British Journal of Nutrition (2012), **108**, 1118–1125
doi:10.1017/S0007114511006295

Predicting urinary creatinine excretion and its usefulness to identify incomplete 24 h urine collections†

Willem De Keyzer[1,2]*, Inge Huybrechts[2], Arnold L. M. Dekkers[3], Anouk Geelen[4], Sandra Crispim[4], Paul J. M. Hulshof[4], Lene F. Andersen[5], Irena Řehůřková[6], Jiří Ruprich[6], Jean-Luc Volatier[7], Georges Van Maele[8], Nadia Slimani[9], Pieter van't Veer[4], Evelien de Boer[3] and Stefaan De Henauw[1,2] on behalf of the EFCOVAL consortium

[1]*Department of Nutrition and Dietetics, Faculty of Health Care Vesalius, University College Ghent, Keramiekstraat 80, B-9000 Ghent, Belgium*
[2]*Department of Public Health, University Hospital, Ghent University, 2 Blok A, De Pintelaan 185, B-9000 Ghent, Belgium*
[3]*National Institute for Public Health and the Environment (RIVM), PO Box 1, Bilthoven, 3720 BA, The Netherlands*
[4]*Division of Human Nutrition, Wageningen University, Bomenweg 2, Wageningen, 6703 HD, The Netherlands*
[5]*Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, PO Box 1046, 0316 Oslo, Norway*
[6]*Department for Food Safety and Nutrition, National Institute of Public Health, Palackeho 1-3, 61242 Brno, Czech Republic*
[7]*Office of Scientific Support for Risk Assessment, French Agency for Food, Environmental and Occupational Health Safety (ANSES), 27–31 Avenue du Général Leclerc, F-94701 Maisons-Alfort Cedex, France*
[8]*Department of Medical Informatics and Statistics, University Hospital Ghent, Zwijnaardsesteenweg 314, Blok F, B-9000 Ghent, Belgium*
[9]*International Agency for Research on Cancer (IARC), 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France*

(Submitted 31 May 2011 – Final revision received 7 September 2011 – Accepted 8 October 2011 – First published online 5 December 2011)

Abstract

Studies using 24 h urine collections need to incorporate ways to validate the completeness of the urine samples. Models to predict urinary creatinine excretion (UCE) have been developed for this purpose; however, information on their usefulness to identify incomplete urine collections is limited. We aimed to develop a model for predicting UCE and to assess the performance of a creatinine index using *para*-aminobenzoic acid (PABA) as a reference. Data were taken from the European Food Consumption Validation study comprising two non-consecutive 24 h urine collections from 600 subjects in five European countries. Data from one collection were used to build a multiple linear regression model to predict UCE, and data from the other collection were used for performance testing of a creatinine index-based strategy to identify incomplete collections. Multiple linear regression (n 458) of UCE showed a significant positive association for body weight ($\beta = 0{\cdot}07$), the interaction term sex × weight ($\beta = 0{\cdot}09$, reference women) and protein intake ($\beta = 0{\cdot}02$). A significant negative association was found for age ($\beta = -0{\cdot}09$) and sex ($\beta = -3{\cdot}14$, reference women). An index of observed-to-predicted creatinine resulted in a sensitivity to identify incomplete collections of 0·06 (95 % CI 0·01, 0·20) and 0·11 (95 % CI 0·03, 0·22) in men and women, respectively. Specificity was 0·97 (95 % CI 0·97, 0·98) in men and 0·98 (95 % CI 0·98, 0·99) in women. The present study shows that UCE can be predicted from weight, age and sex. However, the results revealed that a creatinine index based on these predictions is not sufficiently sensitive to exclude incomplete 24 h urine collections.

Key words: Creatinine index: *Para*-aminobenzoic acid: European Food Consumption Validation: Sensitivity: Specificity

When biomarkers from 24 h urine collections are used in validation studies of dietary assessment methods or in clinical investigations, it is clearly inappropriate to use incomplete urine collections. Currently, two markers are available to check for a complete urine collection; urinary creatinine excretion (UCE) and *para*-aminobenzoic acid (PABA).

In the past decades, scientific disagreement has been reported on the appropriateness of UCE to serve as a marker for identifying incomplete urine collections[1–6]. Furthermore, different strategies for excluding urine samples based on creatinine excretion have been published[3,7–10]. One strategy to exclude specimens is based on the ratio of

Abbreviations: EFCOVAL, European Food Consumption Validation; PABA, *para*-aminobenzoic acid; UCE, urinary creatinine excretion.

* **Corresponding author:** W. De Keyzer, fax +32 9 220 17 26, email willem.dekeyzer@hogent.be

† Part of this work was presented at the 2nd World Congress of Public Health Nutrition in Porto – Portugal (23–25 September 2010).

observed over expected UCE (creatinine index)[10]. Expected UCE (mg/d) is derived from the observed weight of the person and calculated as: body weight (kg) × 24 (males) or 21 (females)[10]. According to Joossens & Geboers[10], urine collections with a creatinine index falling outside the range of 0·6–1·4 should be discarded.

