

# Validation of protein intake assessed from weighed dietary records against protein estimated from 24 h urine samples in children, adolescents and young adults participating in the Dortmund Nutritional and Longitudinally Designed (DONALD) Study

Beate Bokhof<sup>1,2</sup>, Anke LB Günther<sup>1,3</sup>, Gabriele Berg-Beckhoff<sup>4</sup>, Anja Kroke<sup>2</sup> and Anette E Buyken<sup>1,\*</sup>

<sup>1</sup>Research Institute of Child Nutrition, Affiliated Institute of the University of Bonn, Heinstück 11, D-44225 Dortmund, Germany; <sup>2</sup>Institute of Medical Computer Science, Biometrics and Epidemiology, University Hospital Essen, Essen, Germany; <sup>3</sup>Department of Nutritional, Food and Consumer Sciences, Fulda University of Applied Sciences, Fulda, Germany; <sup>4</sup>Faculty of Public Health Sciences, Department of Epidemiology and International Public Health, University of Bielefeld, Bielefeld, Germany

Submitted 29 June 2009; Accepted 9 December 2009; First published online 15 January 2010

## Abstract

**Objective:** To date, only a few nutritional assessment methods have been validated against the biomarker of urinary-N excretion for use in children and adolescents. The aim of the present study was to validate protein intake from one day of a weighed dietary record against protein intake estimated from a simultaneously collected 24 h urine sample.

**Design:** Cross-sectional analyses including 439 participants of the Dortmund Nutritional and Longitudinally Designed (DONALD) Study from four age groups (3–4, 7–8, 11–13 and 18–23 years). Mean differences, Pearson correlation coefficients ( $r$ ), cross-classifications and Bland–Altman plots were used to assess agreement between methods.

**Results:** Weighed dietary records significantly underestimated mean protein intake by  $-6.4$  (95% CI  $-8.2, -4.7$ ) g/d or  $-11\%$ , with the difference increasing across the age groups from  $-0.6$  (95% CI  $-2.7, 1.5$ ) g/d at age 3–4 years to  $-13.5$  (95% CI  $-18.7, -8.3$ ) g/d at age 18–23 years. Correlation coefficients were  $r = 0.7$  for the total study sample and ranged from  $r = 0.5$  to  $0.6$  in the different age groups. Both methods classified 85% into the same/adjacent quartile for the whole study group (83–86% for the different age groups) and 2.5% into the opposite quartile (1.9–3.1% for the different age groups). Bland–Altman plots for the total sample indicated that differences in protein intake increased across the range of protein intake, while this bias was not obvious within the age groups.

**Conclusions:** Protein intake in children and adolescents can be estimated with acceptable validity by weighed dietary records. In this age-heterogeneous sample, validity was lower among adolescents and young adults.

**Keywords**  
Validation  
Protein intake  
Weighed dietary records  
Urine  
Biomarker  
Children

Nutritional intake is considered to play a decisive role in the development of a number of major chronic diseases, such as adiposity, diabetes mellitus, CVD and cancer. Recently, interest in nutritional intake during childhood as a determinant of these diseases has increased. Valid assessment of nutritional intake in periods potentially critical for later diseases, i.e. in early childhood, mid-childhood and puberty, is thus a major challenge. Methods proven to be valid for assessing nutritional intake in adults may not, however, necessarily be valid for use in children

and adolescents, since their food choice could be particularly prone to fluctuation. Furthermore, the validity of these assessment tools may vary by age and sex or change over the course of time<sup>(1–4)</sup>. Adolescents are assumed to provide less accurate reports of their nutritional intake and/or may not be fully aware of their dietary intake, whereas parental reports of nutritional intake in younger children may be more valid. Furthermore, it has been suggested that girls are more prone to under-reporting their food intake than boys<sup>(5)</sup>. Also, recent trends among

\*Corresponding author: Email buyken@fke-do.de

adolescents to increase fast-food consumption and eat out may no longer be adequately reflected by commonly used assessment tools<sup>(1)</sup>. Finally, the increasing variation in foods offered may not be validly considered<sup>(1)</sup>. Thus, even weighed dietary records – commonly regarded a quasi gold standard in nutrition assessment – require validation for their use in children and adolescents.

In the absence of an assessment tool capable of estimating 'true' intake, biomarkers are increasingly recognised as best practice to assess absolute validity<sup>(6)</sup>. Until now, particularly urinary-N excretion as a recovery biomarker for protein intake has rarely been used among children and adolescents in nutritional studies. Some have applied the method of doubly labelled water in order to examine the validity of energy intake<sup>(5)</sup>. The use of the biomarker urinary-N excretion, however, may yield additional information about the validity of dietary records beyond absolute dietary intake. In the contemporary Dortmund Nutritional and Longitudinally Designed (DONALD) Study, 24 h urine samples are routinely collected on the third day of a 3 d weighed dietary record. Thus, the study offers the potential to validate nutritional intake against a urinary biomarker in a large population of boys and girls with a wide age range.

