustrated in Fig. 4. Although we were interested in the direct relationship between fat-free mass and sEI (upper pathway), restricting sEI to PR created an indirect association between fat-free mass and sEI through body mass/TEE (lower pathway). This indirect association occurred because fat-free mass correlates with body mass, and body mass-based TEE defined the PR group in sEI. Even when whole sample is used, but the dietary restriction groups are used as covariates, one can still observe a similar bias (model 3 in Table 4 and Fig. 3, Fig. 5).

To demonstrate the causal role of this sample restriction pathway, we created a simulation where we varied the correlation between fat-free mass and TEE. Namely, we created normally distributed variables with similar properties as actual data on 10 000 people. This included TEE and fat-free mass that correlated between 0 and 0·90 and EI, which in simulation did not correlate with any variable. We skipped body mass in the simulation, as it had high correlation with TEE (r 0·98). Thereafter, we calculated EB% (EI/TEE) and the association between fat-free mass and EI using full data and by creating dietary reporting groups, which were used for restricting the sample to PR (like model 2) or using dietary reporting groups as covariates (like model 3). Under-reporting was defined as EB% <85 %, plausible reporting was defined as EB% ranging from 85 to 115, and over-reporting was defined as EB% >115. The whole procedure was repeated 10 000 times to test for robustness.

As can be seen in Fig. 5, the association between fat-free mass and EI is absent with complete data (solid line, filled circles). At the same time, restricting data to PR (dashed line, empty circles), or using dietary group variables as control variables (dashed line, empty diamonds), created an artificial association between EI and fat-free mass. The association magnitude depended on the correlation between fat-free mass and TEE, and as that correlation decreased the artificial association between EI and fat-free mass decreased. Nevertheless, the artificial association was present even at the smallest non-zero correlation (r 0·15). This demonstrates that the dietary group variable approaches would be appropriate only when the fat-free mass and TEE correlation is zero. See the online

Supplementary Fig. S3 for non-standardised regressions. Scripts used for simulations are available in the online Supplementary Material.

Discussion

We compared three approaches to control for misreporting of sEI in adolescent boys – exclusion of misreporting groups, controlling for misreporting group status and statistically correcting for misreporting using external predictors of EB%. Our analysis confirmed the exploratory conclusion of Börnhorst *et al.*[20] that if children's energy intake based on sEI data is related to other variables the sEI should be statistically corrected for misreporting using separate predictors of EB%. Such statistical correction recovered an association between fat-free mass and sEI, an association that was expected from both previous studies and oEI data. We further demonstrated the dangers of other approaches that exclude misreporting groups or statistically control for dietary group status; in our analysis, these approaches created selection bias that artificially boosted the expected association between fat-free mass and sEI.

Although exclusion of UR and OR has been suggested previously as a useful technique[17,19], our data suggest that plausibility of dietary interviews does not make the data more correct. Instead, focusing on plausible data might produce artifacts – creating an artificially strong association between sEI and fat-free mass. A likely reason is selection bias – fat-free mass relates to the body mass of a participant, and body mass is used to estimate TEE, on which the dietary groups are based (see formula in the 'Statistical methods' section, Fig. 4). If then participant range is restricted to a narrow range of plausible sEI values based on the sEI:TEE ratio, an artificial association emerges between sEI and fat-free mass. Our simulations showed that a detectable contamination is present even when the fat-free mass would relate to TEE only at 0·15. Similar to Rhee *et al.*[21], our simulation showed that using a narrow range of PR is only reasonable when the predictor of sEI has a correlation of zero with TEE. However, this zero correlation can be difficult to achieve, as TEE has multiple components (BMR, body mass and physical activity), and many physiological variables tend to be related. Therefore, the zero correlation between TEE and predictor of sEI has to be demonstrated before dietary groups-based correction methods are used. To be on the safe side, we suggest using external predictors of EB% to correct sEI instead of approaches based on dietary groups.

