Validation of an automated self-administered 24-hour dietary recall web application against urinary recovery biomarkers in a sample of French-speaking adults of the province of Québec, Canada

Journal:	<i>Applied Physiology, Nutrition, and Metabolism</i>
Manuscript ID	apnm-2021-0445.R1
Manuscript Type:	Article
Date Submitted by the Author:	10-Sep-2021
Complete List of Authors:	Paradis, Frédérique; Laval University, School of Nutrition Lamarche, Benoît; Université Laval, Nutrition Robitaille, Julie; Université Laval, Couillard, Charles; Laval University, INAF Lafrenière, Jacynthe; Université Laval, Institute of Nutrition and Functional Foods Tremblay, André; Laval University, INAF Corneau, Louise; Université Laval, Institute of Nutrition and Functional Foods Lemieux, Simone; Laval University, School of Nutrition
Novelty bullets: points that summarize the key findings in the work:	The validity of an automated self-administered 24-hour dietary recall web application named the R24W was tested using urinary. According to seven criteria, the R24W was found to have a good validity to assess self-reported intakes of sodium, potassium a
Keyword:	24-hour dietary recall, dietary assessment, dietary intakes, recovery biomarkers, 24-hour urine collection, misreporting, individual characteristics
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

SCHOLARONE™
Manuscripts

Title: Validation of an automated self-administered 24-hour dietary recall web application against urinary recovery biomarkers in a sample of French-speaking adults of the province of Québec, Canada

Frédérique Paradis ^{1,2}, Benoît Lamarche ^{1,2}, Julie Robitaille ^{1,2}, Charles Couillard ^{1,2}, Jacynthe Lafrenière ¹, André J. Tremblay¹, Louise Corneau¹ and Simone Lemieux ^{1,2}

¹Centre Nutrition, santé et société (NUTRISS), Institut sur la nutrition et les aliments fonctionnels (INAF), Université Laval, 2440 boulevard Hochelaga, Québec, QC, G1V 0A6, Canada;

²School of Nutrition, Université Laval, 2425 rue de l'Agriculture, Québec, QC, G1V 0A6, Canada.

Corresponding author:

Simone Lemieux, Dt.P., Ph.D.

Simone.lemieux@fsaa.ulaval.ca

Centre NUTRISS, INAF, Université Laval
2440 boulevard Hochelaga
Québec, QC, G1V 0A6
Canada

Short title: Criterion validity of a 24-hour dietary recall application

Abstract

The objective of this study was to validate an automated self-administered 24-hour dietary recall web application (R24W) against recovery biomarkers for sodium, potassium and protein intakes and to identify individual characteristics associated with misreporting in a sample of sixty-one men and sixty-nine women aged 20-65 years from Quebec City, Canada. Each participant completed three dietary recalls using the R24W, provided two 24-hour urinary samples and completed questionnaires to document psychosocial factors. Mean reported intakes were 2.2%, 2.1% and 5.0% lower than the urinary reference values respectively for sodium, potassium and proteins (significant difference for proteins only ($p=0.04$)). Deattenuated correlations between the self-reported intake and biomarkers were significant for sodium ($r: 0.48$), potassium ($r: 0.56$) and proteins ($r: 0.68$). Cross-classification showed that 39.7% (sodium), 42.9% (potassium) and 42.1 % (proteins) of participants were ranked into the same quartile with both methods and only 4.8% (sodium), 3.2% (potassium) and 0.8% (proteins) were ranked in opposite quartiles. Lower body esteem related to appearance was associated with sodium underreporting in women ($r: 0.33$, $p=0.006$). No other individual factor was found to be associated with misreporting. These results suggest that the R24W has a good validity for the assessment of sodium, potassium and protein intakes in a sample of French-speaking adults.

Novelty bullets:

- The validity of an automated self-administered 24-hour dietary recall web application named the R24W was tested using urinary biomarkers.

- According to seven criteria, the R24W was found to have a good validity to assess self-reported intakes of sodium, potassium and proteins.

Keywords: 24-hour dietary recall, dietary assessment, dietary intakes, recovery biomarkers, 24-hour urine collection, misreporting, individual characteristics

Draft

Introduction

In recent decades, epidemiological studies in nutrition have identified many dietary factors associated with morbidity and mortality related to various chronic and metabolic diseases (Micha et al., 2017, U.S. Department of Health and Human Services, 2020). The results and conclusions arising from these studies are, however, the subject of several criticisms from the scientific community. The evaluation of food intake in these studies generally relies on self-assessment methods susceptible to introduce several biases in the measurement of food intake due to random as well as systematic errors (Shim et al., 2014, Trabulsi and Schoeller, 2001, Kant, 2010, Freedman et al., 2014, Freedman et al., 2015). The development of web-based dietary assessment tools, such as self-administered web-based 24-hour dietary recall, is increasingly put forward for many reasons such as the reduction of some bias as well as lower cost and enhanced time effectiveness (Illner et al., 2012).

The automated Self-Administered 24-hour food recall tool (ASA24) developed by the National Cancer Institute and released in 2009 (Zimmerman et al., 2009) has been instrumental in moving forward in the field of automated self-administered dietary assessment. The ASA24 has been used in many studies and has also been extensively validated (Kirkpatrick et al., 2016, Kirkpatrick et al., 2014, Pannucci et al., 2018, Thompson et al., 2015). “Myfood24” is another web-based automated 24-hour food recall tool that was developed to be used in the UK population (Carter et al., 2015) and that has been validated in this population (Wark et al., 2018). The Oxford WebQ is a 24-hour dietary assessment tool that has also been developed in the UK. The Oxford WebQ is inquiring on food and beverages consumed on the previous day by using a finite list of pre-

determined items from 21 food groups (Liu et al., 2011). Therefore, the level of specificity and details that can be obtained with the Oxford WebQ is less than what can be obtained from a 24-hour food recall using a more standard approach.

In 2015, our research team has developed the first French-language fully automated web-based 24-hour dietary recall, named the R24W (Jacques et al., 2016). This application can automatically calculate nutrient intakes based on data from the Canadian Nutrient File and diet quality scores. Among the specific features of the R24W, there is the inclusion in our database of food representative of the diversity of food culture in the population of the Province of Quebec (Canada) as well as home-made recipes found to be the most popular in this population. Similarly to the ASA24 and the “myfood24”, the R24W uses a systematic sequence inspired from the AMPM approach to assess food consumed.

Dietary assessment tools need to be validated using objective reference values such as biomarkers. Recovery biomarkers are completely independent of the biases linked to self-reported measures (Trabulsi and Schoeller, 2001). A validation analysis using recovery biomarkers makes it thus possible to assess the accuracy and reliability of dietary assessment tools. Some studies have validated dietary assessment tools for energy intake with the doubly labeled water technique (DLW), which is considered as the gold standard measurement of total energy expenditure (Trabulsi and Schoeller, 2001, Westerterp, 2017, Thorsen et al., 2011, Timon et al., 2016, Park et al., 2018). Validating dietary assessment tools for some nutrients is also important since dietary misreporting does not only concern energy consumption, it can also affect some particular nutrients in the diet (Krebs-Smith et al., 2000, Bel-Serrat et al., 2016). Sodium (McClean et al., 2017, McClean et al., 2019, Park et al., 2018), potassium (Tasevska et al., 2006, Park et al., 2018) and urea (Stamler et al.,

1996, Tielemans et al., 2013, Park et al., 2018) in 24-hour urine collection can be used as recovery biomarkers for sodium, potassium and protein intake, respectively. The measurement of these biomarkers is considered as the reference method for assessing the intake of these nutrients and reflect their short-term consumption.