The present study assessed the validity of dietary protein intake recorded on the third day of a 3 d weighed dietary record against the protein intake estimated from simultaneously collected 24 h urine samples. Among the DONALD participants, 439 males and females from four different age groups, representing early childhood, mid-childhood and puberty as well as young adulthood (for comparative purposes), were selected to examine whether the validity was related to age or sex.

## Subjects and methods

### *Study population*

The DONALD Study is an ongoing, open cohort study that was started in 1985 in the area of Dortmund, Germany. Details of the study design have been published previously<sup>(7)</sup>. In brief, on average forty to fifty healthy infants are newly recruited each year and are first examined at the age of 3–6 months. Detailed data on nutrition, growth, metabolism and health status are collected at regular intervals between infancy and young adulthood. The study was approved by the Ethics Committee of the University of Bonn, and all assessments are performed with parental consent.

Eligible were all children, adolescents and young adults who had provided both a 3 d weighed dietary record and a 24 h urine collection on the third day of recording. Their urine collections had to be complete and a measurement of urinary-N had to be available. Micturitions with time of collection <20 h were rejected. Completeness was verified by excretion of creatinine in urine<sup>(8)</sup>. For the

purpose of the current analysis children and adolescents were selected who had these data available at ages 3–4, 7–8, 11–13 or 18–23 years. From the final data set comprising 439 participants, one set of records and urine was selected for each person. Distribution of the participants across the age groups is given in Table 1.

### *Nutritional assessment*

In the DONALD Study, dietary intake is assessed by 3 d weighed dietary records. Parents are asked to weigh all foods and beverages consumed by their children, including leftovers, to the nearest 1 g over three consecutive days with the help of regularly calibrated electronic food scales (initially Soehnle Digita 8000, Leifheit SG, Nassau, Germany; now WEDO digi 2000, Werner Dorsch GmbH, Muenster/Dieburg, Germany). Parents are instructed by trained dietitians and semi-quantitative recording using household measures (e.g. number of spoons, scoops) is allowed when exact weighing is not possible. Information on recipes or the types and brands of food items is also requested. The dietary records are analysed using the continuously updated in-house nutrient database LEBTAB<sup>(9)</sup>, which is based on information from standard nutrient tables, product labels or recipe simulations based on the labelled ingredients and nutrients.

### *Urine sampling and urine analysis*

Starting at the children's age of 3 years, annual 24 h urine collections are routinely performed on the third day of the 3 d weighed dietary record. Parents and children are instructed personally and in written form on how to collect complete 24 h urine samples. The first micturition in the morning on the third day of recording is discarded. Parents and children register the time of micturition, thereby defining the start of the 24 h collection which ends with the first micturition of the following morning. All micturitions are stored immediately in Extran-cleaned, preservative-free (Extran MA 03; Merck, Darmstadt, Germany) 1-litre plastic containers at  $\leq -12^{\circ}\text{C}$  before being transported to the Research Institute. A dietitian picks up the urine sample, asks parents and children about the completeness of the urine samples and possible influences on urine collection, and records this information on the protocol sheet. At the institute, the containers are stored at  $\leq -20^{\circ}\text{C}$  before being analysed. All urine samples undergo routine check using a commercial test strip after thawing and stirring. Volume, pH, osmolarity and creatinine are determined. Completeness of urine is ascertained via values of creatinine, based on sex- and age-specific body-weight-related reference values of creatinine for the age groups of 3, 4–5, 6–8, 9–13 and 14–18 years. Hence collection-related errors are identified<sup>(8)</sup>.

For the present study, N in urine was measured by the method of Kjeldahl (model 430 Digestor and model B-324 Distillation Unit; Buechi, Flawil, Switzerland) and concentration was obtained by relating N levels to the volume

**Table 1** General characteristics of the DONALD study sample (*n* 439)

	3–4 years ( <i>n</i> 97)		7–8 years ( <i>n</i> 137)		11–13 years ( <i>n</i> 102)		18–23 years ( <i>n</i> 103)	
	<i>n</i> or Median	% or P25, P75	<i>n</i> or Median	% or P25, P75	<i>n</i> or Median	% or P25, P75	<i>n</i> or Median	% or P25, P75
Female, <i>n</i> and %	56	57.7	58	42.3	47	46.1	59	57.3
Anthropometry								
BMI (kg/m <sup>2</sup> )	15.4	14.7, 16.3	16.0	14.8, 17.7	18.1	16.5, 20.2	22.0	20.3, 24.3
BMI SDS <sup>(13)</sup>	−0.03	−0.06, 0.6	0.17	−0.6, 0.9	0.06	−0.6, 0.7	0.24	−0.4, 1.0
BF% <sup>(13)</sup>	18.3	15.8, 20.7	17.5	14.4, 21.7	18.7	16.7, 21.8	22.1	15.7, 28.6
Overweight†, <i>n</i> and %	8	8.3	27	19.7	15	14.7	16	14.7
3 d weighed records								
Protein intake (g/d)	35.0	29.5, 44.7	48.0	39.9, 58.0	62.4	49.1, 77.1	68.4	50.3, 84.3
Protein intake (% of energy)	13.0	11.2, 14.3	12.4	11.0, 14.3	13.5	11.6, 15.3	13.4	11.4, 15.1
Energy intake (kJ/d)	4661	4075, 5502	6410	5531, 7297	7749	6489, 9192	8485	6699, 10506
EI:BMR	1.4	1.2, 1.5	1.5	1.3, 1.6	1.4	1.3, 1.6	1.6	1.3, 1.8
EI:BMR below age- and sex-specific cut-off values <sup>(4)</sup> , <i>n</i> and %	1	1.0	3	2.2	8	7.8	11	10.7
24 h urine collections								
N excretion (mmol/l) from 24 h urine	739	550, 921	748	571, 950	765	595, 1006	564	389, 776
Protein intake (g/d) estimated from 24 h urinary-N excretion	36.5	31.3, 41.9	50.6	42.1, 60.5	68.6	56.4, 81.9	83.5	69.3, 100.0
Protein intake (g/d) estimated from 24 h urinary-N excretion allowing for retention <sup>(12)</sup>	37.1	31.9, 42.4	52.4	43.6, 62.2	70.7	57.9, 84.2	83.5	69.3, 100.0
Creatinine (mmol/l)	5.0	3.8, 6.4	7.0	5.5, 8.7	9.1	7.4, 12.3	10.0	6.6, 13.2
Volume of 24 h urine collection (ml)	420	325, 600	591	450, 750	735	580, 990	1215	860, 1700

