

Validation of web-based, multiple 24-h recalls combined with nutritional supplement intake questionnaires against nitrogen excretions to determine protein intake in Dutch elite athletes

F. C. Wardenaar^{1*}, J. Steennis^{1,2}, I. J. M. Ceelen¹, M. Mensink², R. Witkamp² and J. H. M. de Vries²

¹*Sports and Exercise Studies, HAN University of Applied Sciences, 6525 AJ Nijmegen, The Netherlands*

²*Division of Human Nutrition, Wageningen University, 6708 PB Wageningen, The Netherlands*

(Submitted 9 March 2015 – Final revision received 20 August 2015 – Accepted 2 September 2015 – First published online 5 October 2015)

Abstract

Information on dietary composition is vitally important for elite athletes to optimise their performance and recovery, which requires valid tools. The aim of the present study was to investigate the validity of assessing protein intake using three web-based 24-h recalls and questionnaires, by comparing these with three urinary N excretions on the same day. A total of forty-seven Dutch elite top athletes, both disabled and non-disabled, aged between 18 and 35 years, with a BMI of 17.5–31 kg/m², exercising >12 h/week were recruited. Estimated mean dietary protein intake was 109.6 (SD 33.0) g/d by recalls and questionnaires *v.* 141.3 (SD 38.2) g/d based on N excretions in urine; the difference was 25.5 (SD 21.3)% between the methods ($P < 0.05$). We found a reasonably good association between methods for protein intake of 0.65 (95% CI 0.45, 0.79). On an individual level, under-reporting was larger with higher protein intakes than with lower intakes. No significant differences were found in reporting absolute differences between subcategories (sex, under-reporting, BMI, collection of recalls within a certain amount of time and using protein supplements or not). In conclusion, combined, multiple, 24-h recalls and questionnaires underestimated protein intake in these young elite athletes more than that reported for non-athlete populations. The method proved to be suitable for ranking athletes according to their protein intake as needed in epidemiological studies. On an individual level, the magnitude of underestimation was about equal for all athletes except for those with very high protein intakes.

Key words: Validation: 24-h recalls: Nitrogen excretion: Athletes

Dietary composition and timing of food intake are vitally important for elite athletes to optimise their performance and recovery^(1–3). Protein requirements, in particular, differ between elite athletes and physically less active people. Recommendations for protein intake are between 1.2 and 2.0 g/kg per d and are frequently tailored to individual needs^(4,5). This demands for adequate tools to advise and monitor the intake of athletes.

Food records and 24-h recalls are commonly used to assess current intake or intake of the recent past. To assess dietary intake by athletes at a group level, both 3–4-d or 7-d food records and multiple 24-h recalls have been used^(6,7). These methods, however, show large differences in validity between different populations⁽⁸⁾.

The twenty-first century athlete, often combining intensive training and competition programmes with college education, demands practical dietary assessment methods such as web-based 24-h recalls. The advantage of the 24-h recall is that the outcome is not affected by awareness and consequently an altered – desirable – intake as seen with the use of a food record⁽⁸⁾. Moreover, in using web-based recalls, athletes can report their intake, within certain limits, at their own

convenience. The preferred five-step, multiple-pass, 24-h recall is highly standardised and is considered to provide a reliable estimation of dietary intake⁽⁹⁾.

At the same time, all dietary assessment methods based on self-reporting are prone to different types of errors⁽¹⁰⁾. In addition to this, specific sports-related factors such as season, day of the week and training load contribute to daily variation in intake, which may contribute to errors if not taken into account properly. Total intake of energy and macronutrients, such as protein, normally display less random or day-to-day variation than other nutrients such as retinol or marine fatty acids⁽¹¹⁾. Systematic errors – that is, over-reporting or under-reporting of intake – are often due to socially desirable reporting, omitting foods and underestimating of portion sizes⁽¹²⁾.

Validation studies are important for identifying the magnitude and type of measurement error of a specific method. For validation purposes, 24-h urinary N is considered the best available biomarker to estimate dietary protein intake^(13–18). The method was first proposed by Isaksson⁽¹⁹⁾ and further developed by Bingham & Cummings⁽¹⁵⁾. When corrected for faecal and skin

Abbreviations: EI, energy intake; FIL, food intake level; PABA, *para*-aminobenzoic acid; PAL, physical activity level.

* **Corresponding author:** F. C. Wardenaar, email Floris.wardenaar@han.nl

losses of N, the method correlated well (r 0.99) with protein intake estimated by 28 d of duplicate portions^(15,19).

Information about the validity of self-reported dietary intake in elite athletes is scarce^(20,21). Their active lifestyle with intensive training programmes generates methodological problems. For example, dietary registration of multiple days is required to capture the large variety in dietary intake caused by differences in the intensity of activities between training days and days of rest and recovery.

To our knowledge, no validation studies have been undertaken thus far to estimate the accuracy of multiple 24-h recalls to assess protein intake in athletes. Therefore, the aim of the present study was to investigate the validity of assessing protein intake by three 24-h recalls plus day questionnaires on dietary supplement intake compared with analysis of urinary N excretion using three different 24-h samples in elite athletes.

Methods

Study population

A total of forty-eight subjects, both disabled and non-disabled Dutch elite athletes, aged between 18 and 35 years, with a BMI of 17.5–31 kg/m² and a minimum exercise duration >12 h/week were recruited. They had an elite top athlete status of the Dutch Olympic Committee (NOC*NSF) and/or had participated in a European or World Championship or had proven to be part of the national top level in their discipline or age group. Subjects were recruited from different national elite teams for athletics (middle distance and sprint), cycling (track sprint, BMX cyclists and Paralympics), archery, speed skating and short-track skating. Before the start of the study, all coaches confirmed that the training programme during the period of the study (3–5 weeks) was part of a stable-condition training phase. The majority of subjects did not focus specifically on building muscle mass or breaking down body fat as it could be a possible interfering factor when assessing N balance. Only four athletes intended to lose weight during this period. None of the subjects was diagnosed with a disease or used medication. Written informed consent to participate was provided by all the subjects after attending a presentation meeting about the study. Subjects received a financial incentive if they delivered at least two complete data sets out of three, consisting of a 24-h urine collection, a 24-h recall plus a questionnaire on dietary supplement and sport nutrition product intake and exercise. The survey was conducted according to the Declaration of Helsinki and was approved by the Medical Ethics Committee of Wageningen University.