Even when the development of a dietary assessment tool is following state-of-the-art standards (Subar et al., 2007), there is always a certain level of inaccuracy that could be explained by different factors related to the characteristics of the tool *per se* including errors found in the database used (Leclercq et al., 2001), issues related to portion sizes (Boushey et al., 2017) and recipes. Some factors related to the individual who is completing the dietary assessment have also been proposed as sources of error. Accordingly, the underestimation of food intake with self-report tools is common and has been related to being a woman as well as to having a higher BMI, cognitive dietary restraint and social desirability (Maurer et al., 2006, Subar et al., 2003). Other factors such as the fear of negative evaluation, body dissatisfaction, and loss of control when eating (i.e., disinhibition) have also been identified as potential factors explaining the misreporting of intakes (Maurer et al., 2006). The data remains, however, insufficient to fully understand their impact since only a limited number of studies that have validated dietary assessment tools against objective recovery biomarkers have also investigated individual characteristics that influence misreporting.

The R24W developed by our research team has already been validated with different strategies such as using serum carotenoids to validate the reported intake of fruits and vegetables (Lafrenière et al., 2019), using fully controlled feeding studies (Lafrenière et al., 2017) and assessing the relative validity as compared with a 3-day food record

(Lafreniere et al., 2018). However, the criterion validity of the R24W using recovery biomarkers has not yet been assessed. Therefore, the primary objective of the study was to validate the self-reported intakes of sodium, potassium and proteins with the R24W against urinary recovery biomarkers. As a secondary objective, we also wanted to identify individual characteristics that influence misreporting.

Methods

Study population and design

A sample of 130 subjects aged 20 to 65 years living in Quebec City area participated in the study. Participants were recruited by mailing lists of the Institute of Nutrition and Functional Foods and the institutional listserv of Laval University's staff and students. Excluded participants were those who had diabetes, kidney diseases or any condition affecting the urine production or excretion (e.g., dysuria, oliguria, urinary incontinence). Additional reasons for exclusion were not having access to a computer and to the Internet at home, being pregnant or breastfeeding, not having maintained a stable weight for the last three months, not being fluent in French and taking medications that may affect kidney or urinary function. It is important to underline that the design of the study was planned to include the validation of the R24W for energy intake using the DLW method. However, results from this objective cannot be presented in this paper as doubly labeled water related measurements have been postponed indefinitely due to various reasons related to the COVID-19 pandemic.

After the initial telephone contact, participants attended a screening visit in order to confirm their eligibility. Anthropometric variables were then measured and a 12-hour fasting blood sample was collected. Following this screening visit, participants attended three visits within a 10-day period. At the first visit (D-0: beginning of the study), anthropometric measurements (weight, height and waist circumference) were taken, and participants received instructions for collecting the first 24-hour urine sample. Visit 2 was scheduled three days after visit 1 for the participants to bring back the first 24-hour urine collection. Participants also received instructions for the collection of a second 24-hour urine sample. Visit 3 occurred exactly 10 days after visit 1. Anthropometric measurements were again performed, and participants brought back the second 24-hour urine collection. During the 10 days, each participant was asked to complete a total of three 24-hour food recalls on unannounced days (2-week days and 1-weekend day). Participants received a total of \$150 for their participation in the study. This study received approval from the Laval University Ethics Committee (2018-217) and participants provided written informed consent prior to taking part in the study.

Anthropometric measurements

Anthropometric measures were taken at the screening visit, at the beginning (Day-0) and the end (Day-10) of the study protocol. Height, weight and waist circumference were measured. For the purpose of the present analyses, we used measurements taken at the beginning of the study (Day-0).

Web-based 24-hour food recalls

Over a 10-day period, participants were asked to complete three unannounced 24-hour food recalls (including a weekend day) using the R24W that was developed to measure nutritional intakes in the French-speaking population of the province of Québec, Canada. The development of the R24W has been previously described in detail (Jacques et al., 2016). Briefly, R24W uses a sequence of questions inspired by the USDA's Automated Multiple Pass Method (AMPM) (Moshfegh et al., 2008) but as opposed to the AMPM approach, the R24W is using a meal-based approach in the first step. Respondents can enter as many meals or snacks as they need to reflect their eating pattern during the previous day. When choosing food items consumed, pictures representing up to eight different portion sizes are shown and cues are provided to help participants remember frequently added items to food/dish (toppings, sauces, spices, etc.) and commonly forgotten foods. A mandatory walk-through tutorial must be watched before completing the first 24-hour food recall using the R24W. Other optional tutorials are also available if needed.

Information about meal or snack context is also gathered: time, location, whether the meal was eaten alone or with others and screen-based activities during meals. Some questions about the use of vitamin and mineral supplements are also asked. Participants have to indicate if they consumed some supplements during the recorded day. If so, they then have to specifically indicate the type of supplement consumed. It is therefore possible to characterize participants according to their vitamin and mineral supplement intake, but we do not have the precise amount of vitamins or minerals consumed through these supplements and these amounts are not included in the evaluation of intakes.

Finally, the database includes 2865 food items that are linked to the Canadian Nutrient File released in 2015 and 687 recipes for mixed dishes (Jacques et al., 2016).

24-hour urine collections

Participants were asked to collect two 24-hour urine samples, one week apart, to measure the concentrations of urea, sodium and potassium respectively. A 24-hour urine collection kit that included containers in which 2.5 g of thymol (Wu et al., 2015) was used as a preservative, a manifold, instructions, three capsules of para-aminobenzoic acid (PABA) and icepacks was provided to participants. On urine collection days, which occurred on days 2 and 9, participants discarded the first void of the morning and began collecting urine for the next 24 hours, including the first void of the next morning. They were asked to keep the urine collection cold in the refrigerator. They were also asked to take an 80 mg PABA tablet, a marker for completeness of 24-hour urine collection (Cox et al., 2018, Bingham and Cummings, 1983), at breakfast, lunch and dinner. Participants had to provide the time of the first discarded void in the morning, the time of the last included void in the urine collection the next morning and the time at which the PABA tablets were taken during the day of the collection.