DONALD, Dortmund Nutritional and Longitudinally Designed; P25, 25th percentile; P75, 75th percentile; SDS, standard deviation score; BF%, percentage of body fat; EI, energy intake.

Values are presented as *n* and % as indicated; or median and P25, P75.

†According to the International Obesity Taskforce.

of the 24 h collections (Table 1: N excretion mmol/l from 24 h urine). To obtain protein intake estimated from the urine, the assumption was made that excreted N accounts for 80% of the ingested protein due to extra-renal N losses<sup>(6)</sup>. Thus, excreted N (mmol/l) was converted to g protein/d (protein g/d = (N g/d × 6.25 × 14)/1000) and then divided by 0.8 (Table 1: protein intake estimated from 24 h urinary-N excretion)<sup>(6,10)</sup>. In a further step, an age- and sex-specific reference value for protein requirements for growth was added to allow for protein retention during growth in the paediatric age groups (Table 1: protein intake estimated from 24 h urinary-N excretion allowing for retention)<sup>(11)</sup>.

### **Anthropometry and parental data**

Body weight is assessed to the nearest 100 g with an electronic scale, height to the nearest 0.1 cm and skinfold thickness to the nearest 0.1 mm. For each child, age- and sex-independent standard deviation scores (SDS) of weight, height and BMI were calculated using the German reference curves<sup>(12)</sup>. Percentage of body fat was calculated using Deurenberg's equations<sup>(13)</sup>. The proportion of overweight children was assessed according to the definitions of the International Obesity Taskforce, which correspond to a cut-off of 25 kg/m<sup>2</sup> in adults<sup>(14)</sup>. On a child's entry to the study, parents are asked to provide information about family characteristics, their educational status and employment, and their weight and height are measured by the same trained nurses who assess the anthropometrics of the participating children.

### **Statistical analysis**

For the present analyses, we used only dietary data recorded on the day when the urine sample was collected, since we expect the protein in the 24 h urine (comprising all excretions until the following morning) to reflect mainly the protein ingested during this day. For comparative purposes, agreement was also assessed by relating protein estimated from the urine to the mean protein intakes derived from all three days of recording.

Descriptive statistics of the study sample are presented as medians with the 25th and 75th percentile, since some variables were not normally distributed. To test for any interaction with age or sex we used linear regression models with protein intake estimated from 24 h urine as the dependent variable and dietary protein intake, age group and sex as independent variables. Analyses indicated a significant interaction with age group only (*P* values for interactions: 0.02), hence all analyses were performed for the total sample as well as for the four age groups.

Mean differences and standard deviations were calculated to allow conclusions on a group level about the absolute extent of over- or underestimation of protein intake by the weighed diet records. Relative median differences (ratio of mean difference and protein intake estimated from 24 h urine × 100) permitted a direct

comparison of the degree of over- or underestimation across the age groups. Pearson correlation coefficients were calculated using both crude and energy-adjusted data<sup>(15)</sup> from dietary records. By means of cross-classification, we computed the percentage of persons who were classified into either the same or the adjacent quartile of protein intake by both methods or misclassified into the opposite quartile<sup>(10)</sup>. Again, cross-classification was performed using both crude and energy-adjusted values<sup>(15)</sup> from weighed dietary records. Bland–Altman plots were used to illustrate the difference between the two methods against the mean of the two methods. The horizontal dashed line indicates the mean of the differences. The upper and lower dotted lines represent the upper and lower 95% limits of agreement, which should comprise 95% of the values in the range of the twofold standard deviation ( $1.96 \times \text{SD}$ ) of the mean differences ( $d \pm 1.96 \times \text{SD}$ ). Ideally, mean difference between the methods would be zero with no discernable bias, i.e. the mean differences would cluster on the horizontal continuous line of equality ( $y=0$ ). Any deviation of the mean difference line from the line of equality indicates a bias. Moreover, any systematic variation of the differences in protein intake across the range of protein intakes suggests the presence of an additional systematic in the bias, which would provide further evidence of a limited agreement between the methods<sup>(16–18)</sup>. As proposed by Bland and Altman, log-transformation of the data for the whole study group was performed to control for such systematic increases or decreases in the variation<sup>(17)</sup>.