Study design

The study was performed between March and July 2013. In the general information letter for candidate subjects, it was explained that the collection of 24-h urine samples served to determine the N content of the consumed diet. After including subjects, no additional information was given about the intended use of the 24-h urine collections or the timing of the 24-h urine collection in relation to the 24-h recalls.

After inclusion, subjects were asked to complete three unannounced web-based non-consecutive, 24-h dietary recalls and questionnaires about dietary supplements and sport nutrition products, in combination with a 24-h urine collection covering the same day. The measurements were scheduled during a 2–4-week period on three separate days, selected by the researchers, and collection days were at least 4 d apart. If the collection (24-h recall plus daily questionnaire and/or 24-h urine collection) was not complete, additional days were scheduled at the end of the 2–4-week period. This gave athletes four opportunities to deliver at least two 24-h recalls, questionnaires and 24-h urine samples.

24-h recalls

The 24-h recalls were collected using 'Compl-eat™', a program built by Wageningen University that guides participants to accurately report all foods and drinks consumed the previous day, and were to be reported within 36 h. It is designed based on the multiple-pass approach^(9,22). Using Compl-eat™, the participant first filled out the quick list indicating which food groups were consumed at what occasion and time of the day without giving details. The next step was to specify all food groups of the quick list meal by meal in chronological order during the day. The participant was allowed to continue with the next meal if all food groups of a meal had been specified and the amounts had been added. The participant was offered the possibility to split up food groups into subgroups by a search tool and scroll-down menus. In this way, the tool enables to select single foods and standard recipes commonly used by Dutch population. The programme also contains a recipe module in which the participants could adapt standard recipes or fill out all ingredients of their own recipe and the part of the recipe they consumed. Yield and retention factors were automatically taken into account when appropriate. Participants could include notes for clarifications when needed. Compl-eat™ reminded the participants to fill out often forgotten foods such as sugar in coffee, snacks, fruit and cooking fat.

Trained dietitians checked all the web-based 24-h recalls for their completeness and unusual portion sizes and processed all the notes made by the participants. Adjustments of errors and notes were made in a standardised manner using standard portion sizes and recipes according to a protocol⁽²³⁾.

The programme included a wide selection of foods commonly used in a Dutch food pattern (Stichting NEVO, 2010) but no dietary supplements or sport nutrition products, which were questioned separately.

Nutritional supplement intake questionnaires

The web-based questionnaires regarding training load (total minutes of exercise per d) and dietary supplement and sports nutrition product use as part of each separate 24-h recall were collected with the 'vitality portal' built by the HAN University of Applied Sciences. Subjects were asked whether they had used any dietary supplements and/or sports nutrition products. If so, they were asked to select the specific brand, type and name of the supplement from a pre-specified list with more than 3400 products. Product information was largely based on the Dutch



Database for Dietary Supplements (RIVM, 2015). Next, subjects were asked to specify the total dose of the product used on that day based on unit dose and type of formulation (capsule, tablet, drops, etc.) in case of dietary supplements or grams and/or millilitres in case of sport nutrition products. If a product was missing in the list, subjects were able to fill out all specifications as described earlier. We were able to retrieve the protein composition of all dietary supplements and sport nutrition products used based on label information given by the manufacturers. Data were checked for unusual values, unspecified dietary supplements and sports nutrition products. If necessary, participants were contacted and asked for additional information. If information was missing, the composition of dietary supplements and sports nutrition products was retrieved from the internet or through local shops.

Combining 24-h recalls and questionnaires

To calculate the mean total energy and protein intakes, results of all 24-h recalls with the questionnaire of the same day were summed and divided by the total number of days collected per person. For conversion of products into energy and nutrients, the Dutch Food Composition Database of 2010 was used (Stichting NEVO, 2010) in combination with the Dutch Database for Dietary Supplements (NES)^(24,25).

Total reported dietary intake per person was converted into total protein (g and g/kg) and total energy (MJ). Food intake level (FIL) was calculated by dividing estimated energy intake (EI) through BMR based on the formula of Schofield⁽²⁶⁾. The FIL was used as an indication of the quality of reporting in comparison with expected physical activity level (PAL) values for athletes. Based on studies using doubly labelled water, PAL values are expected to be between 1.75 and 2.0 for these young adult athletes⁽²⁷⁾. We compared the protein intake of the group above and below the median value for FIL to evaluate the influence of reporting quality.

Urine collection

Each subject received labelled containers (at least two), one funnel to help the collection, one safety pin to be fixed in the underwear as a reminder for collection and a diary scheme booklet to register the timing, observations (e.g. use of medication and supplements) and possible deviations (e.g. missing urine) of the urine collection protocol. Subjects were instructed to keep the urine samples at approximately 4°C at all times. Boric acid (3 g/2 litre bottle) was used as preservative.

The evening before the collection, subjects received a short text message between 22.00 and 23.00 hours as a reminder. The collection of the 24-h urine started with sampling and discarding the first urine in the morning after waking up. Subsequently, the urine produced during the next 24 h, up to and including the first sampling of the following day, was collected. The subjects handed in their urine samples at their training facilities, at maximum 5 h after the end of the collection.

To check for completeness of urinary collections, subjects took three tablets of 80 mg *para*-aminobenzoic acid (PABA) (PABAcheck; Laboratories for Applied Biology) on the day of urine collection, one during each of their main meals. The total

amount of 240 mg of PABA was expected to be almost completely excreted in 24 h^(28,29).

Analysis of urine samples

In our laboratory, urine fractions of the same person and of the 24-h period were mixed, weighed and aliquoted. Each sample of 24-h urine was measured to the nearest 1.0 g on a digital scale. The urine was homogenised and two samples of 4.5 ml were taken. Samples were stored at -20°C until further analysis. Total N was analysed using the automated Kjeldahl method⁽³⁰⁾. Total urinary N was calculated by relating N levels to the volume of the 24-h collections. The assumption was made that excreted N accounts for 81% of ingested protein due to extra-renal N losses^(15,31). Daily dietary protein was calculated based on the assumption that all proteins on average contain 16% N^(14,32). Thus, excreted N (mg/d) was converted to g protein/d (protein g/d = (N (mg/d) × 6.25)/1000) and then divided by 0.81⁽³³⁾.