The 24-hour urine collections were brought to the Clinical Investigation Unit (CIU) of the Institute of Nutrition and Functional Foods (INAF) the day after each collection (i.e., on day 3 (visit 2) and day 10 (visit 3)). The total volume of urine was measured, separated into aliquots in 1.5 mL (urea, sodium, potassium) and stored at -80°C until measurement. Urinary urea was measured by an enzymatic method (Dimension Vista System, ref. K1021, BUN, Siemens Healthcare Diagnostics Inc.) and urinary potassium and sodium were measured by an integrated multidetector methodology using flame photometry,

spectrophotometry and potentiometry of ion-selective electrodes (Dimension Vista System, V-LYTE Integrated Multisensor, ref. K800A, Siemens Healthcare Diagnostics Inc.). The dietary protein intake was calculated given that 46.67% of urea is urea nitrogen, urea nitrogen constitutes 85% of nitrogen, 81% of nitrogen intake is excreted in urine and nitrogen constitutes 16% of protein (Bingham and Cummings, 1985a). Allowing that 80% of consumed potassium, and 86% of consumed sodium are excreted in urine (Freedman et al., 2004, Holbrook et al., 1984), we converted the urinary values into dietary intakes for sodium and potassium by applying the respective conversion factors. PABA was measured at INAF facilities by HPLC-UV and the percentage of PABA excretion was calculated with the total volume assuming that 100% excretion was equal to 240 mg. The urine collections with less than 85% of predicted concentration of PABA were considered incomplete, i.e., meaning that participants did not provide all their urine during the 24-hour urine collection (Bingham and Cummings, 1983, Prentice et al., 2011, Neuhouser et al., 2008, Bingham and Cummings, 1985b) and were excluded from the analyses. PABA concentrations also had to be below 110% of the predicted value for urine collections to be considered as valid for being included in the analyses (Knuiman et al., 1986). Therefore, if a participant had one out of two urine collections found to be invalid based on PABA concentrations (i.e., PABA lower than 85% or higher than 110% of the predicted value), analyses were then conducted using only measurements derived from the valid urine collection. If the two urine collections of a participant were considered as invalid, this participant was excluded from the analyses.

Questionnaires

Once the food recalls were completed and urine collected, participants had to complete online questionnaires that included a medical questionnaire, a socio-economic and demographic questionnaire, a questionnaire on eating habits with a question about adding salt at the table (never, rarely, often, always), the 40-item Balanced Inventory of Desirable Responding scale (Paulhus, 1984), the Brief Fear of Negative Evaluation Scale (Leary, 1983, Lacroix and Zhou, 2015), the subsection “Appearance” of the Body Esteem Scale (Mendelson et al., 2001) and the Three-Factor Eating Questionnaire (Stunkard and Messick, 1985).

The Balanced Inventory of Desirable Responding questionnaire was used to measure social desirability. This questionnaire includes two subscales: the first one is the self-deceptive enhancement subscale which is the tendency to provide self-reports that are honest but positively biased and the second one is the impression management subscale which is a tendency to present a favorable image of oneself to others. Each question of the Balanced Inventory of Desirable Responding questionnaire was answered using a seven-point Likert scale (1= strongly disagree, 7=strongly agree). After reversing the negatively keyed items, each rated question with an answer of 6 or 7 was assigned one point. The sum of each score corresponds to the social desirability score for each participant for a maximum of 40 points (high social desirability) and a maximum of 20 points for each subscale.

The Brief Fear of Negative Evaluation Scale questionnaire (12 items) was used to determine the degree of concerns with being evaluated unfavorably by others (fear of negative evaluation) (Leary, 1983). Each question was answered using a 5-point Likert scale (1= not at all characteristic of me, 5=extremely characteristic of me), after reversing

negatively scored items, the sum of each score for each question was calculated for a maximum of 60 points (greater fear of negative evaluation) and a minimum of 12 (Rodebaugh et al., 2004).

The appearance subscale from the Body Esteem Scale was used to rate body esteem related to appearance with 10 items using a 5-point Likert scale (1=never, 5=always). The sum of each item score was calculated after reversing negative items. The total score ranges from 0 to 40 and a higher score means a higher body esteem related to appearance.

The Three-Factor Eating Questionnaire (51 items) was used to measure three different dimensions of eating behavior: cognitive dietary restraint (20 items), disinhibition (16 items) and hunger (15 items) using “true” or “false” for the first 36 items, a 4-point Likert scale for 14 items (e.g., 1=never/not at all, 4=always/very much), and a 6-point Likert scale for one item related to cognitive dietary restraint. After reversing negative items, each answer “true” and each answer greater than 3 for all other items were given one point (Stunkard and Messick, 1985).

Statistical analyses

Student’s paired t-test was used to determine if there was a significant difference between sodium, potassium and protein mean daily intake calculated from the 24-hour food recalls and the recovery biomarker values. Raw and deattenuated Pearson correlations were calculated to evaluate the strength of the associations between the self-reported intakes from the R24W and the recovery biomarker values for each nutrient. The correction for attenuation was used to reduce the measurement error due to within-person variability observed between the measurement days. More precisely, in order to obtain the

deattenuated correlation coefficient, one has to divide the raw correlation by the square root of the product between the reliability of the self-reported value and the reliability of the biomarker-derived value (Trafimow, 2015, Rosner and Willett, 1988). Cross-classification analyses were performed in order to examine the ability of the R24W to classify participants into concordant quartiles of recovery biomarkers. The weighted Cohen's kappa (k_w) was then calculated to assess the agreement of the cross-classification analyses (Gleason et al., 2010). Bland-Altman plots were used to assess agreement between methods (Giavarina, 2015, Bland and Altman, 1986).

Further analyses were performed to determine whether the difference between the biomarkers and the self-reported intakes could be explained by some individual characteristics of our participants. To facilitate the description of these analyses, this difference is referred to as misreporting, assuming that the biomarker value is the true intake value. Misreporting was calculated by subtracting the intake as determined by the biomarker from the reported intake as evaluated by the R24W. Correlation analyses were then performed to evaluate the association between misreporting values and the scores reflecting particular characteristics of participants potentially associated with misreporting such as social desirability, fear of negative evaluation, body esteem related to appearance, and eating behaviors. For adding salt at the table, we compared by Student's t-test the misreporting value for sodium between the two following groups i.e., often or always adding salt at the table vs. never or rarely using salt at the table. Since missing data were less than 10% of the subjects, simple imputation i.e., imputation by person mean, was used (Eekhout et al., 2014). If more than 10% of questions were missing for a given subject, the subject's score was excluded from the analyses. Therefore, imputation was used for

missing data (i.e., social desirability, $n=4$; fear of negative evaluation, $n=4$; body esteem related to appearance, $n=6$, variables of the Three-factor eating questionnaire, $n=8$). One participant was excluded from analyses related to cognitive dietary restraint and one participant was excluded from analyses related to disinhibition. If an interaction term was found between a subjects' characteristic and sex or BMI, subsequent correlation analyses were performed by sex or by BMI groups.

To evaluate and to provide an overview of the validity of the R24W, each statistical test was compared to the criteria proposed by Lombard et al. and interpreted as good, acceptable or poor outcomes (Lombard et al., 2015). Some tests were used to assess the validity at a group level (Student's t -test, percent difference and Bland Altman plot), while others reflected the validity at an individual level (correlation, cross classification (two criteria) and weighted Kappa). The validity was interpreted as being good when the following results were obtained: correlation coefficient ≥ 0.50 , non-significant paired Student's t -test ($p > 0.05$), classification of $\geq 50\%$ of respondents in the same quartile, classification of $\leq 10\%$ of respondents in opposite quartiles, weighted Kappa ≥ 0.61 and non-significant slope in the Bland-Altman plot ($p > 0.05$). Moreover, the validity was interpreted as being acceptable when the correlation coefficient was between 0.20 and 0.49, the difference between both methods was $\leq 10\%$ and the weighted Kappa was between 0.20 and 0.60. Finally, the validity was interpreted as being poor when the following results were obtained: correlation coefficient < 0.20 , significant paired Student's t -test ($p \leq 0.05$), difference between both methods $> 10\%$, classification of $< 50\%$ of respondents in the same quartile, classification of $> 10\%$ of respondents in opposite quartiles, weighted Kappa < 0.20 .

and a significant slope in the Bland-Altman plot ($p \leq 0.05$). Statistical analyses were conducted with the statistical software package SAS university studio (version 3.8).