Energy adjustment was carried out using the residual method<sup>(15)</sup>. We further calculated the ratio of energy intake to BMR and applied age- and sex-specific cut-off-points in order to assess potential under-reporting of energy intake<sup>(4)</sup>.

All statistical analyses were carried out using the SAS statistical software package version 8.2 (SAS Institute Inc., Cary, NC, USA), and a *P* value of <0.05 was considered statistically significant.

## **Results**

Table 1 gives the general characteristics of the study sample. Protein intakes recorded from the weighed dietary records were generally lower than protein intakes estimated from 24 h urines, regardless of whether protein retention was accounted for. The percentage of under-reported energy intake, based on age- and sex-specific cut-off-values<sup>(4)</sup>, ranged from 1% in children aged 3–4 years to 11% in young adults aged 18–23 years. The BMI SDS close to zero indicates that the BMI of the present sample was comparable to that of the German reference population<sup>(12)</sup>. Overweight prevalence ranged from 8 to 20% in the study participants, while 23–36% had an

**Table 2** Agreement between protein intake estimated from one day of a weighed dietary record and protein intake estimated from 24 h urinary-N excretion: differences; Pearson correlation coefficients and cross-classifications for agreement (*n* 439)

	3–4 years ( <i>n</i> 97)	7–8 years ( <i>n</i> 137)	11–13 years ( <i>n</i> 102)	18–23 years ( <i>n</i> 103)	Total ( <i>n</i> 439)
<b>Differences</b>					
Absolute mean difference (g/d)	−0.6	−5.0	−6.8	−13.5	−6.4
95% CI	−2.7, 1.5	−7.3, −2.7	−10.5, −3.1	−18.7, −8.3	−8.2, −4.7
Relative median difference (%)	−2.9	−10.1	−9.8	−19.4	−10.5
P25, P75†	−18.0, 17.6	−21.5, −7.1	−24.1, −7.6	−33.6, 1.5	−24.8, 7.2
<b>Pearson correlation coefficients</b>					
Crude <i>r</i>	0.52*	0.56*	0.59*	0.48*	0.70*
Energy-adjusted <i>r</i>	0.52*	0.50*	0.56*	0.30*	0.30*
<b>Cross-classification into quartiles</b>					
<b>Crude same/adjacent</b>					
<i>n</i>	83	114	85	89	371
%	85.6	83.2	83.3	86.4	84.5
<b>Crude opposite</b>					
<i>n</i>	3	4	2	2	11
%	3.1	2.9	2.0	1.9	2.5
<b>Energy-adjusted same/adjacent</b>					
<i>n</i>	80	112	83	79	354
%	82.5	81.8	81.4	76.7	80.6
<b>Energy-adjusted opposite</b>					
<i>n</i>	2	7	3	3	15
%	2.1	5.1	2.9	2.9	3.4

P25, 25th percentile; P75, 75th percentile.

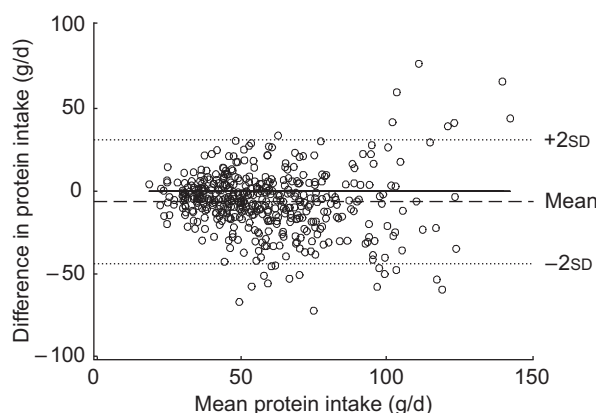
\* $P < 0.0001$ .

†The relative mean difference was not normally distributed and is thus presented as relative median difference.

overweight mother. The sample was characterised by a high parental educational level.

Table 2 gives the results of the validity analyses. Protein intake according to weighed dietary records was underestimated by  $-6.4$  (95% CI  $-8.2, -4.7$ ) g/d, which corresponded to a relative median difference of  $-10.5\%$ . Mean differences increased across age groups, both absolutely and relatively. Pearson correlation coefficients for the total study sample were  $r = 0.70$  ( $P < 0.0001$ ) using crude data and  $r = 0.30$  ( $P < 0.0001$ ) for energy-adjusted data, with the corresponding correlation coefficients ranging from  $r = 0.30$  to  $0.59$  across the age groups. Concerning cross-classification, over 80% of the participants in the total sample and in the different age groups were classified into the same or adjacent category using crude protein intake data, whereas classifications with energy-adjusted protein intake data indicated slightly lower agreement between the methods (77–83%). Misclassification into the opposite quartile was  $\leq 5\%$  in the total study group and all age groups for both crude and energy-adjusted data.