HPLC was used for the determination of PABA. This method includes alkaline hydrolysis for conversion of PABA metabolites in the urine samples to PABA, followed by separation on a reverse-phase column, detection at 290 nm and quantification using an internal standard. Total urinary PABA was calculated by relating the PABA concentration of the samples to the volume of the 24-h collections.

Using a minimum PABA recovery of 78% proposed in the literature as the cut-off point for complete urine collection⁽³⁴⁾, 13% of the urine samples were considered incomplete. A sensitivity analysis was performed comparing the results from the complete urine data set with the results after exclusion of urine samples with <78% of PABA recovery.

The mean protein intake based on recalls and questionnaires was 10 g lower if subjects with PABA recovery <78% were excluded as compared with the total data set. However, excluding subjects based on PABA recovery did not affect estimates for protein intake based on 24-h urine excretions or the association between protein intake derived from recalls and urine samples. Therefore, we decided to report the results on the complete data set.

Other measurements

To check whether the assumption of N balance was justified, self-reported body weight for each day was recorded as an indicator for stable body weight over the total period. In addition, creatinine levels were analysed in a small subgroup (*n* 19 with two to three samples within each person) to estimate protein muscle balance as an indicator for N balance. Urinary creatinine was analysed by (SHO) using a spectrometric chemical analyser (Modular P800; Roche). All the samples were measured in duplicate and each run contained two control samples. The analyses were considered valid if duplicates differed by <5% and control samples were within the expected range. Total daily urinary creatinine content was calculated for the total urine collected over 24 h.

Statistical analyses

Statistical analysis was carried out on the data of all subjects delivering two to three complete data sets using the statistical

software program SPSS (version 19). Values for protein and energy are presented as continuous variables, expressed as means and standard deviations or 95 % CI. Log transformations were used, if necessary, to obtain a normal distribution for additional statistical analysis. Protein ratio was calculated by dividing the estimated mean protein intake based on 24-h recalls by protein based on 24-h N excretions.

To assess whether athletes were in N balance, a repeated-measures ANOVA (using mixed models with Bonferroni's correction) was performed using values of creatinine and body weight.

Paired *t* test was performed to determine the difference between total mean protein intake by 24-h recall and 24-h urine analysis with statistical significance levels set at $P \leq 0.05$. Furthermore, differences for protein intake were examined within several subgroups – sex, reporting, BMI category and recalls – by univariate analysis.

To investigate ranking of individuals according to their intake, Pearson's correlation coefficient was calculated between N estimated by the 24-h recall method and the 24-h urine sample analysis. A partial Pearson's correlation coefficient (*r*) was calculated for estimated crude and energy-corrected protein intakes, and the coefficients were computed using regression calibration-adjusted correlation. For all correlations, 95 % CI were calculated using Fisher's *Z* transformation.

For total protein intake (g/d and g/kg per d), attenuation factors and 95 % CI were calculated. These factors were estimated as the slope in the linear regression of the log-transformed biomarker values plotted against the log-transformed reported intakes. Although no formal cut-off values exist, values <0.4 were considered of no use.

A Bland–Altman plot was made to evaluate the agreement of the assessment of protein intake between recalls and urinary excretions in individuals. For this plot, the mean intake of protein estimated by the methods was plotted against the difference in protein between the methods. We added 95 % limits of agreement (mean (SD 2)) to the plot.

Results

The vast majority of the forty-seven athletes delivered three 24-h recalls, three questionnaires and three urine samples, except for two subjects who delivered only two complete data sets. This final study population consisted of both male (66 %

and female (34 %) athletes; participants were young adult elite athletes, with a BMI between 20 and 31 kg/m² (Table 1). Subjects practised a large variety of sports: five middle-distance runners, eight track cyclists, six BMX cyclists, eight Paralympic cyclists (of which three were physically disabled, one having Möbius syndrome, one visually disabled and three non-disabled tandem pilots), three athletic sprinters, seven archers, seven long-distance ice-skaters and three short-track skaters. Most recalls (74.5 %) were collected within 36 h, and all recalls were collected within 72 h. A proportion of 78.7 % of the subjects reported an FIL <1.75, indicating possible under-reporting of EI.

Mean protein intake assessed by the 24-h recalls combined with the nutritional supplement intake questionnaires was underestimated in comparison with protein estimated by the 24-h urine N excretion analysis. Estimated dietary mean protein intake was 109.6 (SD 33.0) g/d or 1.49 (SD 0.35) g/kg per d by recalls *v.* a 141.3 (SD 38.2) g/d or 1.90 (SD 0.39) g/kg per d based on N excretions in urine, with a protein ratio of 0.78. The absolute underestimation was –31.7 (SD 30.0) g/d, reflecting an absolute significant mean difference of 25.5 (SD 21.3) % between the methods in g/d ($P < 0.001$), which is in line with the previously suggested under-reporting of EI. Estimated mean EI was 16.9 (SD 4.2) MJ, with a range of 8.54–26.6 MJ.

We found a rather high correlation of 0.65 (95 % CI 0.45, 0.79) for crude protein intake in g/d between the multiple 24-h recall method and the 24-h N excretion analysis (Table 2). If protein intake was expressed as protein per kg body weight, the correlation was slightly lower, but still reasonably good (Table 2). Attenuation factors were high for total protein intake and good-to-acceptable for protein intake in g/kg per d. To predict true protein intake, as a gold standard is not available in daily practice, a calibrated regression function formula was calculated for both crude protein and energy-adjusted protein as shown in Table 2.

Fig. 1 depicts the agreements for difference of protein intake based on 24-h recalls and 24-h urine excretion analysis against the mean of both methods, in a so-called Bland–Altman plot. The 95 % limits of agreement were between –90.4 and +27.1 g. In addition, the Bland–Altman plot shows that results were very similar for both men and women. Underestimation of protein intake was related to the amount of protein intake ($r = -0.20$; 95 % CI –0.46, 0.09), indicating that higher protein intakes are possibly prone to higher under-reporting than lower intakes.