Results

A total of 319 individuals who were interested in participating in the study contacted the research team by email or phone. About half of them ($n=167$) were excluded after the initial contact as they did not meet the inclusion criteria ($n=70$), were not interested after learning more about the project ($n=50$) or were unreachable ($n=47$). The remaining 152 individuals attended a screening visit to confirm their eligibility. Of them, 22 were excluded and therefore 130 participants were included in the study.

All participants provided the two 24-hour urine samples. Three participants had one of their urine samples that was not sent to the laboratory for the analysis of biomarkers because total urine volume was not appropriately measured. One participant reported having a technical problem with one urine collection and had therefore only one valid urine collection. A total of 23 participants had one invalid urine collection as one of their urine samples had less than 85% of the predicted concentration of PABA. Three participants had two invalid urine collections based on the PABA concentration criterion. Also, one of the participants did not take the PABA tablets for any of the urine collections, and urine samples from this participant were therefore not included in the analyses. As a result, 99 participants had two valid urine samples, 27 had only one valid sample (for them, analyses were performed only on the valid collection) and four had two invalid collections. These four participants were consequently excluded from the analyses. Therefore, data from 126

subjects were used for the analyses comparing reported intakes to biomarker concentrations.

The main characteristics of the participants are shown in Table 1. Almost half of the participants were men (46.9 %). Mean age of the participants was 40.5 years, mean BMI was 26.7 kg/m² and mean waist circumference was 92.5 cm for men and 88.7 cm for women. Most of the participants were Caucasian (90.8 %) and had a university degree (72.3 %).

Mean self-reported intakes for sodium, potassium and proteins from the R24W, and mean intakes as determined by their respective biomarker are presented in Table 2. Mean self-reported intake of sodium was 2.2 % lower than the value obtained with the recovery biomarker. For potassium, the mean self-reported intake was 2.1 % less than the mean value derived from the biomarker. Mean self-reported intake of proteins was 5.0% lower than values from the recovery biomarker ($p < 0.05$). Raw correlations between the self-reported intakes and intakes based on recovery biomarkers were all significant ($p < 0.01$) for sodium ($r = 0.41$), potassium ($r = 0.47$) and proteins ($r = 0.61$). Deattenuated correlations were slightly higher ($r = 0.48$ for sodium; $r = 0.56$ for potassium; and $r = 0.68$ for proteins).

Cross-classification analyses into quartiles of the distribution for intakes determined either with the self-reported data or data from recovery biomarkers are presented in Table 3. Regarding sodium intake, 78.6 % of participants were in the same or adjacent quartiles with the two methods. For potassium and protein intakes, it was found that 81.0 and 85.7 % were respectively in the same or adjacent quartiles. Misclassification (i.e., when a participant is classified in the 1st quartile with a method and in the 4th with the other method) occurred in 4.8 % of the participants for sodium, 3.2 % for potassium and 0.8 % for

proteins. Weighted kappa values were 0.30, 0.36 and 0.41 for sodium, potassium and proteins, respectively.

The Bland-Altman analyses are shown in Figures 1, 2 and 3. In these figures, the Y axis corresponds to the difference between the self-reported and the biomarker-derived intakes while the X axis corresponds to the average between the self-reported and the biomarker-derived intakes. As shown in Figure 1, the mean difference for sodium (i.e., bias) between the self-reported and the biomarker-derived intakes was -76 mg and is represented by the line just below the X axis which corresponds to a difference of zero. It can be seen in Figure 2 that the bias is -70 mg for potassium while Figure 3 shows a bias of - 5.2 g for proteins. Also, it was found for all three nutrients that no proportional bias ($p > 0.05$) was observed meaning that the difference between both methods was not significantly different through the range of intakes (Figures 1, 2 and 3).

When participants ($n=11$) taking dietary supplements likely to contain small amounts of potassium and sodium (i.e., multivitamins with minerals) were excluded, similar results were obtained (results not shown).

Analyses were performed to evaluate the associations between some selected individual characteristics and the misreporting in sodium, potassium and proteins. Correlation analyses showed that neither total social desirability scores nor its subscales were associated with misreporting in sodium, potassium and protein intakes ($-0.10 \leq r \leq 0.11$). Similarly, no significant associations were found with the fear of negative evaluation scores ($r \leq 0.16$) and with variables derived from the Three-factor eating questionnaire ($-0.16 \leq r \leq 0.06$). No significant associations were observed between body esteem related to

appearance and misreporting in potassium and proteins. For all these non-significant associations, no moderation effects of sex or BMI were observed.

The association between the score of body esteem related to appearance and misreporting value in sodium was significantly moderated by sex ($p=0.03$) and BMI ($p=0.046$). Further analyses were therefore performed within each sex and BMI group separately for sodium misreporting (Table 4). In women only, a positive and significant association was observed between body esteem related to appearance and the misreporting in sodium ($r: 0.33$, $p=0.006$). This means that a lower body esteem related to appearance was associated with a more important under-reporting of sodium. No significant association between body esteem related to appearance and the misreporting in sodium was found within any of the BMI groups, although a trend for significance was observed for BMI ≥ 30.0 ($r: 0.36$, $p=0.06$). Further analyses conducted in the whole sample showed that the misreporting for sodium was not different ($p=0.12$) between participants using salt at the table, often or always (underreporting of 504 ± 1804 mg) vs. those using salt at the table never or rarely (overreporting of 58 ± 1303 mg). Finally, misreporting values were not significantly different between men and women (sodium: $p=0.80$, potassium: $p=0.28$, proteins: $p=0.16$) and among BMI groups (sodium: $p=0.08$, potassium: $p=0.87$, proteins: $p=0.67$).

The number of validation criteria considered as either good, acceptable or poor outcomes are shown for each nutrient in Table 5. Out of the seven validation criteria that were analyzed, it was found that sodium had three good outcomes, three acceptable outcomes and one poor outcome. For potassium, four good outcomes, two acceptable outcomes and one poor outcome were reported. Finally, for proteins, three good outcomes, two acceptable outcomes and two poor outcomes were observed.

Discussion

The objective of this study was to assess the criterion validity of an automated web-based 24-hour dietary recall developed by our research group, the R24W, using recovery biomarkers. Our results suggest that the R24W provides a valid assessment of sodium, potassium and protein intakes since the majority of validity criteria that we evaluated corresponded to either good or acceptable outcomes.

Our results showed that the criteria used to assess validity at a group level (percent difference, Student's-t test, Bland Altman slope) were generally indicative of acceptable or good agreement between the reported intakes and the values derived from the recovery biomarkers (Lombard et al., 2015). The percent difference between the two methods was -2.2 %, -2.1 % and -5.0 % for sodium, potassium and proteins respectively. The magnitude of these differences corresponds to acceptable outcomes according to Lombard et al. (Lombard et al., 2015). The only validity criterion at a group level that corresponded to a poor outcome was the statistically significant difference between methods for proteins. We therefore have to keep in mind that the R24W may be likely to underestimate the protein intake in particular. However, as discussed in detail below, the difference in proteins between methods that we observed in this study is within the range of values obtained in other studies (Wark et al., 2018, Greenwood et al., 2019, Park et al., 2018).