Because some investigators argued that creatinine excretion is not a reliable marker to detect incomplete urine collections, use of the exogenous marker PABA has been proposed[11]. As PABA is actively absorbed and rapidly and completely excreted in urine, it is suitable to verify completeness of urine collections based on the simple principle of an expected minimal PABA recovery in sampled urine[12].

Currently, there are no recent data available on creatinine excretion in a large European sample, or on its potential value to serve as a marker for detection of incomplete urine collections. Also, models to predict UCE have been published before, although information on their performance to correctly identify incomplete urine samples is limited.

The objective of the present study is, first, to develop a multiple linear regression model which is able to predict UCE using readily available subject information and, secondly, to investigate the suitability of UCE as a marker for identifying incomplete 24 h urine collections using PABA as a reference.

Subjects and methods

Subjects

The study group comprised 600 healthy European adults (297 males and 303 females aged 45–65 years) participating in the European Food Consumption Validation (EFCOVAL) study. Data were collected in five European countries: Belgium, the Czech Republic, France (Southern part), the Netherlands and Norway. In all countries, a convenience sample was recruited targeting equal numbers of men and women, and inclusion of subjects from lower, intermediate and higher educational levels. Exclusion criteria comprised the use of diuretics, receiving diet therapy, simultaneous participation in another study, pregnancy or lactation, having diabetes mellitus or a kidney disease and donation of blood or plasma during (or less than 4 weeks before) the study. In addition, because of PABA administration during urine collections, use of sulphonamide-based antibiotics or acetaminophen painkillers (e.g. paracetamol) was not allowed and subjects hypersensitive to sulphonamides or PABA were also excluded.

Urine collections

All subjects were carefully instructed to keep two 24 h urine collections according to a standardised protocol. Days of urine collection were randomly assigned, though with a time interval of at least 4 weeks between the two collections. Subjects were asked to urinate upon rising in the morning; this micturition was completely discarded and the time was registered. This time was taken as the starting point of the 24 h urine collection. Subsequently, all urine produced during the next 24 h was collected up to, and including, the first voiding of the following day. Subjects were provided a diary to register the time of rising, observations (e.g. use of medication) and possible deviations (e.g. missing urine) from the urine collection protocol. All urine voidings were stored in light-protected receptacles containing boric acid (3 g/2 litre bottle) as a preservative. In order to verify the completeness of urine collections, an 80 mg PABA tablet (PABAcheck, Laboratories for Applied Biology) was taken three times over the day, orally during a meal. The underlying assumption to use PABA as a marker for detection of incomplete urine collections is that it is excreted almost quantitatively in 24 h. Hence, we expected that approximately 240 mg of PABA would be present in every 24 h urine sample[11]. After 1 month, the same procedure was repeated so that every subject yielded two 24 h urine collections. Subjects' body weight (kg) and height (cm) were measured at the study centres before the first urine collection (wearing light clothes and no shoes). This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving human subjects were approved by the ethical committees in each country involved in the data collection. Written informed consent was obtained from all subjects.

Quantification of urinary analytes

At the study centre, urine samples were weighed and well mixed before aliquoting into 10 ml cryo-storage tubes. Aliquots were then frozen at −20°C until shipment on dry ice to the central laboratory at Wageningen University (the Netherlands), where they were kept at the same temperature. On the day of chemical analysis, aliquots were rapidly thawed at room temperature. PABA was measured using the colorimetric diazocoupling method described by Bingham & Cummings[11]. The method is based on total quantification of aromatic amines derived from PABA and its metabolites after alkaline hydrolysis. Urinary creatinine concentrations were measured at 520 nm on a Synchron LX20 (Beckman Coulter) using a commercial kit which is a modification of the kinetic Jaffé procedure.

Stability of PABA in urine was verified by storing urine samples of three participants at four different temperatures (−20, 6, 20 and 30°C) for 8 d. PABA concentrations were measured at five moments (days 0, 1, 2, 4 and 7). No significant changes in PABA concentrations were observed during the storage period at each temperature. In addition, previous studies have shown that in all but extreme cases (e.g. storage at 55°C for 30 d), urinary creatinine is virtually unaffected by storage time and temperature in both acidified as in unpreserved urine samples[13,14].

Dietary data

Two 24 h recalls were collected using EPIC-Soft software (version 9.16; IARC). The structure and standardisation procedure of EPIC-Soft have been described elsewhere[15,16]. Briefly, EPIC-Soft is an assisted dietary tool that follows standardised steps when describing, quantifying, probing and calculating food intakes across countries[15]. Dietary recalls followed a

randomised schedule, which included all days of the week. Protein intake was calculated using mainly the country-specific food composition tables and missing information from a food was gathered from another similar food or another food composition tables. The protein coverage of national food composition tables was 100 % for all foods reported.