The current results also highlight the usefulness of data re-sampling. Selection bias or any other bias can be difficult to detect, because understanding indirect associations between physiological variables can be a complex task. Data re-sampling provides a quick and simple test to check, whether the used data correction mechanism has created artifacts – an association that is different from zero. If a correction procedure creates an association between two variables, which should have zero association as one of them is random noise, then this correction mechanism should not be used.

Simulation provides further opportunities to test the mechanism of the artifact. In this case, we suspected that a correlation between a predictor variable (fat-free mass) and a variable used for determining PR (TEE) could cause selection bias – that is, overestimation of effect size. Simulation provided an opportunity to test how the overestimation would change with different correlation magnitudes between fat-free mass and TEE. Such testing is difficult in real data, as various types of predictors have to be available (although see Rhee *et al.*[21] for an example). On the basis of simulation, we now know that even a small correlation between fat-free mass and TEE would have caused a selection bias and overestimation of the sEI–fat free mass association's effect size.

The current study once again documented high under-reporting in adolescents. While the current high estimate (76 %) might be lower when a different sEI estimation method is used[19,21], the adolescent under-reporting problem is still widely known from previous literature. The under-reporting mechanism is hard to capture – 'the detection of under-reporting does not automatically reveal the process responsible'[16]. A previously outlined reason could be that the task of tracking food for 3 d could be cognitively too demanding for adolescents[66]. They might forget food items or not comply with the task. However, current data cannot provide evidence to the reasons for under-reporting. To properly understand the mechanisms of under-reporting, future research should simultaneously measure both sEI and oEI[67] for the same meals, and experimentally manipulate or randomise possible mechanisms, such as perception bias[68] or cognitive ability.

Intriguingly, controlling for predictors of EB% recovered the fat-free mass and sEI association rather well. Although the association between fat-free mass and oEI was even stronger ($\beta = 0{\cdot}51$), oEI was measured only for a single meal. Single meal association with fat-free mass has been similarly strong previously (r 0·42, 0·29). sEI at the same time was assessed for 3 d and averaged for a single day. On the basis of previous studies, one could expect that the association between fat-free mass and full day EI ranges between $\beta = 0{\cdot}28$ and 0·33[27]. In the current study, the corrected effect size was $\beta = 0{\cdot}35$, which is surprisingly close. At the same time, such success might be the peculiarity of the current sample and has to be replicated.

The current study has several limitations. Our study group included adolescent boys, and therefore the effect sizes seen pertain to this study group. At the same time, our empirical data and simulations show that the basic principle of selection bias should remain, and that statistical correction using external predictors of EB% is likely the best approach. We were unable to obtain data from questionnaires and oEI from all participants. However, groups with and without more detailed data did not differ in terms of basic sample statistics (Table 1), suggesting that this is not a major concern. oEI and sEI were based on different measurements – sEI was based on 3-d self-observed dietary records, whereas oEI was based on one breakfast comprised of convenience food, which is likely not the most optimal choice of food. Further, as we captured only one meal for oEI, we were unable to evaluate the EB% for oEI. Nevertheless, replicating previously known findings that current oEI can be predicted by fat-free mass allowed us to be certain that oEI was measured reasonably well. However, future studies should have (a) more naturalistic food and (b) oEI and sEI data

https://doi.org/10.1017/S0007114516003317 Published online by Cambridge University Press
British Journal of Nutrition

should be based on the same food consumed; people should report what they ate at the same time their eating habits are objectively captured (e.g. Stubbs *et al.*[67]). Finally, TEE was somewhat imprecise, as it was estimated from an equation, as opposed to measuring actual resting metabolic rate. Measuring actual resting metabolic rate could have decreased the association between TEE and fat-free mass, decreasing the size of the artifactual association between sEI and fat-free mass, if only PR are considered. Similarly, the TEE equation was derived from a younger sample than the current study sample, possibly making the TEE less accurate. Nevertheless, our simulations showed that any association between fat-free mass and TEE would have caused artifacts, when sEI is related to fat-free mass in a subsample of PR; therefore, despite the inaccuracies in TEE measurement, the major conclusion of the paper remains.