In general, our results showed smaller underestimation values than those documented in a study using another automated self-administered 24-hour recall, the ASA24, and using the same recovery parameters for biomarkers assessment as the ones we used in the present study. In fact, Park et al. observed that using the ASA24 leads to an underestimation reaching -13% for sodium, -16% for potassium and -10% for proteins in men and -9%, -6

% and -5% in women when comparing self-reported intakes provided by three 24-hour recalls to biomarkers (Park et al., 2018). Of note, some characteristics of our population differ from the ones in the study of Park et al. In fact, our sample was younger and less diverse than participants included in their study. These differences could have influenced the results obtained. Some web-based food recalls rather documented an overestimation of sodium, potassium and protein intakes. More specifically, Wark et al. found an overestimation of 22% for sodium, 26% for potassium and 1% for protein with myfood24 (Wark et al., 2018). Greenwood et al. also observed an overestimation associated with the Oxford WebQ as the mean difference from the biomarker-derived value reached 53% for potassium and 12% for proteins (Greenwood et al., 2019).

Moreover, we observed a mean difference in sodium intake of 76 mg between our two measurement methods, which is less than the pooled weighted mean difference of 602 mg that was observed in a meta-analysis of 28 studies comparing self-reported data to recovery biomarker-derived values (McClean et al., 2019). Furthermore, when only studies classified as having an excellent quality using a scoring system including five elements (sample size, statistics used, data collection procedure, consideration of seasonality and supplement use) (Serra-Majem et al., 2009), the mean difference (59 mg) was similar to the one observed in the present study. To our knowledge, no such meta-analysis exists for the comparison between self-reported and biomarker-derived values for proteins and potassium. This would be helpful to further assess the validity of dietary instruments for the evaluation of protein and potassium intakes.

Results obtained with the Bland-Altman analyses were indicative of good validity outcomes for sodium, potassium and proteins (Lombard et al., 2015) as no proportional

bias was observed. In contrast, Cornejo et al. observed a negative slope for sodium with the Bland-Altman plot, meaning that high consumers of sodium had a larger difference between measurements (i.e., food recalls vs. urinary biomarker) than lower consumers (Cornejo et al., 2014). The fact that the latter study used an interviewer-administered format for 24-hour recall could explain why results obtained were different from ours.

Some analyses were performed to assess the validity of the R24W at the individual level. The data showed that nine out of 12 results obtained corresponded to either acceptable or good outcomes. For deattenuated correlation coefficients, it was found that sodium had an acceptable outcome whereas potassium and proteins had good outcomes (Lombard et al., 2015). The stronger correlation that we observed for proteins than for sodium and potassium was also reported when intakes evaluated with 24-hour dietary recalls using myfood24 were correlated with intakes measured with biomarkers (Wark et al., 2018). Similarly, using the Oxford WebQ, the correlation for proteins was stronger than the one for potassium (Greenwood et al., 2019). The deattenuated correlation coefficient for sodium was the only one below the threshold for a good outcome which has also been observed in other studies (Wark et al., 2018, Campino et al., 2016, Dennis et al., 2003, Freedman et al., 2015, Lassale et al., 2015). This finding may be explained by the important variability in sodium content within different recipes and food products as suggested by Ahuja et al. who observed a vast range of sodium content among similar food types and brands (Ahuja et al., 2015). Also, sodium content in food products may change quickly meaning that values included in food databases are often outdated. Accordingly, the “other foods” group (i.e., food not recommended by the Canada’s Food Guide) has been shown to be the main source of sodium in the diet of Quebecers (Plante et al., 2019). This food

group includes mostly processed and highly processed food that often come in different brands. Also, these foods are likely to experience rapid changes in their sodium composition as they can be the target of national sodium reduction strategies. Therefore, the fact that this type of foods constitutes the main food sources of sodium can particularly enhance the risk of measurement errors that are described above. In addition, according to McLean et al., all dietary assessment methods are characterized by the difficulty to estimate discretionary salt, which is the salt added at home during food preparation or at the table, especially when the proportion of discretionary salt is high in the respondents diet (McLean et al., 2018). In our study, it was apparently not a significant factor since the misreporting in sodium was not different between participants who “always” or “often” added salt at the table compared to those who mentioned adding salt either “never” or “rarely”.

Although biomarkers measurement was used as the reference method in the present study, we have to acknowledge that there is always some degree of measurement error (e.g., assay imprecision) and biological variability associated with biomarkers. For example, studies have shown day-to-day variability in sodium excretion even when sodium intakes were clamped under highly controlled conditions (Rakova et al., 2013, Lucko et al., 2018). Therefore, it is possible that a part of the difference observed between the values from the R24W and the biomarker-derived values could be explained by measurement error and biological variability of biomarkers.

As previously documented, some individual factors can contribute to misreporting. In this study, the only significant association found was between body esteem related to appearance and sodium misreporting in women. The fact that this association was found only in women could be explained by some gender differences in the factors leading to low

body esteem related to appearance (Porrás-García et al., 2019, Chang and Christakis, 2003, Camirand et al., 2016). Further studies will be needed to better understand how a low body esteem related to appearance could be specifically associated with the underreporting of sodium in women.

Social desirability has been associated with misreporting in many studies, as documented by the review of Maurer et al. (Maurer et al., 2006). However, we did not observe such an association. The fact that the completion of 24-hour recalls was performed in an automated way (i.e., without an interviewer) could explain, at least partially, why social desirability was not associated with misreporting in our study. According to Tooze et al., the presence of an interviewer can have an impact on subjects, and they could then report dietary intake in a more “socially desirable” way, consciously or not (Tooze et al., 2004). Another reason that could explain our result is the fact that participants knew that some biomarkers of their food intake were being measured, which could have led to more honest self-report of their food intake.

The major strength of our study is the use of objective reference values to validate the R24W for sodium, potassium and proteins. The fact that we used a marker of completeness of urine collections (i.e., PABA) is another strength of this study allowing more accurate analyses. Accordingly, *McLean et al.* reported that studies that used a marker for completeness had a pooled mean difference of 488 ± 238 mg for sodium compared with 985 ± 515 mg for those that did not have one (McLean et al., 2019). Another strength of this study is that our sample has characteristics that are representative of our population. In fact, the percentage of people living with obesity in our study (22.3%) is very similar to the prevalence observed in the province of Québec (23.1%) (Statistics Canada, 2015).

Similarly, the prevalence of participants classified as being overweight is very similar in our sample and in the population of Québec (37.6% vs. 36.5%) (Statistics Canada, 2015). Also, the proportion of Caucasian participants in our study (90.8%) is only slightly higher than what is found at the population level in Quebec (88.2%) (Statistics Canada., 2016a). In terms of age, our adult sample aged between 18 and 65 y shows a pattern of distribution between the age groups (i.e., 18-34 y; 35-49y; 50-65y) that is similar to data collected at a population level. In fact, in our sample the proportion of participants in these three age groups are respectively 37.7%, 33.1% and 29.2% while the corresponding numbers at the population level are 32.6%, 30.4% and 37.0% (Statistics Canada., 2016b). Finally, the analyses including psychosocial factors as potential correlates of misreporting are novel and can be considered as a strength of this study.

The fact that our sample has generally a high level of education is a limitation of this study as it is not fully representative of our population. Also, the context of the study, namely the fact that subjects knew that we were measuring their intakes through biomarkers, might have prompted the participants to respond more conscientiously than they would usually do. This can be considered as a limitation as it might have positively influenced the validation results obtained. Finally, the fact that we were not able to conduct validation analyses for energy intake is a limitation of our analysis.