Statistical analysis

A stepwise regression approach was used for developing and validating a model for creatinine prediction. Data from the second urine collection were used for model development, and data from the first collection for model validation. The first collection was chosen to validate the model because this corresponds with data that are available if only one urine collection would be performed. When the first urine collection was used for model development, the results did not differ significantly from the present ones.

First, multiple regression with UCE from the second urine collection as dependent variable was performed. The independent variables – age, sex, weight, height and an interaction term sex × weight – were entered in the model. The interaction term was included because regression of weight to UCE showed different slopes for both sexes. Because protein-rich foods are able to influence UCE[17] (arginine and glycine are amino acid precursors of creatinine, and creatine is present in meat), mean protein intake derived from the dietary intake data was also included as an explanatory variable. Only cases with complete urine collections according to PABA recovery ≥ 85 % were used to develop a model for predicting UCE. Cook's distance and studentised residuals were used to analyse residuals and to assess the influence of outliers.

Second, the predictive ability of the model was evaluated. Therefore, expected UCE was calculated for every participant by using the regression equation of the model. The standard error of the model was used to calculate 95 % CI around predicted UCE. Next, the percentage of actual values for UCE from collection 1 that were located within 95 % CI of the predicted UCE was calculated.

Subsequently, the suitability of creatinine to serve as a marker for identifying incomplete urine collections was investigated. Therefore, the performance of the ratio observed over expected UCE was tested using PABA recovery as a reference. The two ratios were calculated and compared. For the first ratio, expected UCE was body weight-derived as proposed by Joossens & Geboers[10]. In the second ratio, expected UCE was derived from the regression equation of the final regression model. For both ratios, collections were classified as incomplete using the cut-off by Murakami *et al.*[3] (creatinine index < 0·7 = incomplete collection). Finally, sensitivity was calculated as the proportion of participants with incomplete urine collections correctly classified, and specificity as the proportion of participants with complete urine collections correctly classified. For comparison of groups, the likelihood ratio of a positive test was also calculated (sensitivity/(1 − specificity)). For all figures, 95 % CI are reported.

Results

Participant characteristics are listed in Table 1. The mean age of the participants was 55·2 (SD 5·7) years for men and 54·7 (SD 5·8) years for women. Mean BMI was, respectively, 26·7 (SD 3·5) and 24·7 (SD 4·1) kg/m^2 for men and women. In men, 34 % of the participants had a BMI below 25 kg/m^2, and in women this was 61 %. All participants (n 600) collected their urine the first time; but two participants failed to perform the second collection. Diaries were analysed to inventory deviations from the urine collection protocol. Overall, four and ten participants reported not having taken all three PABA tablets during collections 1 and 2, respectively. Prohibited medication intake during urine collections 1 and 2 was reported by two and six participants, respectively. During collections 1 and 2, respectively, fifty-four and fifty-eight participants reported that not all urine could be collected. Finally, in three urine specimens (two from collection 1; one from collection 2), creatinine could not be determined. Taking all deviations from the first and second urine collections and analysis protocol into account, respectively, 541 (90·2 %) and 528 (88·3 %) specimens were available for data analysis. Mean self-reported collection time was close to 24 h (23·90 h for men during both collections, 23·9 and 24 h during the first and second collections for women; Table 2). PABA recovery was proximate to 100 % (range of means: 98·0 (SD 15·6) to 102·2 (SD 11·7) %), and 86·2 % (women during the second urine collection) to 93·5 % (men during the first urine collection) of PABA recoveries had values equal to or above 85 %. Mean weight-adjusted creatinine excretion was 21 mg/kg in men and 17 mg/kg in women.

For the prediction of UCE, the multiple regression model showed a significant positive association for the interaction term sex × weight ($\beta = 0{\cdot}09$, $P < 0{\cdot}001$), weight ($\beta = 0{\cdot}07$, $P < 0{\cdot}001$) and protein intake ($\beta = 0{\cdot}02$, $P < 0{\cdot}001$). A significant negative association was found for age ($\beta = -0{\cdot}09$, $P < 0{\cdot}001$) and sex ($\beta = -3{\cdot}14$, $P = 0{\cdot}006$; Table 3). The variables in the model explained more than three-quarters (adjusted R^2 0·76) of the variance in predicted UCE with a

Table 1. Characteristics of participants

(Mean values and standard deviations; percentages and number of participants)

	Men (n 297)		Women (n 303)	
	Mean	SD	Mean	SD
Age (years)	55·2	5·7	54·7	5·8
Weight (kg)	83·6	12·5	67·0	11·9
Height (cm)	177	7·5	165	7·4
BMI (kg/m^2)	26·7	3·5	24·7	4·1
	n	%	n	%
BMI category (kg/m^2)				
< 25	100	34	185	61
25–29·9	151	51	89	29
≥ 30	46	15	29	10
Educational level				
Low	50	17	66	22
Intermediate	73	24	96	32
High	174	59	141	47