The Bland–Altman plot for the total study sample indicated that differences in protein intake were scattered and increased across the range of protein intake. The limits of agreement comprised less than 95% of the values (Fig. 1). Mean log-transformed protein intake (g/d) plotted against differences of log-transformed protein intake (g/d) for the total group (Fig. 2) removed the increase in variation, hence confirming the presence of a systematic bias for the total group. However, this increase in variation appeared to result from an age-related increase in

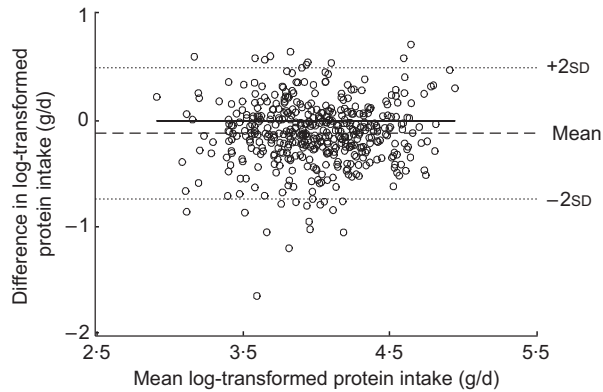


**Fig. 1** Bland–Altman plot for the total study group. The difference between protein intake (g/d) calculated from a weighed dietary record (test method) and protein intake level (g/d) estimated from 24 h urinary-N excretion (reference method) for each person (*y*-axis) is plotted against the mean protein intake averaged from the two methods (*x*-axis). Data are presented for the total study sample (*n* 439). The horizontal solid line ( $y = 0$ ) represents ideal agreement, where the differences between methods are zero; the horizontal dashed line indicates the mean of the differences; the upper and lower dotted lines show the upper and lower 95% limit of agreement, respectively, presented as twofold standard deviations ( $\pm 1.96 \times \text{sd}$ )

differences between the two methods, since age-specific Bland–Altman plots did not confirm such an obvious bias (Figs 3(a) to (d)).

We repeated the analyses, comparing mean protein intakes from the complete 3 d dietary record with protein intake estimated from the urine. The mean difference for





**Fig. 2** Bland–Altman plot of log-transformed data for the total study group. The difference between log-transformed protein intake (g/d) calculated from a weighed dietary record (test method) and log-transformed protein intake level (g/d) estimated from 24 h urinary-N excretion (reference method) for each person (y-axis) is plotted against the mean of log-transformed protein intake averaged from the two methods (x-axis). Data are presented for the total study sample ( $n$  439). The horizontal solid line ( $y=0$ ) represents ideal agreement, where the differences between methods are zero; the horizontal dashed line indicates the mean of the differences; the upper and lower dotted lines show the upper and lower 95% limit of agreement, respectively, presented as twofold standard deviations ( $\pm 1.96 \times \text{sd}$ )

the total study sample was 0.6 (95% CI  $-0.6$ , 1.9) g/d or  $-13\%$ , the correlation coefficient was  $r=0.8$  and both methods classified 91% into the same/adjacent quartile for the whole study group and 1% into the opposite. Bland–Altman plots were similar to those for the comparison based on the third day of recording only.

## Discussion

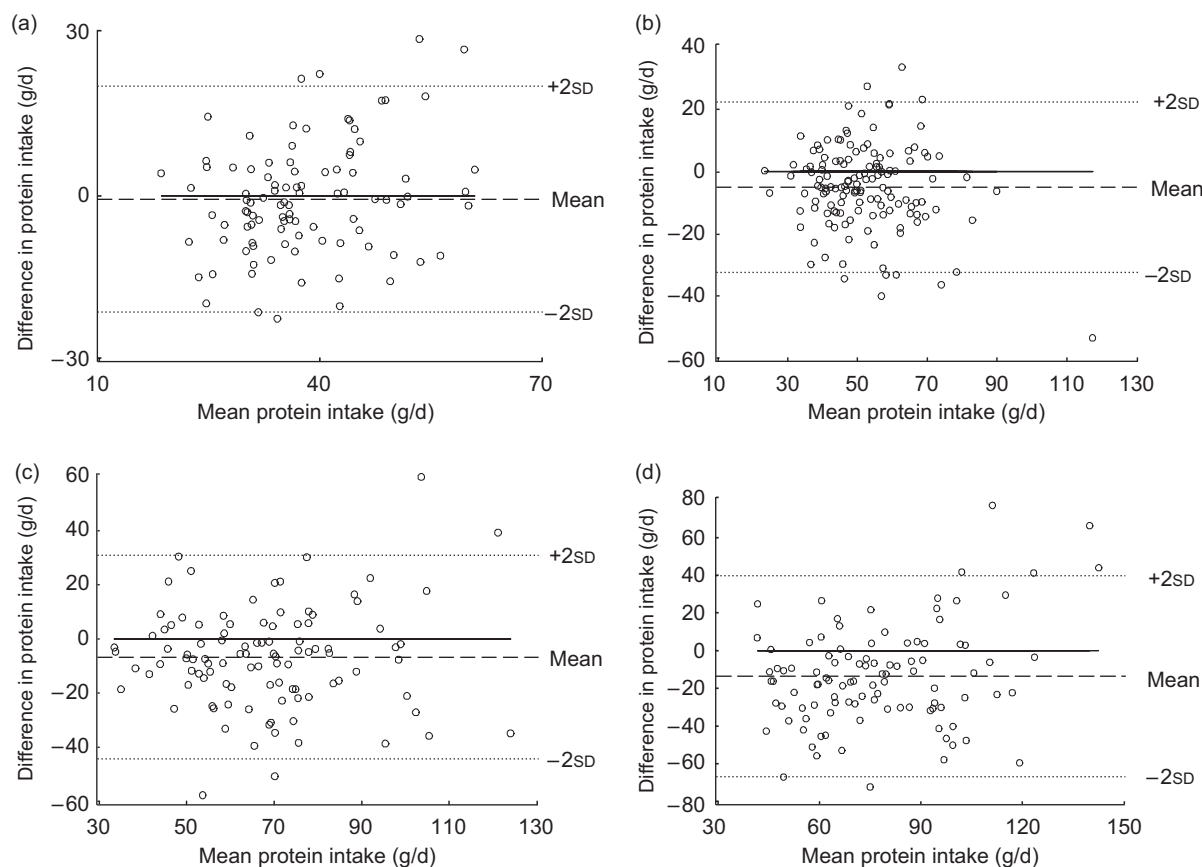
The results of our study suggest a moderate overall validity of protein intake estimated from weighed dietary records in a large sample of healthy children, adolescents and young adults as evidenced from a comparison with urinary-N excretion levels as a biomarker. Agreement between the methods tended to be lower in older age groups.