In conclusion, this study suggests that our web-based 24-hour recall application, the R24W, is valid to assess sodium, potassium and protein intakes in a French-speaking adults from the Province of Québec. The upcoming validation for energy with DLW will be helpful to better appreciate the overall validity of the R24W.

Acknowledgements

The authors would like to thank Steeve Larouche for his laboratory technical support, Élise Carbonneau for her help with statistical analyses and all participants for their participation. We would also like to acknowledge the contribution of undergraduate students Andrée-Anne Doyon, Frédérique Audet, Émilie Richard-Bilodeau and Élisabeth Demers-Potvin to the data collection, laboratory measurements and coding.

Competing interests statement

None

Contributor's statement

FP: Investigation, formal analysis, writing-original draft, visualization. **BL, JR, CC:** Conceptualization, methodology, writing-review and editing, funding acquisition. **JL:** Methodology, writing-review and editing. **AJT:** Investigation, writing-review and editing. **LC :** Conceptualization, methodology, writing-review and editing, project administration. **SL:** Conceptualization, methodology, writing-review and editing, supervision, project administration, funding acquisition.

All authors reviewed and approved the final version of the manuscript.

Funding statement

This work was funded by the Canadian Institutes of Health Research (Grant no. FHG 129921). Funders of the studies had no role in the design, conduct and management of the study, in data collection and analysis, or in the writing and reviewing of the present article. Frédérique Paradis received a studentship from the Centre NUTRISS.

Data availability statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Draft

References

Ahuja, J. K., Wasswa-Kintu, S., Haytowitz, D. B., Daniel, M., Thomas, R., Showell, B., Nickle, M., Roseland, J. M., Gunn, J., Cogswell, M. & Pehrsson, P. R. 2015. Sodium content of popular commercially processed and restaurant foods in the United States. *Prev Med Rep*, 2, 962-7. doi: 10.1016/j.pmedr.2015.11.003.PMID: 26844175

Bel-Serrat, S., Julián-Almárcegui, C., González-Gross, M., Mouratidou, T., Börnhorst, C., Grammatikaki, E., Kersting, M., Cuenca-García, M., Gottrand, F., Molnár, D., Hallström, L., Dallongeville, J., Plada, M., Roccaldo, R., Widhalm, K., Moreno, L. A., Manios, Y., De Henauw, S., Leclercq, C., Vandevijvere, S., Lioret, S., Gutin, B. & Huybrechts, I. 2016. Correlates of dietary energy misreporting among European adolescents: the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) study. *Br J Nutr*, 115, 1439-52. doi: 10.1017/S0007114516000283.PMID: 26888046

Bingham, S. & Cummings, J. H. 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, 629-35. DOI: 10.1042/cs0640629. PMID 6601560

Bingham, S. A. & Cummings, J. H. 1985a. 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, 1276-89. doi: 10.1093/ajcn/42.6.1276.PMID: 4072961

Bingham, S. A. & Cummings, J. H. 1985b. The use of creatinine output as a check on the completeness of 24-hour urine collections. *Hum Nutr Clin Nutr*, 39, 343-53.

PMID: 3877037

Bland, J. M. & Altman, D. G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307-10. PMID: 2868172

Boushey, C. J., Spoden, M., Zhu, F. M., Delp, E. J. & Kerr, D. A. 2017. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. *Proc Nutr Soc*, 76, 283-294. DOI: 10.1017/S0029665116002913.

PMID: 27938425

Camirand, H., Traoré, I. & Bauline, J. 2016. L'Enquête québécoise sur la santé de la population, 2014-2015: pour en savoir plus sur la santé des Québécois. Résultats de la deuxième édition. Québec: Institut de la statistique du Québec.

Campino, C., Hill, C., Baudrand, R., Martínez-Aguayo, A., Aglony, M., Carrasco, C. A., Ferrada, C., Loureiro, C., Vecchiola, A., Bancalari, R., Grob, F., Carvajal, C. A., Lagos, C. F., Valdivia, C., Tapia-Castillo, A., Fuentes, C. A., Mendoza, C., Garcia, H., Uauy, R. & Fardella, C. E. 2016. Usefulness and Pitfalls in Sodium Intake Estimation: Comparison

of Dietary Assessment and Urinary Excretion in Chilean Children and Adults. *Am J Hypertens*, 29, 1212-7. DOI: 10.1093/ajh/hpw056. PMID: 27279009

Carter, M. C., Albar, S. A., Morris, M. A., Mulla, U. Z., Hancock, N., Evans, C. E., Alwan, N. A., Greenwood, D. C., Hardie, L. J., Frost, G. S., Wark, P. A., Cade, J. E. & Myfood, C. 2015. Development of a UK Online 24-h Dietary Assessment Tool: myfood24. *Nutrients*, 7, 4016-32. DOI: 10.3390/nu7064016. PMID: 26024292

Chang, V. W. & Christakis, N. A. 2003. Self-perception of weight appropriateness in the United States. *Am J Prev Med*, 24, 332-9. DOI: 10.1016/s0749-3797(03)00020-5. PMID: 12726871

Cornejo, K., Pizarro, F., Atalah, E. & Galgani, J. E. 2014. [Assessment of dietary intake and urinary excretion of sodium and potassium in adults]. *Rev Med Chil*, 142, 687-95. DOI: 10.4067/S0034-98872014000600001. PMID: 25327312

Cox, L., Guberg, K., Young, S., Nicholson, S., Steer, T., Prentice, A. & Page, P. 2018. Validation of the use of p-aminobenzoic acid to determine completeness of 24 h urine collections in surveys of diet and nutrition. *Eur J Clin Nutr*, 72, 1180-1182. DOI: 10.1038/s41430-018-0195-x. PMID: 29872160

Dennis, B., Stamler, J., Buzzard, M., Conway, R., Elliott, P., Moag-Stahlberg, A., Okayama, A., Okuda, N., Robertson, C., Robinson, F., Schakel, S., Stevens, M., Van Heel, N., Zhao, L., Zhou, B. F. & Group, I. R. 2003. INTERMAP: the dietary data--process and quality control. *J Hum Hypertens*, 17, 609-22. DOI: 10.1038/sj.jhh.1001604. PMID: 13679951

Eekhout, I., De Vet, H. C., Twisk, J. W., Brand, J. P., De Boer, M. R. & Heymans, M. W. 2014. Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *J Clin Epidemiol*, 67, 335-42. DOI:10.1016/j.jclinepi.2013.09.009. PMID: 24291505

Freedman, L. S., Commins, J. M., Moler, J. E., Arab, L., Baer, D. J., Kipnis, V., Midthune, D., Moshfegh, A. J., Neuhouser, M. L., Prentice, R. L., Schatzkin, A., Spiegelman, D., Subar, A. F., Tinker, L. F. & Willett, W. 2014. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. *Am J Epidemiol*, 180, 172-88. DOI: 10.1093/aje/kwu116. PMID: 24918187

Freedman, L. S., Commins, J. M., Moler, J. E., Willett, W., Tinker, L. F., Subar, A. F., Spiegelman, D., Rhodes, D., Potischman, N., Neuhouser, M. L., Moshfegh, A. J., Kipnis, V., Arab, L. & Prentice, R. L. 2015. Pooled results from 5 validation studies of dietary

self-report instruments using recovery biomarkers for potassium and sodium intake. *Am J Epidemiol*, 181, 473-87. DOI: 10.1093/aje/kwu325. PMID: 25787264