The current study also has several strengths, which allowed us to conclude that misreporting of sEI data should be statistically corrected using external predictors of EB%. Compared with Börnhorst *et al.*[20], we extended the results in several ways. We related sEI to a different predictor – fat-free mass. The supposed EI and fat-free mass association was first independently verified using oEI data, before we set to recover the association from sEI. Such an approach enabled us to know what type of effect size to look for. For methodological strengths, fat-free mass was objectively measured with DXA, and TEE for EB% was calculated based on objective physical activity. We also extended the previous findings by first using a different measure of sEI, 3-d dietary interview, which ensured that statistical correction applies for multiple measures of sEI. Second, we included various psychological predictors of EB% not included by Börnhorst *et al.*[20]. Despite these methodological differences, we reached a very similar conclusion, suggesting the robustness of using statistical correction.

Another strength was the use of several methods to scrutinise the appearance of selection bias. We first re-sampled our analysed data, which should have eliminated any association between sEI and fat-free mass. However, some associations remained, suggesting the existence of selection bias. We further demonstrated the causal role of selection bias by varying association strength between fat-free mass and TEE in a simulation study.

In summary, we suggest that future studies on sEI should plan ahead to include the known predictors of EB% in their data collection procedures. These could include BMI, restraint, social desirability or other relevant variables[12]. Our empirical data and simulation indicated that studying only PR groups can artificially increase the regression coefficient in certain conditions due to selection bias. Until more accurate and easily applicable EI measures are developed, statistically correcting sEI remains the best approach in large-scale studies.

Acknowledgements

The authors thank Triin Rääsk, Maarja Aarlaid and Liisi Panov for helping with data collection, Kairi Kreegipuu and Jüri Allik for their support and Caroline Uhler for pointing us to the selection bias literature. Finally, the authors thank Yashar Zeighami, Selin Neseliler and the three anonymous reviewers for their feedback.

This research was supported by the Estonian Ministry of Education and Science Institutional Grants IUT 20-58 and IUT2-13 and by the Doctoral School of Behavioral, Social and Health Sciences created under the auspices of European Social Fund.

Formulating the current research question(s): U. V., K. K.; designing the study: U. V., E. L., J. M., P. P., J. J.; carrying out the study: U. V., E. L., J. M., P. P., J. J.; analysing the data: U. V.; and writing the article: U. V. All the authors have read and commented on the manuscript.

The authors declare that there are no conflicts of interest.

Supplementary material

For supplementary material/s referred to in this article, please visit http://dx.doi.org/doi:10.1017/S0007114516003317