For all the categories shown in Table 3, no significant differences were found between protein intakes estimated by

Table 1. Characteristics of the forty-seven athletes (Numbers and percentages; mean values and standard deviations)

	Sex		Age (years)		Height (cm)		Body weight (kg)		BMI (kg/m ²)		FIL†				
	<i>n</i>	%	Mean	SD	Mean	SD	mean	SD	Mean	SD	BMI >25 (%)*	Recall <36 h (%)	Mean	SD	Reporters <PAL 1.75 (%)‡
Female	16	34	20.8	4.0	172.6	4.7	66.3	8.1	21.7	2.9	18.8	81.3	1.7	0.5	68.9
Male	31	66	21.4	3.9	182.8	5.5	78.4	8.9	21.5	4.6	19.4	71.0	1.5	0.3	83.7
Total	47	100	21.2	3.9	179.3	7.2	74.3	10.3	21.6	4.1	19.1	74.5	1.6	0.4	78.7

FIL, food intake level; PAL, physical activity level.

* >25.0 and <31.0 kg/m².

† FIL = energy intake/BMR, reported values between 0.88 and 2.81.

‡ PAL 1.75 is selected as a cut-off value to determine under-reporting for active younger adults based on FIL.

Table 2. Associations and attenuation factors (AF) for the comparison of self-reported protein intakes with urinary nitrogen excretions for crude protein intake (Correlation coefficients and 95 % confidence intervals)

Protein intake	Pearson's correlation		Attenuation		Calibrated regression function
	<i>r</i>	95 % CI*	AF	95 % CI	
Crude intake (g/d)	0.65	0.45, 0.79	0.76	0.49, 1.02	$y = 58.3 + (0.76 \times \text{g protein 24-h recall and questionnaires})\dagger$
Crude intake (g/kg per d)	0.58	0.35, 0.74	0.65	0.36, 0.95	$y = 0.43 + (0.52 \times \text{g/kg protein 24-h recall and questionnaires})\dagger$

* 95 % CI computed by Fisher's Z transformation.

† $y = \beta_0 + (\beta_1 \times \text{protein 24-h recall})$.

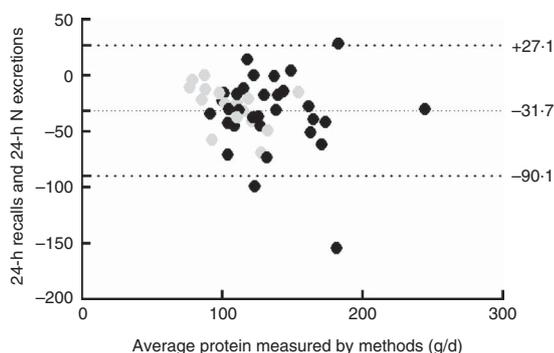


Fig. 1. Difference in protein intake based on 24-h recalls and 24-h urine excretions against the mean of both methods. Data points for women (●) and for men (●).

both methods between subcategories. The *P* values were 0.51 for sex, 0.58 for under-reporting according to a cut-off value of 1.75, 0.29 for BMI category, 0.08 for reporting within 36 h or more and 0.63 for protein supplement use or no use. When we split the population into two groups of low reporters and high reporters (below and above the median FIL of 1.52), the mean FIL was, respectively, 1.26 (sd 0.17) for the group below the median and 1.76 (sd 0.29) for the group above the median. The mean protein intake of the high reporters was significantly higher than that of the low reporters (*P* = 0.04).

For ranking based on correlations, some categories showed better results, as the other categories within subcategories. In all, thirteen out of forty-seven subjects used one or more protein supplement(s). Two subjects reported the use of protein supplements on all 3 d, three subjects on 2 d and eight subjects on only 1 d. For these thirteen subjects, the mean intake of protein including supplements was 129.3 (sd 41.0) g/d, which was 27.2 g/d higher than that of non-users of protein supplements (*n* 34). Protein supplement users showed a slightly higher correlation (*r* 0.70; 95 % CI 0.24, 0.90) but also a wider 95 % CI compared with non-users (*r* 0.65; 95 % CI 0.40, 0.81).

Total mean daily urinary creatinine was 19.4 (sd 4.69) mmol/d (with a mean of 18.6 (sd 4.17) mmol/d for women and 19.8 (sd 5.05) mmol/d for men), which is within the normal range of 10–20 mmol/d⁽³⁵⁾. Per sport category, small differences were observed. Speed skaters and short-track skaters had mean creatinine levels of approximately 14.5 mmol/d. The middle-distance runners, Paralympic cyclist, track cycling sprinters, athletic sprinters and archers had slightly higher mean creatinine levels of approximately 20 mmol/d (range 13–28). In addition, BMX riders had the highest creatinine levels in comparison with

others (average approximately 24 mmol/d, with a range of 19–32 mmol/d). Although some of the athletes had higher creatinine levels than within normal limits, their urinary creatinine levels did not differ between single recall days (*P* = 1.00), confirming that they were in N balance. In addition, according to the interviews with coaches (results not shown), none of the athletes gained or lost weight as part of a personal intervention, and daily body weight did not differ between recall days (*P* > 0.05). Based on the results of these combined measurements, we consider this group of subjects to be in N balance.

Discussion

As far as we know, this is the first validation study in elite athletes to evaluate a web-based 24-h recall with questionnaires for protein intake. Mean protein intake was underestimated by 25.5 %. The multiple 24-h recalls and questionnaires proved to be suitable for ranking athletes according to their protein intake, as reflected by the reasonably high correlation coefficient. At the individual level, we observed considerable underestimations of protein intake, which tended to be more pronounced in situations where intake was higher. Small differences were seen for subcategories of athletes.

Quality of results

It is difficult to compare the outcomes of the tools used in the present study – the web-based recall module Compl-eat™ and questionnaires – with those of other web-based tools^(36–40) because of different study designs. Recently, Compl-eat™ was compared with telephone recalls performed on randomly chosen days, independent of the web-based recalls, as part of an ongoing cohort study (NQplus). In this study, it was found that the tool under-reported EI by 8 % and protein intake by 9 %. Possible explanations for these differences included not defining serving sizes or ignoring components such as dairy or cooking fat. Thompson *et al.*⁽⁴¹⁾ evaluated a comparable method, Automated Self-Administered 24-h dietary recall, against an Automated Multiple-Pass Method (AMPM) interview by telephone and found an average under-reporting of 6 % for reported energy. Thus, a part of the under-reporting in our study may be explained by collection using a web-based tool instead of an AMPM-based interview by telephone.