Freedman, L. S., Midthune, D., Carroll, R. J., Krebs-Smith, S., Subar, A. F., Troiano, R. P., Dodd, K., Schatzkin, A., Bingham, S. A., Ferrari, P. & Kipnis, V. 2004. Adjustments to improve the estimation of usual dietary intake distributions in the population. *J Nutr*, 134, 1836-43. DOI: 10.1093/jn/134.7.1836. PMID: 15226478

Giavarina, D. 2015. Understanding Bland Altman analysis. *Biochem Med (Zagreb)*, 25, 141-51. DOI: 10.11613/BM.2015.015. PMID: 26110027

Gleason, P. M., Harris, J., Sheean, P. M., Boushey, C. J. & Bruemmer, B. 2010. Publishing nutrition research: validity, reliability, and diagnostic test assessment in nutrition-related research. *J Am Diet Assoc*, 110, 409-19. DOI: 10.1016/j.jada.2009.11.022. PMID: 20184991

Greenwood, D. C., Hardie, L. J., Frost, G. S., Alwan, N. A., Bradbury, K. E., Carter, M., Elliott, P., Evans, C. E. L., Ford, H. E., Hancock, N., Key, T. J., Liu, B., Morris, M. A., Mulla, U. Z., Petropoulou, K., Potter, G. D. M., Riboli, E., Young, H., Wark, P. A. & Cade, J. E. 2019. Validation of the Oxford WebQ Online 24-Hour Dietary Questionnaire Using Biomarkers. *Am J Epidemiol*, 188, 1858-1867. DOI: 10.1093/aje/kwz165.

PMID: 31318012

Holbrook, J. T., Patterson, K. Y., Bodner, J. E., Douglas, L. W., Veillon, C., Kelsay, J. L., Mertz, W. & Smith, J. C., Jr. 1984. Sodium and potassium intake and balance in adults consuming self-selected diets. *Am J Clin Nutr*, 40, 786-93. DOI: 10.1093/ajcn/40.4.786.

PMID: 6486085

Illner, A. K., Freisling, H., Boeing, H., Huybrechts, I., Crispim, S. P. & Slimani, N. 2012. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. *Int J Epidemiol*, 41, 1187-203. DOI: 10.1093/ije/dys105.

PMID: 22933652

Jacques, S., Lemieux, S., Lamarche, B., Laramée, C., Corneau, L., Lapointe, A., Tessier-Grenier, M. & Robitaille, J. 2016. Development of a Web-Based 24-h Dietary Recall for a French-Canadian Population. *Nutrients*, 8. DOI: 10.3390/nu8110724. PMID: 27854276

Kant, A. K. 2010. Dietary patterns: biomarkers and chronic disease risk. *Appl Physiol Nutr Metab*, 35, 199-206. DOI: 10.1139/H10-005. PMID: 20383233

Kirkpatrick, S. I., Potischman, N., Dodd, K. W., Douglass, D., Zimmerman, T. P., Kahle, L. L., Thompson, F. E., George, S. M. & Subar, A. F. 2016. The Use of Digital Images in 24-Hour Recalls May Lead to Less Misestimation of Portion Size Compared with Traditional Interviewer-Administered Recalls. *J Nutr*, 146, 2567-2573. DOI: 10.3945/jn.116.237271. PMID: 27807039

Kirkpatrick, S. I., Subar, A. F., Douglass, D., Zimmerman, T. P., Thompson, F. E., Kahle, L. L., George, S. M., Dodd, K. W. & Potischman, N. 2014. Performance of the Automated Self-Administered 24-hour Recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. *Am J Clin Nutr*, 100, 233-40. DOI: 10.3945/ajcn.114.083238. PMID: 24787491

Knuiman, J. T., Hautvast, J. G., Van Der Heyden, L., Geboers, J., Joossens, J. V., Tornqvist, H., Isaksson, B., Pietinen, P., Tuomilehto, J. & Poulsen, L. 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, 229-37. PMID: 3487532

Krebs-Smith, S. M., Graubard, B. I., Kahle, L. L., Subar, A. F., Cleveland, L. E. & Ballard-Barbash, R. 2000. Low energy reporters vs others: a comparison of reported food intakes. *Eur J Clin Nutr*, 54, 281-7. DOI: 10.1038/sj.ejcn.1600936. PMID: 10745278

Lacroix, F. & Zhou, B. 2015. Scales translated into French: social anxiety and Taijin Kyofusho. Concordia University, Montreal University.

Lafrenière, J., Couillard, C., Lamarche, B., Laramée, C., Vohl, M. C. & Lemieux, S. 2019. Associations between self-reported vegetable and fruit intake assessed with a new web-based 24-h dietary recall and serum carotenoids in free-living adults: a relative validation study. *J Nutr Sci*, 8, e26. DOI: 10.1017/jns.2019.23. PMID: 31428333

Lafrenière, J., Lamarche, B., Laramée, C., Robitaille, J. & Lemieux, S. 2017. Validation of a newly automated web-based 24-hour dietary recall using fully controlled feeding studies. *BMC Nutr*, 3, 34. : DOI: 10.1186/s40795-017-0153-3. PMID: 32153814

Lafrenière, J., Laramée, C., Robitaille, J., Lamarche, B. & Lemieux, S. 2018. Assessing the relative validity of a new, web-based, self-administered 24 h dietary recall in a French-Canadian population. *Public Health Nutr*, 21, 2744-2752. DOI: 10.1017/S1368980018001611. PMID: 29976261

Lassale, C., Castetbon, K., Laporte, F., Camilleri, G. M., Deschamps, V., Vernay, M., Faure, P., Hercberg, S., Galan, P. & Kesse-Guyot, E. 2015. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br J Nutr*, 113, 953-62. DOI: 10.1017/S0007114515000057. PMID: 25772032

Leary, M. R. 1983. A Brief Version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, 9, 371-75. DOI:10.1177/0146167283093007

Leclercq, C., Valsta, L. M. & Turrini, A. 2001. Food composition issues--implications for the development of food-based dietary guidelines. *Public Health Nutr*, 4, 677-82. DOI: 10.1079/phn2001153. PMID: 11683561

Liu, B., Young, H., Crowe, F. L., Benson, V. S., Spencer, E. A., Key, T. J., Appleby, P. N. & Beral, V. 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. DOI: 10.1017/S1368980011000942. PMID: 21729481

Lombard, M. J., Steyn, N. P., Charlton, K. E. & Senekal, M. 2015. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. *Nutr J*, 14, 40. DOI: 10.1186/s12937-015-0027-y. PMID: 25897837

Lucko, A. M., Doktorchik, C., Woodward, M., Cogswell, M., Neal, B., Rabi, D., Anderson, C., He, F. J., Macgregor, G. A., L'abbe, M., Arcand, J., Whelton, P. K., Mclean, R., Campbell, N. R. C. & Consortium, T. 2018. Percentage of ingested sodium

excreted in 24-hour urine collections: A systematic review and meta-analysis. *J Clin Hypertens (Greenwich)*, 20, 1220-1229. DOI: 10.1111/jch.13353. PMID: 30101426

Maurer, J., Taren, D. L., Teixeira, P. J., Thomson, C. A., Lohman, T. G., Going, S. B. & Houtkooper, L. B. 2006. The psychosocial and behavioral characteristics related to energy misreporting. *Nutr Rev*, 64, 53-66. DOI: 10.1301/nr.2006.feb.53-66. PMID: 16536182

McLean, R., Cameron, C., Butcher, E., Cook, N. R., Woodward, M. & Campbell, N. R. C. 2019. Comparison of 24-hour urine and 24-hour diet recall for estimating dietary sodium intake in populations: A systematic review and meta-analysis. *J Clin Hypertens (Greenwich)*, 21, 1753-1762. DOI: 10.1111/jch.13729. PMID: 31769168 .