References

1. Ng M, Fleming T, Robinson M, *et al.* (2014) Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **384**, 766–781.
2. Blundell JE, Caudwell P, Gibbons C, *et al.* (2012) Role of resting metabolic rate and energy expenditure in hunger and appetite control: a new formulation. *Dis Model Mech* **5**, 608–613.
3. Vainik U, Dagher A, Dubé L, *et al.* (2013) Neurobehavioural correlates of body mass index and eating behaviours in adults: a systematic review. *Neurosci Biobehav Rev* **37**, 279–299.
4. Dagher A (2012) Functional brain imaging of appetite. *Trends Endocrinol Metab Tem* **23**, 250–260.
5. Buckeridge DL, Charland K, Labban A, *et al.* (2014) A method for neighborhood-level surveillance of food purchasing. *Ann N Y Acad Sci* **1331**, 270–277.
6. Harris KM, Perreira K & Lee D (2009) Obesity in the transition to adulthood: predictions across race-ethnicity, immigrant generation, and sex. *Arch Pediatr Adolesc Med* **163**, 1022–1028.
7. Cohen D & Farley TA (2008) Eating as an automatic behavior. *Prev Chronic Dis* **5**, A23.
8. Wansink B (2004) Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annu Rev Nutr* **24**, 455–479.
9. Livingstone MBE & Black AE (2003) Markers of the validity of reported energy intake. *J Nutr* **133**, 895S–920S.
10. Livingstone MBE, Robson PJ & Wallace JMW (2004) Issues in dietary intake assessment of children and adolescents. *Br J Nutr* **92**, Suppl. S2, S213–S222.
11. Archer E, Hand GA & Blair SN (2013) Validity of US nutritional surveillance: National Health and Nutrition Examination Survey caloric energy intake data, 1971–2010. *PLOS ONE* **8**, e76632.
12. Börnhorst C, Huybrechts I, Ahrens W, *et al.* (2013) Prevalence and determinants of misreporting among European children in proxy-reported 24 h dietary recalls. *Br J Nutr* **109**, 1257–1265.
13. Noel SE, Mattocks C, Emmett P, *et al.* (2010) Use of accelerometer data in prediction equations for capturing implausible dietary intakes in adolescents. *Am J Clin Nutr* **92**, 1436–1445.