To our knowledge, the present study is the first to validate protein intake reported from weighed dietary records against protein intake estimated from 24 h urine samples in a large sample of healthy children and adolescents. To date, validation studies in children and adolescents comparing protein intake from weighed dietary records to the recovery biomarker urinary-N excretion have been carried out only in children suffering from diabetes mellitus or renal insufficiency<sup>(19,20)</sup>. The scarcity of comparable studies in healthy children is probably due to the high level of compliance and considerable effort required from the children, adolescents and their parents for (repeated) 24 h urine collections. Furthermore, the positive N balance of growing children and adolescents

represents a specific methodological challenge. Therefore, our study took advantage of an ongoing longitudinal study, while validation studies in adults were mostly experimental and especially designed to assess validity<sup>(21–23)</sup>.

The overall level of validity in our study as evidenced from Pearson correlation coefficients and cross-classifications was higher than that seen in other studies with children<sup>(19,24)</sup> and close to that found in experimental studies conducted in adults<sup>(21,22)</sup>. In the present study, weighed dietary records underestimated protein intake by  $-6$  g/d or  $-11\%$ . The overall magnitude of underestimation appears acceptable; however, the limits of agreement ranged from  $+31$  g to  $-44$  g. Five per cent of the data were located beyond these limits. Our overall conclusion of an acceptable agreement is likely based on conservative estimations, since agreement was even higher when dietary records from all three days were considered, i.e. in a comparison including protein ingested and at least partly excreted before the urine collection was started.

Hackett *et al.* reported an overestimation of 2.9 g protein/d among 3–16-year-old participants with diabetes mellitus<sup>(19)</sup>. Among comparable validation studies conducted in healthy adults, one reported a slight overestimation of protein intake by 1.31 g/d<sup>(22)</sup>, while another study found an underestimation by 10% when comparing protein intake assessed by weighed dietary records to urinary-N excretion levels<sup>(21)</sup>. Thus, published literature does so far not suggest an explicit tendency towards under- or overestimation of protein intake by weighed dietary records. Nevertheless, our results seem plausible, since various studies with different designs suggest that nutritional intake tends to be underestimated by a number of different nutrition assessment methods<sup>(25)</sup>. Underreporting has been reported to occur especially among overweight persons and girls during puberty<sup>(4,5)</sup>. In the present study, we did not observe any sex differences in the association between protein intake reported in weighed dietary records and protein intake estimated from urinary-N excretion levels. On the other hand, we did observe a lower level of agreement between the methods in older age groups. Hackett *et al.* also reported age-related differences in validity comparable to our observations<sup>(19)</sup>. Furthermore, in a study conducted in 10–15-year-old girls to validate energy intake assessed by dietary records against doubly labelled water, the authors arrived at the conclusion that the accuracy of the data obtained from dietary records was lower in older age groups<sup>(5)</sup>. The lower level of validity among adolescents could be due to a lower level of compliance in accurately weighing and recording all foods and beverages consumed or to the fact that adolescents tend to consume more foods away from home, which are estimated with a lower level of precision. In contrast, among younger children, nutritional intake was recorded by their parents, who are still relatively well informed about the nutritional



**Fig. 3** Bland–Altman plots for different age groups. The difference between protein intake calculated from a weighed dietary record (test method) and protein intake level estimated from 24 h urinary-N excretion (reference method) for each person (*y*-axis) is plotted against the mean protein intake averaged from the two methods (*x*-axis). Data are presented for different age groups: (a) age 3–4 years (*n* 97); (b) age 7–8 years (*n* 137); (c) age 11–13 years (*n* 102); (d) age 18–23 years (*n* 103). The horizontal solid line ( $y = 0$ ) represents ideal agreement, where the differences between methods are zero; the horizontal dashed line indicates the mean of the differences; the upper and lower dotted lines show the upper and lower 95% limits of agreement, respectively, presented as twofold standard deviations ( $\pm 1.96 \times \text{SD}$ )

intake of their children, thus potentially contributing to the observed higher validity.