As no comparable studies are available for athletes using N excretion as a validation method for web-based 24-h recalls, we discuss our results in comparison with studies performed in other groups. A study in women using two 24-h recalls and

Table 3. Mean protein intake, absolute difference and associations for sex, reporting, BMI category and recall time (Mean values and standard deviations; correlation coefficients and 95 % confidence intervals)

Category	Estimate of protein intake								Absolute difference (g/d)*		Correlation coefficient*		n
	24 h urine collection				24 h recall and questionnaires				Mean	SD	r	95 % CI†	
	Mean (g/d)	SD	Mean (g/kg per d)	SD	Mean (g/d)	SD	Mean (g/kg per d)	SD					
Sex													
Females	119.8	27.3	1.81	0.37	92.5	19.3	1.41	0.29	-27.3	19.2	0.69	0.30, 0.88	16
Males	152.4	38.6	1.95	0.43	118.5	35.2	1.53	0.38	-33.9	25.7	0.56	0.26, 0.76	31
Reporting													
FIL < PAL 1.75‡	142.6	37.5	1.88	0.41	108.5	31.5	1.45	0.33	-34.1	27.4	0.63	0.40, 0.79	41
FIL > PAL 1.75‡	132.0	44.9	2.00	0.45	117.1	44.6	1.74	0.39	-14.8	12.5	0.78	-0.09, 0.97	6
FIL low (<1.52)	139.4	38.8	1.78	0.44	101.7	26.4	1.31	0.28	-37.7	33.9	0.56	0.19, 0.79	23
FIL high (>1.52)	143.2	38.3	2.01	0.36	117.9	37.5	1.67	0.32	-25.3	24.4	0.76	0.51, 0.89	24
BMI category													
BMI <20 kg/m ²	126.9	38.1	1.84	0.47	92.7	23.3	1.41	0.34	-34.1	34.9	0.44	-0.15, 0.80	13
BMI 20–24.9 kg/m ²	146.3	34.1	2.00	0.35	113.4	29.4	1.55	0.32	-32.9	32.0	0.60	0.27, 0.80	25
BMI >25 kg/m ²	148.1	47.4	1.71	0.45	122.3	46.8	1.41	0.46	-25.7	14.4	0.93	0.70, 0.99	9
Recalls													
Recall <36 h	140.6	41.1	1.89	0.44	112.3	35.2	1.53	0.37	-28.3	22.5	0.66	0.42, 0.81	35
Recall >36 and <72 h	143.4	29.6	1.91	0.32	101.0	25.2	1.34	0.27	-42.4	20.4	0.74	0.29, 0.92	12
Protein supplement use													
Yes	158.9	42.3	2.08	0.40	129.3	41.0	1.72	0.32	-29.6	21.3	0.70	0.24, 0.90	13
No	134.5	34.8	1.83	0.40	102.1	26.3	1.40	0.33	-32.5	27.1	0.65	0.40, 0.81	34

FIL, food intake level; PAL, physical activity level.

* Energy-adjusted Pearson's partial correlation coefficients on log-transformed protein estimates (g/d or in g/kg per d).

† 95 % CI computed by Fisher's Z transformation.

‡ PAL 1.75 is selected as the cut-off value to determine under-reporting for active younger adults based on FIL.

F. C. Wardenar *et al.*

eight 24-h N excretions in urine⁽⁴²⁾ reported a smaller mean difference than our study. The difference found by Bingham *et al.*⁽⁴²⁾ was only 7.8 % between methods, which is substantial lower than the values of our women (24.9 %).

The European Food Consumption Validation (EFCOVAL) study included data of a population in the Netherlands that may be comparable with our results because of a similar dietary pattern⁽⁴³⁾. The study included two non-consecutive 24-h recalls and 24-h urine collections covering the same reference day. Under-reporting for mean protein intake in men was 12.4 % and in women it was 8.2%. Although total under-reporting was lower in this overweight older population, sex differences are pointing in the same direction.

In addition, the reported protein ratio of our study of 0.78 is in line with a range of 0.54–0.99 reported as part of the European Prospective Investigation into Cancer and Nutrition study in non-athletes⁽¹³⁾. Therefore, the extent of under-reporting at a group level in the present study is fairly acceptable in comparison with others using 24-h recalls^(13,43,44).

Three 24-h recalls covering the same reference day as the biomarker appeared to be good for ranking elite athletes according to their protein intake, as indicated by a rather high correlation with the reference method (r 0.65; 95 % CI 0.45, 0.79). Pooled results of five validation studies in non-athletes showed an average correlation for men of 0.39 (95 % CI 0.31, 0.47) and for women of 0.42 (95 % CI 0.36, 0.48)⁽⁴⁵⁾. The Dutch data subset of the EFCOVAL study also resulted in lower unadjusted correlations for men, r 0.42 (95 % CI 0.18, 0.61), and women, r 0.51 (95 % CI 0.29, 0.67)⁽⁴³⁾. Thus, based on the comparisons with these studies, the results of our validation study for protein are comparable or better than that of others^(43,45). However, most of the other studies did not cover the same reference day for recall and biomarker testing, as was the case in our study.

The observed attenuation factors for total protein intake are almost two times as high as in most others⁽⁴⁵⁾. This indicates that relative risks assessed in our population will not be substantially de-attenuated because of measurement error. High correlations and attenuation factors indicate that the method is suitable for ranking athletes according to their protein intake.

On an individual level, the magnitude of underestimation is quite large as shown by the levels of agreement in the Bland–Altman plots (with a range of –105.7 up to 17.3 g). This was also observed in other studies⁽¹⁴⁾. The higher protein intakes based on the means of the combined methods tended to be underestimated more than the lower intakes. An explanation for this finding might be that the 24-h recalls make use of standardised portion sizes, which may not be representative for athletes⁽¹⁴⁾. It is not very likely that we have missed specific protein-rich foods, because our questionnaire was specifically designed to cover all dietary supplements and sports nutrition products containing protein. Nevertheless, in future studies, estimation of protein intake needs specific attention.