Table 2. Characteristics of urine collections*

(Mean values and standard deviations; percentages and number of participants)

	Urine collection							
	First				Second			
	Men (*n* 278)		Women (*n* 263)		Men (*n* 267)		Women (*n* 261)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Total urine volume (ml/d)	2114	849	2336	911	2105	817	2313	944
Reported collection time (h)	23·9	1·1	23·9	1·1	23·9	1·0	24·0	0·7
Urinary PABA (mg/d)	245·2	28·0	239·5	30·9	240·3	39·0	234·5	37·0
PABA recovery (%)	102·2	11·7	99·6	13·2	100·1	16·2	98·0	15·6
< 85 %								
n	18		19		26		36	
%	6·5		7·2		9·7		13·8	
≥ 85 %								
n	260		244		241		225	
%	93·5		92·8		90·3		86·2	
Urinary creatinine (mmol/l)	8·6	3·6	5·0	2·5	8·5	3·6	5·0	2·3
Total urinary creatinine (mmol/d)	15·7	3·2	9·9	1·9	15·6	3·3	9·9	2·0
Urinary creatinine:body weight (mg/kg)	21·4	3·6	17·1	3·2	21·2	3·6	17·1	3·5

PABA, *para*-aminobenzoic acid.

* Only collections that fully complied with the protocol are presented (total n_{first} 541, total n_{second} 528).

correlation coefficient of 0·87. The standard error of the estimate was 1·884 mmol/d. UCE values from seven participants were identified as outliers and excluded before model development, and from one participant no dietary intake data were available (*n* 458).

When UCE measurements from the first collection were compared to the predicted UCE using the regression equation, 93·5 % of the measurements fell within the 95 % CI of the prediction.

Sensitivity and specificity of UCE and the likelihood ratio of a positive test for identification of incomplete urine collectors are reported in Table 4 using different ratios. First, the performance of the strategy proposed by Murakami *et al.*[3] is presented. When a cut-off of 0·7 was used for the creatinine index, sensitivity and specificity was, respectively, 0·49 (95 % CI 0·34, 0·63) and 0·88 (95 % CI 0·84, 0·89). For females, sensitivity was higher (0·63) and specificity lower (0·81) compared to the total group. In males, sensitivity was markedly lower (0·33) and specificity slightly higher (0·94) compared to total group figures. The likelihood of an incomplete urine collection was increased 4-fold, given a positive test result (5·8 and 3·4 for males and females, respectively). When the estimated UCE was calculated using the regression equation and the same cut-off for detection of incomplete urine collections was used, sensitivity was very low (0·08) and specificity almost excellent (0·98). Explorations of other cut-offs accompanied with their respective measures for sensitivity, specificity and positive likelihood ratios are presented in Table 5. Finally, Fig. 1 shows a scatter-plot from observations for PABA recoveries and ratios of observed-to-predicted creatinine. Both cut-offs for the detection of incomplete urine collections by PABA and creatinine ratio are indicated by a horizontal and vertical line, respectively. The large scatter shows that shifting the cut-off of the ratio observed over expected creatinine to the right (i.e. using a higher cut-off) will lead to an important increase of false-positive results.

Discussion

In the present study, measures from duplicate 24 h urine collections are used to develop a prediction model for estimating urinary creatinine excretion on one hand and to test the performance of urinary creatinine to serve as an indicator for detecting incomplete urine collections on the other. For the latter, PABA is used as a reference and is assumed to correctly identify incomplete urines. The model presented in this paper is able to explain up to 76 % of the variance in urinary creatinine excretion when using sex, age, weight and sex × weight as independent variables. Although 93·5 % of the observed urinary creatinine values from urine collection 1 fell within 95 % CI of the model predictions from urine collection 2, the sensitivity of the ratio of observed-to-predicted creatinine was very low when using traditional cut-offs for the creatinine index. Sensitivity improved after increasing the cut-off (e.g. to 0·95); however, given the trade-off between sensitivity and specificity, this also decreases specificity.

Table 3. Model coefficients of multiple linear regression*

	β			
UCE (mmol/d)†	β	SE	95 % CI	*P*
Constant	9·36	1·132	7·136, 11·585	< 0·001
Sex‡	−3·14	1·141	−5·377, −0·892	0·006
Weight (kg)	0·07	0·011	0·048, 0·093	< 0·001
Sex × weight	0·09	0·015	0·058, 0·118	< 0·001
Age (years)	−0·09	0·015	−0·123, −0·063	< 0·001
Protein (g/d)	0·02	0·003	0·009, 0·021	< 0·001

UCE, urinary creatinine excretion; SE, sensitivity.