14. Stice E, Palmrose CA & Burger KS (2015) Elevated BMI and male sex are associated with greater underreporting of caloric intake as assessed by doubly labeled water. *J Nutr* **145**, 2412–2418.
15. Tooze JA, Subar AF, Thompson FE, *et al.* (2004) Psychosocial predictors of energy underreporting in a large doubly labeled water study. *Am J Clin Nutr* **79**, 795–804.
16. Macdiarmid J & Blundell J (1998) Assessing dietary intake: who, what and why of under-reporting. *Nutr Res Rev* 11, 231–253.
17. Huang TT-K, Roberts SB, Howarth NC, *et al.* (2005) Effect of screening out implausible energy intake reports on relationships between diet and BMI. *Obes Res* **13**, 1205–1217.
18. Nielsen SJ & Adair L (2007) An alternative to dietary data exclusions. *J Am Diet Assoc* **107**, 792–799.
19. Mendez MA, Popkin BM, Buckland G, *et al.* (2011) Alternative methods of accounting for underreporting and overreporting when measuring dietary intake-obesity relations. *Am J Epidemiol* **173**, 448–458.
20. Börnhorst C, Huybrechts I, Hebestreit A, *et al.* (2013) Diet–obesity associations in children: approaches to counteract attenuation caused by misreporting. *Public Health Nutr* **16**, 256–266.
21. Rhee JJ, Sampson L, Cho E, *et al.* (2015) Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies. *Am J Epidemiol* **181**, 225–233.
22. Greenland S & Pearl J (2011) Adjustments and their consequences-collapsibility analysis using graphical models: adjustments and their consequences. *Int Stat Rev* **79**, 401–426.
23. Hernán MA, Hernández-Díaz S & Robins JM (2004) A structural approach to selection bias: epidemiology. **15**, 615–625.
24. Elwert F & Winship C (2014) Endogenous selection bias: the problem of conditioning on a collider variable. *Annu Rev Sociol* **40**, 31–53.
25. Mendez MA (2015) Invited commentary: dietary misreporting as a potential source of bias in diet-disease associations: future directions in nutritional epidemiology research. *Am J Epidemiol* **181**, 234–236.
26. Rhee JJ & Willett WC (2015) Rhee and Willett respond to ‘Dietary Misreporting’. *Am J Epidemiol* **181**, 237.
27. Blundell JE, Caudwell P, Gibbons C, *et al.* (2012) Body composition and appetite: fat-free mass (but not fat mass or BMI) is positively associated with self-determined meal size and daily energy intake in humans. *Br J Nutr* **107**, 445–449.
28. Cuenca-García M, Ortega FB, Ruiz JR, *et al.* (2014) More physically active and leaner adolescents have higher energy intake. *J Pediatr* **164**, 159–166.e2.
29. Fearnbach SN, Thivel D, Meyermann K, *et al.* (2015) Intake at a single, palatable buffet test meal is associated with total body fat and regional fat distribution in children. *Appetite* **92**, 233–239.
30. Ivuškāns A, Jürimäe T, Lätt E, *et al.* (2014) Role of physical activity in bone health in peripubertal boys. *Pediatr Int* **56**, 763–767.
31. Ivuškāns A, Mäestu J, Jürimäe T, *et al.* (2014) Sedentary time has a negative influence on bone mineral parameters in peripubertal boys: a 1-year prospective study. *J Bone Miner Metab* **33**, 85–92.
32. Jürimäe J, Lätt E, Mäestu J, *et al.* (2015) Osteocalcin is inversely associated with adiposity and leptin in adolescent boys. *J Pediatr Endocrinol Metab* **28**, 571–577.
33. Lätt E, Mäestu J, Rääsk T, *et al.* (2013) Association of physical activity to cardiovascular fitness and fatness in 12–13-year-old boys in different weight status. *J Public Health* **21**, 231–239.
34. Lätt E, Mäestu J, Ortega FB, *et al.* (2015) Vigorous physical activity rather than sedentary behaviour predicts overweight and obesity in pubertal boys: a 2-year follow-up study. *Scand J Public Health* **43**, 276–282.
35. Rääsk T, Konstabel K, Mäestu J, *et al.* (2015) Tracking of physical activity in pubertal boys with different BMI over two-year period. *J Sports Sci* **33**, 1649–1657.
36. Rääsk T, Lätt E, Jürimäe T, *et al.* (2015) Association of subjective ratings to objectively assessed physical activity in pubertal boys with differing BMI. *Percept Mot Skills* **121**, 245–259.
37. Remmel L, Tillmann V, Mäestu J, *et al.* (2015) Associations between bone mineral characteristics and serum levels of ghrelin and peptide YY in overweight adolescent boys. *Horm Res Paediatr* **84**, 6–13.
38. Utsal L, Tillmann V, Zilmer M, *et al.* (2012) Elevated serum IL-6, IL-8, MCP-1, CRP, and IFN-γ Levels in 10- to 11-year-old boys with increased BMI. *Horm Res Paediatr* **78**, 31–39.
39. Vaitkeviciute D, Lätt E, Mäestu J, *et al.* (2014) Physical activity and bone mineral accrual in boys with different body mass parameters during puberty: a Longitudinal Study. *PLOS ONE* **9**, e107759.
40. World Health Organization (2015) *WHO|The WHO Child Growth Standards*. Geneva: WHO. http://www.who.int/childgrowth/standards/en/ (accessed October 2015).
41. Marshall WA & Tanner JM (1970) Variations in the pattern of pubertal changes in boys. *Arch Dis Child* **45**, 13–23.
42. Ojiambo R, Konstabel K, Veidebaum T, *et al.* (2012) Validity of hip-mounted uniaxial accelerometry with heart-rate monitoring vs. triaxial accelerometry in the assessment of free-living energy expenditure in young children: the IDEFICS Validation Study. *J Appl Physiol* **113**, 1530–1536.
43. Akkermann K, Herik M, Aluoja A, *et al.* (2010) Söömishäirete Hindamise Skaala (Eating Disorders Assessment Scale). Institute of Psychology, University of Tartu, Tartu.
44. Podar I, Hannus A & Allik J (1999) Personality and affectivity characteristics associated with eating disorders: a comparison of eating disordered, weight-preoccupied, and normal samples. *J Pers Assess* **73**, 133–147.
45. Price M, Higgs S & Lee M (2015) Self-reported eating traits: underlying components of food responsivity and dietary restriction are positively related to BMI. *Appetite* **95**, 203–210.
46. Vainik U, Neseliler S, Konstabel K, *et al.* (2015) Eating traits questionnaires as a continuum of a single concept. Uncontrolled eating. *Appetite* **90**, 229–239.
47. Epel ES, Tomiyama AJ, Mason AE, *et al.* (2014) The Reward-Based Eating Drive scale: a self-report index of reward-based eating. *PLOS ONE* **9**, e101350.
48. Laidra K, Allik J, Harro M, *et al.* (2006) Agreement among adolescents, parents, and teachers on adolescent personality. *Assessment* **13**, 187–196.
49. John OP, Caspi A, Robins RW, *et al.* (1994) The ‘Little Five’: exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Dev* **65**, 160–178.
50. Konstabel K, Aavik T & Allik J (2006) Social desirability and consensual validity of personality traits. *Eur J Personal* **20**, 549–566.
51. Pitsi T, Kambek L & Jõelecht A (2014) NutriData toidu koostise andmebaas (NutriData Food Composition Database). Estonian National Institute for Health Development. www.nutridata.ee (accessed December 2014).
52. Brooks GA, Butte NF, Rand WM, *et al.* (2004) Chronicle of the Institute of Medicine physical activity recommendation: how a physical activity recommendation came to be among dietary recommendations. *Am J Clin Nutr* **79**, 921S–930S.
53. Huang TT-K, Howarth NC, Lin BH, *et al.* (2004) Energy intake and meal portions: associations with BMI percentile in U.S. children. *Obesity Res* **12**, 1875–1885.