The Bland–Altman plots underpinned this observation, indicating a tendency towards larger differences between the methods at higher levels of protein intake, which were largely observed in the adolescents and young adults. Consistent with this view, the Bland–Altman plots for each age group – i.e. with age-specific smaller overall ranges of protein intake – did not reveal additional systematic in the bias with higher protein intake levels. We therefore conclude that a moderate agreement in the absolute differences may be observed as long as single age groups are examined. By contrast, when large ranges of protein intake are considered (e.g. when examining a wide age range) intake levels estimated from weighed dietary records could be systematically biased.

Interestingly, in the present study, energy adjustment did not enhance validity. The number of participants with ‘implausible energy intake levels’, based on age- and sex-specific cut-off values, was rather low<sup>(4)</sup>. Thus, energy adjustment may not improve validity of the dietary data in

populations with largely plausible dietary records and should be challenged in further studies.

A particular strength of the present analysis lies in the use of different statistical methods to determine agreement. Pearson correlation coefficients indicate the size of correlation between the test and reference method and should thus only be used in combination with other validation methods to measure the extent of agreement between two methods. In cross-classification the percentage of agreement also comprises randomly originated agreements but this method is independent of the data distribution and potential artefacts resulting from statistics<sup>(26)</sup>. Furthermore, in nutritional epidemiology participants are commonly grouped by their intake level, thus supporting the relevance of conclusions drawn from cross-classification analyses. Finally, Bland–Altman plots enable assessment of the extent of variance in agreement and bias. Even though Bland–Altman plots are considered a gold standard in studies of methodical comparisons<sup>(18)</sup>, validation studies in nutritional epidemiology have only started to use this method during the past few years<sup>(27)</sup>.

Finally, some limitations of our study need to be mentioned. While we were able to compare 24 h urine to 24 h diet records on the same day in a large sample of healthy children and adolescents we could not address intra-personal day-to-day variation in protein intake, since our results are based on a single 24 h urine collection. In our view this may have resulted in an underestimation of the 'true' validity as indicated by the fact that agreement was already higher when protein estimated from urinary-N excretion was compared with recorded protein intake averaged from three days. According to Bingham and Cummings, intra-personal variations in protein can be eliminated by collection of eight or more complete 24 h urine samples, allowing 81 (SD 5) %<sup>(28)</sup> of the consumed protein to be found in the urine samples. However, conduction of a validity study under metabolic ward conditions over a period of at least a week does not seem feasible among healthy children and adolescents.

Since the present study was not specifically designed to determine the validity of recorded protein intake, we judged the completeness of the 24 h urine collection by means of creatinine excretion. Use of an external marker such as *p*-aminobenzoic acid would have allowed a more reliable verification of completeness. Hence, misclassification of samples collected over notably longer or shorter periods as 'complete 24 h samples' may have introduced additional variation in protein intake levels estimated from urinary-N excretion. Potential bias due to energy under-reporting remains a possibility; however, analyses excluding under-reporters yielded essentially comparable results (data not shown). Nevertheless we are aware that conclusions in this regard are hampered by the fact that the present analyses are based on one day of a dietary record for which a broad range of energy intakes has to be regarded as plausible<sup>(4)</sup>.

Furthermore, we based our allowances for extra-renal losses on percentages adopted from adults in the absence of values for children and adolescents. Finally, parents of children and adolescents participating in DONALD have relatively high socio-economic status<sup>(7)</sup> and are interested in health and nutrition issues<sup>(29)</sup>. As a consequence, the DONALD sample is not representative for Germany and conclusions from the present results may not be extended to other populations, e.g. to samples comprising less motivated children and parents. Nevertheless, previous analyses have suggested that nutritional intake in the DONALD Study is largely comparable to that observed in nationwide representative studies<sup>(29)</sup>.

In summary, our study suggests an acceptable validity of protein intake from one day of a weighed dietary record as compared with protein intake estimated from the recovery biomarker 24 h urinary-N excretion in a large population of children and adolescents. Since lower levels of validity were found among adolescents and young adults, caution should be exercised when analysing study samples with broad age ranges.

## Acknowledgements

The DONALD Study is funded by the Ministry of Science and Research of North Rhine Westphalia, Germany. Support for additional lab analyses was provided by WCRF International. No author had any conflicts of interest. A.K. conceived the project. B.B. and A.L.B.G. performed initial data analyses. B.B. conducted further analyses and drafted the manuscript. A.E.B. supervised the study. G.B.-B. and A.L.B.G. provided statistical expertise. All authors made substantial contributions to the study design and/or the interpretation of the results.