* Model based on data from the second urine collection. Only collections that fully complied with the protocol and *para*-aminobenzoic acid recoveries equal to or above 85 % were included. Overall, seven cases were identified as outliers based on studentised residuals and Cook's distance and excluded from the model.

† Model summary: *R* 0·87, adjusted R^2 0·76, SE estimated 1·884 mmol/d, *n* 458.

‡ Sex = 0 for females, 1 for males.

British Journal of Nutrition

Table 4. Number of participants with incomplete and complete 24 h urine collections by *para*-aminobenzoic acid (PABA) and two test strategies, and sensitivity (SE), specificity (SP) and positive likelihood ratios (LR+) of both test strategies for identifying incomplete 24 h urine collections
(Number of participants and 95 % confidence intervals)

	Completeness of 24 h urine collection by PABA										
	Incomplete		Complete								
	Completeness of 24 h urine collection by test strategy		Completeness of 24 h urine collection by test strategy								
Test strategy	Incomplete	Complete	Incomplete	Complete	*n*	SE	95 % CI	SP	95 % CI	LR+	95 % CI
Murakami *et al.*[3]*											
All	18	19	61	443	541	0·49	0·34, 0·63	0·88	0·87, 0·89	4·0	2·6, 5·7
Male	6	12	15	245	278	0·33	0·17, 0·53	0·94	0·93, 0·96	5·8	2·5, 11·9
Female	12	7	46	198	263	0·63	0·42, 0·80	0·81	0·80, 0·83	3·4	2·1, 4·6
Regression model†											
All	3	34	11	493	541	0·08	0·03, 0·17	0·98	0·97, 0·99	3·7	1·1, 11·6
Male	1	17	7	253	278	0·06	0·01, 0·20	0·97	0·97, 0·98	2·1	0·3, 11·7
Female	2	17	4	240	263	0·11	0·03, 0·22	0·98	0·98, 0·99	6·4	1·4, 28·3

* Participants with a value for (urinary creatinine (mg/d)/(G × body weight (kg)) of < 0·7 were identified as having incomplete urine collections. G = 21 for females and 24 for males.
† Ratio of observed:expected urinary creatinine excretion. Expected urinary creatinine excretion was calculated as (9·36 − 3·14 × sex + 0·07 × weight + 0·09 × sex × weight − 0·09 × age + 0·02 × protein intake). Participants with a ratio < 0·7 were identified as having incomplete urine collections.

In a study performed by Murakami *et al.*[3], the sensitivity and specificity of different strategies using creatinine as an indicator for identification of incomplete urine collections against PABA were investigated. The strategy using a creatinine index with a cut-off of 0·7 was considered to be potentially useful for identifying incomplete 24 h urine collections. The sensitivity and specificity among 654 Japanese girls (mean age 19·7 (SD 1·1) years) were, respectively, 0·47 and 0·99. In the present study, sensitivity and specificity in females using the strategy by Murakami *et al.*[3] was 0·63 and 0·81, respectively. The sample characteristics from this latter study are different as compared to those from the present sample in terms of age, ethnicity and educational diversity.

The sensitivity in men (0·33) was roughly half the sensitivity in women (0·63) in the present study. A lower sensitivity in men compared to women was also demonstrated in a small-scale study (*n* 83), comprising data from six European countries, performed by Knuiman *et al.*[2], where sensitivity was 0·07 for men compared to 0·13 for women. It is noteworthy to mention that in their study a different cut-off (i.e. 0·6) was used. When a cut-off of 0·6 was used on the present data, sensitivity was 0·05 for men and 0·36 for women (data not shown). Regardless of the cut-off or study sample, it is clear that when exclusion of collections is based on UCE, a large part of incomplete collections remains unidentified because of low sensitivity, resulting in an underestimation of urine-based analytes used during further analysis.

In this study, the fraction of incomplete collections was 6·5 and 9·7 % for the first and second collections in men, and 7·2 and 13·8 % in women, respectively. Earlier studies have reported prevalences of incomplete collections similar to the ones reported here[3,5,18,19]. A possible explanation for the

Table 5. Sensitivity (SE), specificity (SP) and positive likelihood ratios (LR+) from a range of cut-offs from observed to expected creatinine calculated by the regression model*

	All			Male			Female		
Cut-off	SE	SP	LR+	SE	SP	LR+	SE	SP	LR+
0·50	0·00	1·00	0·0	0·00	1·00	0·0	0·00	1·00	–
0·55	0·00	0·99	0·0	0·00	0·99	0·0	0·00	1·00	0·0
0·60	0·05	0·99	6·8	0·00	0·99	0·0	0·05	1·00	12·8
0·65	0·05	0·99	4·5	0·00	0·98	0·0	0·11	1·00	25·6
0·70	0·08	0·98	4·5	0·06	0·97	0·0	0·11	0·98	12·8
0·75	0·22	0·97	6·8	0·17	0·97	5·4	0·26	0·97	8·0
0·80	0·38	0·93	5·6	0·33	0·93	4·8	0·42	0·94	6·8
0·85	0·49	0·87	3·9	0·39	0·88	3·3	0·58	0·87	4·4
0·90	0·65	0·77	2·8	0·56	0·81	2·9	0·74	0·73	2·7
0·95	0·81	0·61	2·1	0·78	0·62	2·1	0·84	0·60	2·1
1·00	0·89	0·47	1·7	0·94	0·49	1·9	0·84	0·45	1·5
1·05	0·95	0·33	1·4	0·94	0·34	1·4	0·95	0·32	1·4
1·10	0·97	0·22	1·3	0·94	0·22	1·2	1·00	0·23	1·3