54. Schafer JL (1999) Multiple imputation: a primer. *Stat Methods Med Res* **8**, 3–15.
55. Honaker J, King G & Blackwell M (2011) Amelia II: a program for missing data. *J Stat Softw* **45**, 1–47 (accessed August 2016).
56. Buuren S van & Groothuis-Oudshoorn K (2011) Mice: multivariate imputation by chained equations. *R. J Stat Softw* **45**, 1–20.
57. Barnard J & Rubin DB (1999) Small-sample degrees of freedom with multiple imputation. *Biometrika* **86**, 948–955.
58. R Core Team (2013) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/(accessed June 2015).
59. Wickham H & Francois R (2014) RStudio. dplyr: a grammar of data manipulation. http://cran.r-project.org/web/packages/dplyr/index.html (accessed November 2014).
60. Lemon J (2006) Plotrix: a package in the red light district of R. *R-News* **6**, 8–12.
61. Venables WN & Ripley BD 2002, Modern applied statistics with S. In *Statistics and Computing* [J Chambers, W Eddy, W Härdle, S Sheather and L Tierney, editors]. New York, NY: Springer New York. http://link.springer.com/10.1007/978-0-387-21706-2 (accessed May 2016).
62. Trautmann H, Steuer D, Mersmann O, *et al.* (2014) Truncnorm: truncated normal distribution. https://cran.r-project.org/web/packages/truncnorm/index.html (accessed May 2016).
63. Wagenmakers E-J & Gronau QF (2016) A compendium of clean graphs in R. http://shinyapps.org/apps/RGraphCompendium/index.php (accessed May 2016).
64. Field AP (2012) *Discovering Statistics Using R*. London and Thousand Oaks, CA: Sage.
65. Fox J (1991) *Regression Diagnostics: An Introduction*. Newbury Park, CA: Sage.
66. McPherson RS, Hoelscher DM, Alexander M, *et al.* (2000) Dietary assessment methods among school-aged children: validity and reliability. *Prev Med* **31**, S11–S33.
67. Stubbs RJ, O'Reilly LM, Whybrow S, *et al.* (2014) Measuring the difference between actual and reported food intakes in the context of energy balance under laboratory conditions. *Br J Nutr* **111**, 2032–2043.
68. Chandon P & Wansink B (2007) Is obesity caused by calorie underestimation? A psychophysical model of meal size estimation. *J Mark Res* **44**, 84–99.

British Journal of Nutrition