## References

1. Mensink GB & Burger M (2004) What do you eat? Food frequency questionnaire for children and adolescents. *Bundesgesundheitsblatt* **47**, 219–226.
2. Oja P (1997) Health-related physical activity and fitness among European children and adolescents. *World Rev Nutr Diet* **81**, 98–104.
3. Alexy U, Sichert-Hellert W & Kersting M (2002) Fifteen-year time trends in energy and macronutrient intake in German children and adolescents: results of the DONALD study. *Br J Nutr* **87**, 595–604.
4. Sichert-Hellert W, Kersting M & Schoch G (1998) Underreporting of energy intake in 1 to 18 year old German children and adolescents. *Z Ernahrungswiss* **37**, 242–251.
5. Bandini LG, Must A, Cyr H *et al.* (2003) Longitudinal changes in the accuracy of reported energy intake in girls 10–15 y of age. *Am J Clin Nutr* **78**, 480–484.
6. Bingham SA (2003) Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* **133**, Suppl. 3, 921S–924S.
7. Kroke A, Manz F, Kersting M *et al.* (2004) The DONALD Study. History, current status and future perspectives. *Eur J Nutr* **43**, 45–54.
8. Remer T, Neubert A & Maser-Gluth C (2002) Anthropometry-based reference values for 24-h urinary creatinine excretion during growth and their use in endocrine and nutritional research. *Am J Clin Nutr* **75**, 561–569.
9. Sichert-Hellert W, Kersting M, Chada C *et al.* (2007) German food composition data base for dietary evaluations in children and adolescents. *J Food Compos Anal* **20**, 63–70.
10. Gibson R (2005) Validity in dietary assessment methods. In *Principles of Nutritional Assessment*, pp. 147–196 [R Gibson, editor]. New York: Oxford University Press.
11. Garlick PJ (2006) Protein requirements of infants and children. *Nestle Nutr Workshop Ser Pediatr Program* **58**, 39–47.
12. Kromeyer-Hauschild K, Wabitsch M, Kunze D *et al.* (2001) Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschr Kinderheilkd* **149**, 807–818.
13. Deurenberg P, Pieters JJ & Hautvast JG (1990) The assessment of the body fat percentage by skinfold thickness measurements in childhood and young adolescence. *Br J Nutr* **63**, 293–303.
14. Cole TJ, Bellizzi MC, Flegal KM *et al.* (2000) Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* **320**, 1240–1243.
15. Willett W & Stampfer M (1998) Implications of total energy intake for epidemiologic analyses. In *Nutritional*



- Epidemiology*, pp. 273–301 [W Willett, editor]. New York: Oxford University Press.
16. Bland JM & Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.
  17. Bland JM & Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* **8**, 135–160.
  18. Grouven U, Bender R, Ziegler A *et al.* (2007) Comparing methods of measurement. *Dtsch Med Wochenschr* **132**, Suppl. 1, e69–e73.
  19. Hackett AF, Morton S, McCowen C *et al.* (1987) Measurement of food intake in children with diabetes mellitus: a comparison of protein intake and urine nitrogen. *Diabetes Res* **6**, 29–32.
  20. Wingen AM, Fabian-Bach C & Mehls O (1993) Evaluation of protein intake by dietary diaries and urea-N excretion in children with chronic renal failure. European Study Group for Nutritional Treatment of Chronic Renal Failure in Childhood. *Clin Nephrol* **40**, 208–215.
  21. Bingham SA, Cassidy A, Cole TJ *et al.* (1995) Validation of weighed records and other methods of dietary assessment using the 24h urine nitrogen technique and other biological markers. *Br J Nutr* **73**, 531–550.
  22. Black AE, Welch AA & Bingham SA (2000) Validation of dietary intakes measured by diet history against 24h urinary nitrogen excretion and energy expenditure measured by the doubly-labelled water method in middle-aged women. *Br J Nutr* **83**, 341–354.
  23. Kroke A, Klipstein-Grobusch K, Voss S *et al.* (1999) Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* **70**, 439–447.
  24. Lietz G, Barton KL, Longbottom PJ *et al.* (2002) Can the EPIC food-frequency questionnaire be used in adolescent populations? *Public Health Nutr* **5**, 783–789.
  25. Livingstone MB & Black AE (2003) Markers of the validity of reported energy intake. *J Nutr* **133**, Suppl. 3, 895S–920S.
  26. Block G (2001) Invited commentary: another perspective on food frequency questionnaires. *Am J Epidemiol* **154**, 1103–1104.
  27. Cade JE, Burley VJ, Warm DL *et al.* (2004) Food-frequency questionnaires: a review of their design, validation and utilisation. *Nutr Res Rev* **17**, 5–22.
  28. Bingham SA & Cummings JH (1985) Urine nitrogen as an independent validity measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. *Am J Clin Nutr* **42**, 1276–1289.
  29. Kersting M, Sichert-Hellert W, Alexy U *et al.* (1998) Macronutrient intake of 1 to 18 year old German children and adolescents. *Z Ernahrungswiss* **37**, 252–259.