* Regression equation of the model: expected urinary creatinine excretion (mg/d) = 9·36 − 3·14 × sex + 0·07 × weight + 0·09 × sex × weight − 0·09 × age + 0·02 × protein intake (sex: male = 1, female = 0; weight (kg); age (years); protein intake (g/d)).

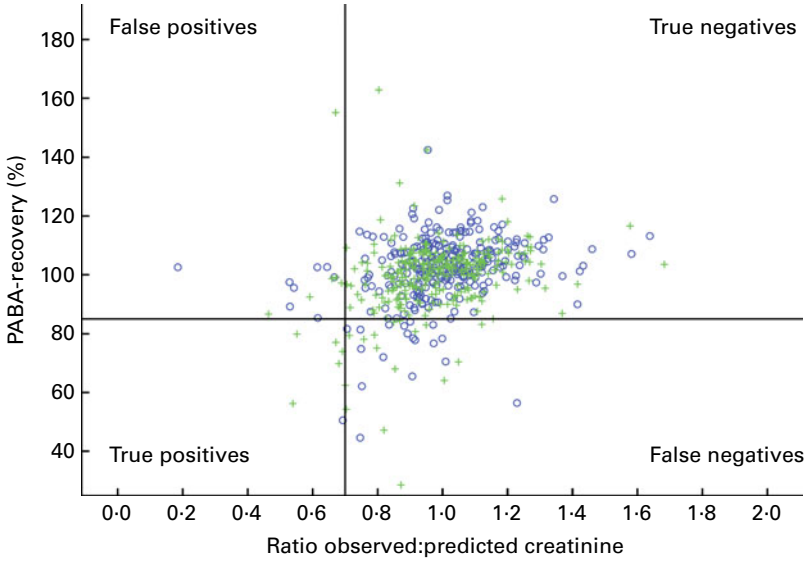

Fig. 1. Comparing *para*-aminobenzoic acid (PABA)-recovery and creatinine ratio cut-offs in their ability to identify incomplete urine collections in our sample of 541 adults. The horizontal line marks 85 % PABA-recovery, the vertical line represents the cut-off for the ratio of observed-to-predicted creatinine $<0{\cdot}7$ from Murakami *et al.*[3]. ○, Male; +, female.

higher prevalence of incomplete urines in the second collection might be that participants did not follow the instructions as consequent as the first time. Also, the collection protocol was not explained to participants for the second collection, which might have decreased the attention or motivation of participants to collect all urine. Therefore, repeating of instructions and stressing the necessity of complete urine collections is advised when multiple collections are requested. Mean PABA recoveries (ranging from 98·0 % in women to 102·2 % in men) were close to 100 % and in agreement with those from past observations in a European sample[2] and somewhat lower than the one reported by Murakami *et al.*[3] in a Japanese sample (103·8 % after adjustment for self-reported missed urine and collection time). In the literature, models to predict creatinine have been described before[20–23] and reported correlations between observed and predicted UCE range from 0·50 to 0·64[21–23]. However, no reference method was used to assess the performance of these models. From the present study, it can be concluded that when an internal standard like PABA is used to check for completeness of urine collections, sensitivity of urinary creatinine to identify incomplete urine collections is very disappointing. This is the case for both cut-offs based on simple calculations as for more advanced predictions based on multiple regression modelling. However, both the ratio of observed-to-predicted creatinine as total UCE from incomplete collections are significantly lower than those from complete collections (data not shown).