Athletes may register their dietary intake more carefully than non-athletes as they represent a motivated and disciplined population. There is some evidence that motivated subjects report better than less-motivated subjects⁽⁴⁶⁾, but not much is known about the validity of protein intake values assessed by

multiple 24-h recalls in elite athletes. Mean under-reporting of 25.5 % is higher than that in the non-athlete literature, but these studies also differ in the methodology used compared with ours^(13,45). An explanation could be that under-reporting was larger because of a higher true mean protein intake of athletes compared with less-active people. This seems reasonable, because all the other results responded within the normal range in comparison with other publications.

Protein intake values in our study are in line with those reported by other groups of athletes. However, only a few studies used multiple 24-h recalls^(7,47). One study using three 24-h recalls reported an absolute mean protein intake of 1.6 (SD 0.4) g/kg per d⁽⁷⁾. This is slightly higher than the mean protein intake found in the present study. The only Dutch study examining elite athletes reported 1.0–3.0 g/kg per d for different types of athletes, based on 3–4-d food records⁽⁴⁸⁾.

Mean protein intake of dietary supplement users in our study was substantially higher than that of non-users. However, as the overall number of participants using dietary supplements was small, the contribution of protein supplements to the mean protein intake was limited.

Urinary N as marker for dietary protein as proposed by Isaksson⁽¹⁹⁾ is well accepted as a marker for actual protein intake^(13,14,16–18), but assumes N balance. Especially in athletes, this assumption may not always be applicable because athletes potentially have a higher protein turnover than more sedentary subjects⁽⁴⁹⁾. We used urinary creatinine levels as an indicator of protein turnover. Higher creatinine levels are often a sign of higher protein turnover⁽⁵⁰⁾. In athletes, higher urinary creatinine values supported by serum creatinine values compared with sedentary people have been found in several studies⁽⁵¹⁾. In our study, urinary creatinine levels in the athletes did not differ between the recall days. In addition, although the mean creatinine levels were high, they were still in line with reference values (10–20 mmol/d) for the normal population⁽⁵²⁾. Athletes performing sports characterised by a high BMI, indicating a high muscle mass, have higher serum creatinine levels than endurance athletes, who generally have a lower BMI^(53–55). Most athletes in the present study had pronounced muscular bodies probably resulting in higher net protein turnover and higher creatinine levels. However, we assume that our athletes were in N balance. This was also supported by stable body weights during the study period and confirmation by coaches that training schedules were stable. If N was overestimated, under-reporting of protein would actually have been smaller than that measured.

We used the data of all the subjects, despite incomplete PABA recovery in a small number of subjects. Regarding this, our results were not completely in line with previous findings⁽⁵⁶⁾, because we found a difference in estimated mean protein intake of the total population compared with that of the population excluding subjects with a low PABA recovery. However, as other results were not affected, we chose to follow the recommendation of Subar *et al.*⁽⁵⁶⁾ to include all subjects. This resulted in a smaller mean difference (11.3 g/d) without influencing the correlation coefficients.

We found no significant differences for mean protein intake within BMI or other subcategories. However, results for these subcategories should be interpreted with care because of small



sample sizes, resulting in a wide 95% CI in some, but not all cases. The higher association between protein intake and the biomarker found in athletes with a BMI >25 kg/m² compared with the other two subcategories is interesting, because in most validation studies overweight people misreport more than their leaner counterparts. A possible explanation is that athletes with a high BMI are mostly not overweight, but are very muscular and perhaps more aware of their protein intake than others.

We chose to compare the estimated FIL with a PAL of 1.75 to distinguish between accurate reporting and under-reporting. PAL for active people was estimated at approximately 1.75, besides this no training study reported individual PAL values over 2.5, and mean PAL values in young people reached an upper limit approximately 2.0⁽²⁷⁾. Although slightly conservative, this PAL seems reasonable for this mixed group of active athletes. We indeed found that reports below an FIL of 1.75 were less accurate than those of 1.75 or higher, suggesting under-reporting of energy as well.

Additional questionnaires were used to assess intake via supplements and/or sport nutrition products as the web-based 24-h recall module was in our opinion not specific enough in this respect. Owing to the ability to select out of >3400 pre-selected products, we think that the enquired types of foods were rather complete and could not explain the underestimated intake. The used food composition tables and ingredient declarations of the dietary supplements and sport nutrition products can have a significant influence on reporting. The contribution of these products was based on labelling information only, which may not have been accurate. However, because of the small number of users and the amounts used, these are considered of minor influence on the group mean protein intake. On the other hand, estimation of the normal diet based on food composition tables is also prone to under-reporting as is known from studies with duplicate portions⁽⁵⁷⁾.

Further, the combined 24-h recalls and questionnaires were unannounced, but followed consistently the day after the 24-h urine collection. It cannot be excluded that subjects were aware of this, and adapted their protein intake during the second and third collection. However, this would not have affected the difference in the estimated protein intake between recalls and urine sample collections.

It is unlikely that delay of reporting was of great influence on total group reporting. Some of the 24-h recalls were not collected within the requested 36 h. Owing to the strict planning of training schedules of the athletes, with different types of training days, it was not possible in some specific cases to perform re-scheduling to a later phase of the study. Exceptionally, a new time slot was opened and data collection took place within 72 h from the first invitation, but mostly within 48 h. No significant difference was seen between categories, although protein intake was substantially lower in the group reporting after 36 h.

Although our study population did not cover all types of sports, it was a fairly good mixture of different types of elite athletes competing at an international level. The mixed exercise types and training programmes reflect a representative workload for elite-level athletes. The men:women ratio was representative for athletes with a Dutch Olympic Committee (NOC*NSF) status.

The study was designed to assess current intake collecting all the data within 4 weeks. We believe this gives the best insight into the actual accuracy of multiple 24-h recalls in athletes, because athletes undergo training programmes with changing exercise load and volume throughout the year. Although the study design did not take into account all different seasons, the effect on the results is expected to be small at the population level. The study covered winter, spring and summer during a basic conditional training phase of different specific types of sports. Nevertheless, we can only draw conclusions about the validity of this method within a 4-week period.