For model development, data from the second urine collection were used. Subsequently, UCE from the first urine collection was compared to predicted UCE using the equation generated by the model. This way, both model development sample and model validation sample were large, comprising 458 and 504 cases, respectively. As mentioned before, there was only a difference in prevalence of incomplete collections between the first and second urine collections. Because incomplete urine collections were excluded during both model development and validation, this difference was removed, making both collections equal. Due to the fact that between-person variance in UCE is higher than within-person variance, the high number of participants in the present study is a major strength ($CV_{between}$: 19·1 % for males, 18·3 % for females; CV_{within}: 9·1 % for males, 9·2 % for females, in the present study). Also, body weight of participants was accurately weighed instead of self-reported. In addition, an internal standard (PABA) was used to identify incomplete urine collections. Nevertheless, PABA has also some limitations. First, over-collection of urine cannot be detected. For instance, participants might collect the first urine voiding instead of discarding it. This will dilute urine concentration leading to underestimations of its constituents. Second, impaired renal function leads to under-collection. Therefore, kidney disease was an exclusion criterion for enrolment in the study. Third, it has been suggested that late night meals interfere with PABA-resorption due to decreasing of metabolism and excretion of PABA[2]. This problem was overcome in the present study by asking participants to take the PABA tablet before 19.00 hours. Fourth, when the colorimetric method is used for PABA-analysis, aromatic amines are also determined[24]. Because these compounds can originate from drugs like paracetamol and sulphonamide, use of these drugs was prohibited during study participation. Fifth, previous studies[10,11] have shown excretions of PABA (15–24 mg) originating from foods; so determination of baseline PABA excretion, at least in a subsample, is advised because any natural presence of PABA can result in an underestimation of incomplete urine collections. Sixth, a major drawback of PABA use is that intake can be refused or forgotten by study participants. Also, it can be considered

https://doi.org/10.1017/S0007114511006295 Published online by Cambridge University Press

inappropriate because of unknown interferences with biological samples of interest (e.g. genetic studies).

No self-reported dietary assessment method is capable of capturing all protein intakes. When comparing protein intake to urinary nitrogen, the EFCOVAL study found an underestimation of the EPIC-Soft guided 24 h recall varying from 1·4 % up to 14·7 % in men and 1·8 % up to 12·2 % in women after adjusting for age, BMI and educational level[25]. Because of the small influence of dietary protein intake on UCE[17], and the fact that underestimation of protein intake is not large, the anticipated consequences for the present results are assumed to be limited.

Development of a model for the prediction of urinary creatinine based on anthropometrics logically implies the inclusion of participants expected to deliver complete urine collections only. Application of this method while assessing the performance of a test strategy, however, will yield a lower prevalence of incomplete urine collections, thereby potentially affecting sensitivity and specificity calculations. Additional analyses of our data nevertheless demonstrated a limited impact of including fifty-four participants reporting incomplete urine collections during the first collection, as only four of these also showed to be catalogued as such by the reference method (data not shown).

Finally, adjustments for missed urine have been proposed elsewhere[3], as have corrections for collections with PABA recoveries below 85 %[26]. Given that one of the objectives of the present study was to report on performance of strategies to deal with suspicious urine collections in terms of completeness, it was chosen not to adjust nor correct data and to keep them as straightforward as possible, so that consequences for inferences could be minimised.

The present study has shown that UCE can be predicted from readily available subject information like sex, weight and age. For the first time, the performance of a prediction-based creatinine index in terms of sensitivity and specificity was assessed against a reference method. When such a creatinine index is used to exclude incomplete urines, sensitivity analysis showed that 94–89 % of incomplete 24 h urine collections remain unidentified in males and females, respectively. Also, in this European sample, sensitivity to identify incomplete urine collections of a traditional creatinine index was only 0·33 (95 % CI 0·17, 0·53) and 0·66 (95 % CI 0·42, 0·80) in males and females, respectively. Therefore, based on the present findings, both the prediction-based as the traditional creatinine index can be considered as an unreliable marker for detection of incomplete urine collections when compared to PABA.

Acknowledgements

The EFCOVAL project is supported by European Community funding under the Sixth Framework Program (FOOD-CT-2006-022895); FWO Flanders (Fonds Wetenschappelijk Onderzoek– Vlaanderen, Krediet aan navorsers – 1519110N) and the Research Fund of University College Ghent (WDK) are also acknowledged for their financial contribution. The authors thank the EFCOVAL partners for their useful advice. The EFCOVAL partners are Ghent University (DPH), Belgium; Academy of Medical Sciences (AMZH), Croatia; National Institute of Public Health (NIPH), Czech Republic; National Food Institute, Technical University of Denmark (DTU), Denmark; French Food Safety Authority (AFSSA), France; National Institute for Agricultural Research (INRA), France; German Institute of Human Nutrition (DIfE), Germany; National Research Institute for Food and Nutrition (INRAN), Italy; Wageningen University (WU), The Netherlands; National Institute for Public Health and the Environment (RIVM), The Netherlands; University of Oslo, Norway; Basque Foundation for Health Innovation and Research (BIOEF), Spain; Prima informatics limited (Primainfo), United Kingdom; and International Agency for Research on Cancer (IARC, WHO). E. d. B., N. S. and P. v. V. designed the research; P. J. M. H. was responsible for laboratory analysis; W. D. K., A. L. M. D. and G. V. M. analysed the data; and W. D. K. wrote the paper. All authors read and approved the final manuscript. The present document reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained therein. The authors declare that there are no competing or financial conflicts of interest.

References

1. Bingham SA & Cummings JH (1985) The use of creatinine output as a check on the completeness of 24-hour urine collections. *Hum Nutr Clin Nutr* **39**, 343–353.
2. Knuiman JT, Hautvast JG, van der Heyden L, *et al.* (1986) A multi-centre study on completeness of urine collection in 11 European centres. I. Some problems with the use of creatinine and 4-aminobenzoic acid as markers of the completeness of collection. *Hum Nutr Clin Nutr* **40**, 229–237.
3. Murakami K, Sasaki S, Takahashi Y, *et al.* (2008) Sensitivity and specificity of published strategies using urinary creatinine to identify incomplete 24-h urine collection. *Nutrition* **24**, 16–22.
4. 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.
5. Williams DR & Bingham SA (1986) Sodium and potassium intakes in a representative population sample: estimation from 24 h urine collections known to be complete in a Cambridgeshire village. *Br J Nutr* **55**, 13–22.
6. Edwards OM, Bayliss RIS & Millen S (1969) Urinary creatinine excretion as an index of the completeness of 24-hour urine collections. *Lancet* **294**, 1165–1166.
7. Malekshah AF, Kimiagar M, Saadatian-Elahi M, *et al.* (2006) Validity and reliability of a new food frequency questionnaire compared to 24 h recalls and biochemical measurements: pilot phase of Golestan cohort study of esophageal cancer. *Eur J Clin Nutr* **60**, 971–977.
8. Reinivuo H, Valsta LM, Laatikainen T, *et al.* (2006) Sodium in the Finnish diet: II. Trends in dietary sodium intake and comparison between intake and 24-h excretion of sodium. *Eur J Clin Nutr* **60**, 1160–1167.
9. World Health Organization (1984) *Estimation of Sodium Intake and Output: Review of Methods and Recommendations for Epidemiological Studies*. Geneva: WHO.
10. Joossens JV & Geboers J (1984) Monitoring salt intake of the population: methodological considerations. In *Surveillance of the Dietary Habits of the Population with Regard to Cardiovascular Diseases, EURONUT Report 2*, pp. 61–73

[GG De Backer, HT Pedoe and P Ducimetiere, editors]. Wageningen: Department of Human Nutrition, Agricultural University.
11. 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.
12. Bingham SA (2003) Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* **133**, Suppl. 3, 921S–924S.
13. Miki K & Sudo A (1998) Effect of urine pH, storage time, and temperature on stability of catecholamines, cortisol, and creatinine. *Clin Chem* **44**, 1759–1762.
14. Spierto FW, Hannon WH, Gunter EW, *et al.* (1997) Stability of urine creatinine. *Clinica Chimica Acta* **264**, 227–232.
15. Slimani N, Deharveng G, Charrondiere RU, *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* **58**, 251–266.
16. Slimani N, Ferrari P, Ocke M, *et al.* (2000) Standardization of the 24-hour diet recall calibration method used in the European Prospective Investigation into Cancer and Nutrition (EPIC): general concepts and preliminary results. *Eur J Clin Nutr* **54**, 900–917.
17. Gibson RS (2005) Assessment of protein status. In *Principles of Nutritional Assessment* [RS Gibson, editor]. New York, NY: Oxford University Press.
18. Leclercq C, Maiani G, Polito A, *et al.* (1991) Use of PABA test to check completeness of 24-h urine collections in elderly subjects. *Nutrition* **7**, 350–354.
19. Toft U, Kristoffersen L, Ladelund S, *et al.* (2007) Relative validity of a food frequency questionnaire used in the Inter99 study. *Eur J Clin Nutr* **62**, 1038–1046.
20. Kroos DS, Mays JE & Harris SA (2010) A model to predict 24-h urinary creatinine using repeated measurements in an occupational cohort study. *J Expo Sci Environ Epidemiol* **20**, 516–525.
21. Kawasaki T, Uezono K, Itoh K, *et al.* (1991) Prediction of 24-hour urinary creatinine excretion from age, body weight and height of an individual and its application. *Nippon Koshu Eisei Zasshi* **38**, 567–574.
22. Turner WJ (1975) Total body potassium and 24-hour creatinine excretion in healthy males. *Clin Pharmacol Ther* **18**, 405–412.
23. Dodge WF, Travis LB & Daeschner CW (1967) Comparison of endogenous creatinine clearance with inulin clearance. *Am J Dis Child* **113**, 683–692.
24. Jakobsen J, Ovesen L, Fagt S, *et al.* (1997) *Para*-aminobenzoic acid used as a marker for completeness of 24 hour urine: assessment of control limits for a specific HPLC method. *Eur J Clin Nutr* **51**, 514–519.
25. Crispim SP, de Vries JHM, 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.
26. Johansson G, Bingham S & Vahter M (1999) A method to compensate for incomplete 24-hour urine collections in nutritional epidemiology studies. *Public Health Nutr* **2**, 587–591.