In conclusion, the validity of assessing protein intake by three 24-h recalls plus day questionnaires on dietary supplement intake is fairly good in comparison with the protein intake based on N excretions in three different 24-h urine samples in elite athletes. The combined multiple 24-h recalls with an accompanying questionnaire about dietary supplements and sport nutrition products were shown to under-report protein intake in young adult elite athletes to the same extent as in other non-athlete populations. The method is, however, well suitable for ranking athletes according to their protein intake as is needed in epidemiological studies. When correlation for protein intake was adjusted for total EI based on self-reported intake, no clear improvement in the results was observed, indicating that protein intake is a focus of athletes regardless their total EI. On an individual level, the magnitude of underestimation was about equal for all athletes, except for those with a very high intake of protein. Other determinants related to under-reporting included sex, reporting FIL of <1.75, BMI, time to hand in recalls and using protein-containing dietary supplements or not, and the determinants should be taken into account for the dietary assessments of athletes to obtain better insight into the estimation and monitoring of protein intake.

Acknowledgements

The authors acknowledge Mark Rothuis, Pim Huting, Coen Manders, Victor Mooren and Sjoerd Privee for helping during this study as part of their internship.

This research received no specific grant from any funding agency or the commercial or not-for-profit sectors. The writing of this article was made possible by a regional government grant as part of the Eat2Move project of the province of Gelderland, the Netherlands.

F. C. W. (formulating the research question(s), study design, analysing the data and writing the article), J. S. (running data collection, analysing the data and writing the article), I. J. M. C. (writing the article), M. M. (formulating the research question(s) and writing the article), R. W. (formulating the research question(s) and writing the article), J. H. M. d. V. (formulating the research question(s), study design and writing the article).

There are no conflicts of interest.

References

1. Hawley JA, Dennis SC, Lindsay FH, *et al.* (1995) Nutritional practices of athletes: are they sub-optimal? *J Sports Sci* **13**, S75–S81.



2. Kerksick C, Harvey T, Stout J, *et al.* (2008) International Society of Sports Nutrition position stand: nutrient timing. *J Int Soc Sports Nutr* **5**, 17.
3. Jeukendrup A (2013) The new carbohydrate intake recommendations. *Nestle Nutr Inst Workshop Ser* **75**, 63–71.
4. Campbell B, Kreider RB, Ziegenfuss T, *et al.* (2007) International Society of Sports Nutrition position stand: protein and exercise. *J Int Soc Sports Nutr* **4**, 8.
5. Holwerda AM, van Vliet S & Trommelen J (2013) Refining dietary protein recommendations for the athlete. *J Physiol* **591**, 2967–2968.
6. Magkos F & Yannakoulia M (2003) Methodology of dietary assessment in athletes: concepts and pitfalls. *Curr Opin Clin Nutr Metab Care* **6**, 539–549.
7. Chaouachi A, Chamari K, Roky R, *et al.* (2008) Lipid profiles of judo athletes during Ramadan. *Int J Sports Med* **29**, 282–288.
8. Willett W (2012) *Nutritional Epidemiology*, 3rd ed. New York, NY: Oxford University Press.
9. Moshfegh AJ, Rhodes DG, Baer DJ, *et al.* (2008) The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr* **88**, 324–332.
10. Bingham SA (1994) The use of 24-h urine samples and energy expenditure to validate dietary assessments. *Am J Clin Nutr* **59**, 227S–231S.
11. Van Staveren WA, Ocké MC & De Vries JHM (2012) Estimation of dietary intake. In *Present Knowledge in Nutrition*, 10th ed., pp. 1012–1026 [JW Erdman, IA Macdonald and SH Zeisel, editors]. Ames, IA: Wiley-Blackwell.
12. Livingstone MB & Black AE (2003) Markers of the validity of reported energy intake. *J Nutr* **133**, Suppl. 3, 895S–920S.
13. Slimani N, Bingham S, Runswick S, *et al.* (2003) Group level validation of protein intakes estimated by 24-hour diet recall and dietary questionnaires against 24-hour urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. *Cancer Epidemiol Biomarkers Prev* **12**, 784–795.
14. Bingham SA (2003) Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* **133**, Suppl. 3, 921S–924S.
15. 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.
16. Black AE, Bingham SA, Johansson G, *et al.* (1997) Validation of dietary intakes of protein and energy against 24 hour urinary N and DLW energy expenditure in middle-aged women, retired men and post-obese subjects: comparisons with validation against presumed energy requirements. *Eur J Clin Nutr* **51**, 405–413.
17. Bokhof B, Gunther AL, Berg-Beckhoff G, *et al.* (2010) 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. *Public Health Nutr* **13**, 826–834.
18. Bokhof B, Buyken AE, Dogan C, *et al.* (2012) Validation of protein and potassium intakes assessed from 24 h recalls against levels estimated from 24 h urine samples in children and adolescents of Turkish descent living in Germany: results from the EVET! Study. *Public Health Nutr* **15**, 640–647.
19. Isaksson B (1980) Urinary nitrogen output as a validity test in dietary surveys. *Am J Clin Nutr* **33**, 4–5.
20. Rehrer NJ, Hellemans IJ, Rolleston AK, *et al.* (2010) Energy intake and expenditure during a 6-day cycling stage race. *Scand J Med Sci Sports* **20**, 609–618.
21. Westerterp KR, Saris WH, van Es M, *et al.* (1986) Use of the doubly labeled water technique in humans during heavy sustained exercise. *J Appl Physiol* **61**, 2162–2167.
22. Conway JM, Ingwersen LA, Vinyard BT, *et al.* (2003) Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. *Am J Clin Nutr* **77**, 1171–1178.
23. Donders-Engelen MR, Van der Heijden LJM & Hulshof KFAM (1997) *Maten, gewichten en codenummers (Food Portion Sizes and Coding Instructions)*. Wageningen: Department Human Nutrition, Wageningen Agricultural University.
24. Buurma-Rethans E, Fransen H, Ghameshlou Z, *et al.* (2008) Een databestand voor supplementen: behoeftes en acties (A dataset for dietary supplements: needs and actions). *Voeding Nu* **10**, 21–24.
25. RIVM/Dutch Nutrition Centre (2011) Dutch Food Composition Table 2011/version 3, NEVO-tabel. Bilthoven, the Netherlands: RIVM/Dutch Nutrition Centre.
26. Johansson G, Wikman A, Ahren AM, *et al.* (2001) Under-reporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living. *Public Health Nutr* **4**, 919–927.
27. Westerterp KR (2013) Physical activity and physical activity induced energy expenditure in humans: measurement, determinants, and effects. *Front Physiol* **4**, 90.
28. Runswick S, Slothouber B, Boeing H, *et al.* (2002) Compliance with the urine marker PABAcheck in cancer epidemiology studies. *IARC Sci Publ* **156**, 35–37.
29. Bingham S & Cummings JH (1983) The use of 4-aminobenzoic acid as a marker to validate the completeness of 24 h urine collections in man. *Clin Sci (Lond)* **64**, 629–635.
30. Hambleton LG & Noel RJ (1975) Protein analysis of feed, using a block digestion. *J AOAC* **58**, 143–145.
31. Kipnis V, Midthune D, Freedman LS, *et al.* (2001) Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol* **153**, 394–403.
32. Jones DB (1941) Factors for converting percentages of nitrogen in foods and feeds into percentages of proteins. Circular 183. Washington, DC: United States Department of Agriculture.
33. Gibson R (2005) *Validity of Dietary Assessment Methods, in Principles of Nutritional Assessment*. New York: Oxford University Press.
34. Jakobsen J, Pedersen AN & Ovesen L (2003) Para-aminobenzoic acid (PABA) used as a marker for completeness of 24 hour urine: effects of age and dosage scheduling. *Eur J Clin Nutr* **57**, 138–142.
35. Coban (2008) Creatinine-plus: Cobas. Rotkreuz, Switzerland: Roche Diagnostics Ltd.
36. Arab L, Wesseling-Perry K, Jardack P, *et al.* (2010) Eight self-administered 24-hour dietary recalls using the Internet are feasible in African Americans and Whites: the energetics study. *J Am Diet Assoc* **110**, 857–864.
37. Liu B, Young H, Crowe FL, *et al.* (2011) Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. *Public Health Nutr* **14**, 1998–2005.
38. Subar AF, Kirkpatrick SI, Mittl B, *et al.* (2012) The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. *J Acad Nutr Diet* **112**, 1134–1137.
39. Touvier M, Kesse-Guyot E, Mejean C, *et al.* (2011) Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* **105**, 1055–1064.

40. Zoellner J, Anderson J & Gould SM (2005) Comparative validation of a bilingual interactive multimedia dietary assessment tool. *J Am Diet Assoc* **105**, 1206–1214.
41. Thompson FE, Dixit-Joshi S, Potischman N, *et al.* (2015) Comparison of interviewer-administered and automated self-administered 24-hour dietary recalls in 3 diverse integrated health systems. *Am J Epidemiol* **181**, 970–978.
42. Bingham SA, Gill C, Welch A, *et al.* (1994) Comparison of dietary assessment methods in nutritional epidemiology – weighed records v 24-h recalls, food-frequency questionnaires and estimated-diet records. *Br J Nutr* **72**, 619–643.
43. Crispim SP, de Vries JH, Geelen A, *et al.* (2011) Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres – results from the European Food Consumption Validation (EFCOVAL) study. *Br J Nutr* **105**, 447–458.
44. Crispim SP, Geelen A, de Vries JH, *et al.* (2012) Bias in protein and potassium intake collected with 24-h recalls (EPIC-Soft) is rather comparable across European populations. *Eur J Nutr* **51**, 997–1010.
45. Freedman LS, Commins JM, Moler JE, *et al.* (2014) Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. *Am J Epidemiol* **180**, 172–188.
46. de Vries JH, Zock PL, Mensink RP, *et al.* (1994) Underestimation of energy intake by 3-d records compared with energy intake to maintain body weight in 269 nonobese adults. *Am J Clin Nutr* **60**, 855–860.
47. Darvishi L, Rabbani Z, Goodarzy S, *et al.* (2012) Investigating and comparing energy and macronutrient intake in female aerobic athletes in two different socio-economic regions. *J Educ Health Promot* **1**, 40.
48. van Erp-Baart AM, Saris WH, Binkhorst RA, *et al.* (1989) Nationwide survey on nutritional habits in elite athletes. Part I. Energy, carbohydrate, protein, and fat intake. *Int J Sports Med* **10**, Suppl. 1, S3–S10.
49. Phillips SM (2004) Protein requirements and supplementation in strength sports. *Nutrition* **20**, 689–695.
50. Kim H, Lee S & Choue R (2011) Metabolic responses to high protein diet in Korean elite bodybuilders with high-intensity resistance exercise. *J Int Soc Sports Nutr* **8**, 10.
51. Banfi G & Del Fabbro M (2006) Serum creatinine values in elite athletes competing in 8 different sports: comparison with sedentary people. *Clin Chem* **52**, 330–331.
52. Cobas (2008) Creatinine-plus, vol. 11, V 14. Rotkreuz, Switzerland: Roche Diagnostics Ltd.
53. Witt KA & Bush EA (2005) College athletes with an elevated body mass index often have a high upper arm muscle area, but not elevated triceps and subscapular skinfolds. *J Am Diet Assoc* **105**, 599–602.
54. Banfi G & Del Fabbro M (2006) Relation between serum creatinine and body mass index in elite athletes of different sport disciplines. *Br J Sports Med* **40**, 675–678.
55. Milic R, Banfi G, Del Fabbro M, *et al.* (2011) Serum creatinine concentrations in male and female elite swimmers. Correlation with body mass index and evaluation of estimated glomerular filtration rate. *Clin Chem Lab Med* **49**, 285–289.
56. Subar AF, Midthune D, Tasevska N, *et al.* (2013) Checking for completeness of 24-h urine collection using para-amino benzoic acid not necessary in the Observing Protein and Energy Nutrition study. *Eur J Clin Nutr* **67**, 863–867.
57. Siebelink E, deVries JHM, Trijsburg L, *et al.* (2015) Evaluation of calculated energy and macronutrient contents of diets provided in controlled dietary intervention trials by chemical analysis of duplicate portions. *J Food Compos Anal* **11**.