McLean, R. M., Farmer, V. L., Nettleton, A., Cameron, C. M., Cook, N. R., Campbell, N., TRUE Consortium (International Consortium for Quality Research on Dietary Sodium/Salt) 2017. Assessment of dietary sodium intake using a food frequency questionnaire and 24-hour urinary sodium excretion: a systematic literature review. *J Clin Hypertens (Greenwich)*, 19, 1214-1230. DOI: 10.1111/jch.13148. PMID: 29220554

McLean, R. M., Farmer, V. L., Nettleton, A., Cameron, C. M., Cook, N. R., Woodward, M., Campbell, N. R. C. & Sodium/Salt), TRUE Consortium (International Consortium

for Quality Research on Dietary Sodium/Salt) 2018. Twenty-Four-Hour Diet recall and Diet records compared with 24-hour urinary excretion to predict an individual's sodium consumption: A Systematic Review. *J Clin Hypertens (Greenwich)*, 20, 1360-1376. DOI: 10.1111/jch.13391. PMID: 30298972

Mendelson, B. K., Mendelson, M. J. & White, D. R. 2001. Body-esteem scale for adolescents and adults. *J Pers Assess*, 76, 90-106. DOI: 10.1207/S15327752JPA7601_6. PMID: 11206302

Micha, R., Penalvo, J. L., Cudhea, F., Imamura, F., Rehm, C. D. & Mozaffarian, D. 2017. Association Between Dietary Factors and Mortality From Heart Disease, Stroke, and Type 2 Diabetes in the United States. *JAMA*, 317, 912-924. DOI: 10.1001/jama.2017.0947. PMID: 28267855

Moshfegh, A. J., Rhodes, D. G., Baer, D. J., Murayi, T., Clemens, J. C., Rumpler, W. V., Paul, D. R., Sebastian, R. S., Kuczyński, K. J., Ingwersen, L. A., Staples, R. C. & Cleveland, L. E. 2008. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr*, 88, 324-32. DOI: 10.1093/ajcn/88.2.324. PMID: 18689367

Neuhouser, M. L., Tinker, L., Shaw, P. A., Schoeller, D., Bingham, S. A., Horn, L. V., Beresford, S. A., Caan, B., Thomson, C., Satterfield, S., Kuller, L., Heiss, G., Smit, E., Sarto, G., Ockene, J., Stefanick, M. L., Assaf, A., Runswick, S. & Prentice, R. L. 2008. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol*, 167, 1247-59. DOI: 10.1093/aje/kwn026. PMID: 18344516

Pannucci, T. E., Thompson, F. E., Bailey, R. L., Dodd, K. W., Potischman, N., Kirkpatrick, S. I., Alexander, G. L., Coleman, L. A., Kushi, L. H., Groesbeck, M., Sundaram, M., Clancy, H., George, S. M., Kahle, L. & Subar, A. F. 2018. Comparing Reported Dietary Supplement Intakes between Two 24-Hour Recall Methods: The Automated Self-Administered 24-Hour Dietary Assessment Tool and the Interview-Administered Automated Multiple Pass Method. *J Acad Nutr Diet*, 118, 1080-1086. DOI: 10.1016/j.jand.2018.02.013. PMID: 29803270

Park, Y., Dodd, K. W., Kipnis, V., Thompson, F. E., Potischman, N., Schoeller, D. A., Baer, D. J., Midthune, D., Troiano, R. P., Bowles, H. & Subar, A. F. 2018. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. *Am J Clin Nutr*, 107, 80-93. DOI: 10.1093/ajcn/nqx002. PMID: 29381789

Paulhus, D. L. 1984. Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609. DOI: 10.1037/0022-3514.46.3.598

Plante, C., Rochette, M. & Blanchet, C. 2019. Les apports et les sources alimentaires de sucre, de sodium et de gras saturés chez les Québécois. Collection : Regard sur l'alimentation des Québécois. Numéro 2. Institut national de santé publique du Québec. Québec. 34 p.

Porras-Garcia, B., Ferrer-Garcia, M., Ghita, A., Moreno, M., López-Jiménez, L., Vallvé-Romeu, A., Serrano-Troncoso, E. & Gutiérrez-Maldonado, J. 2019. The influence of gender and body dissatisfaction on body-related attentional bias: An eye-tracking and virtual reality study. *Int J Eat Disord*, 52, 1181-1190. DOI: 10.1002/eat.23136.

PMID: 31301151

Prentice, R. L., Mossavar-Rahmani, Y., Huang, Y., Van Horn, L., Beresford, S. A., Caan, B., Tinker, L., Schoeller, D., Bingham, S., Eaton, C. B., Thomson, C., Johnson, K. C., Ockene, J., Sarto, G., Heiss, G. & Neuhouser, M. L. 2011. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. *Am J Epidemiol*, 174, 591-603. DOI: 10.1093/aje/kwr140. PMID: 21765003

Rakova, N., Juttner, K., Dahlmann, A., Schroder, A., Linz, P., Kopp, C., Rauh, M., Goller, U., Beck, L., Agureev, A., Vassilieva, G., Lenkova, L., Johannes, B., Wabel, P., Moissl, U., Vienken, J., Gerzer, R., Eckardt, K. U., Muller, D. N., Kirsch, K., Morukov, B., Luft, F. C. & Titze, J. 2013. Long-term space flight simulation reveals infradian rhythmicity in human Na(+) balance. *Cell Metab*, 17, 125-31. DOI: 10.1016/j.cmet.2012.11.013. PMID: 23312287

Rodebaugh, T. L., Woods, C. M., Thissen, D. M., Heimberg, R. G., Chambless, D. L. & Rapee, R. M. 2004. More information from fewer questions: the factor structure and item properties of the original and brief fear of negative evaluation scale. *Psychol Assess*, 16, 169-81. DOI: 10.1037/1040-3590.16.2.169. PMID: 15222813

Rosner, B. & Willett, W. C. 1988. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am J Epidemiol*, 127, 377-86. DOI: 10.1093/oxfordjournals.aje.a114811. PMID: 3337089

Serra-Majem, L., Frost Andersen, L., Henríque-Sánchez, P., Doreste-Alonso, J., Sánchez-Villegas, A., Ortiz-Andrelluchi, A., Negri, E. & La Vecchia, C. 2009. Evaluating the quality of dietary intake validation studies. *Br J Nutr*, 102 Suppl 1, S3-9. DOI: 10.1017/S0007114509993114. PMID: 20100366

U.S. Department of Agriculture and U.S. Department of Health and Human Services
2020. Dietary guidelines for Americans 2020-2025. 9th edition. available at
dietaryguidelines.gov.

Shim, J. S., Oh, K. & Kim, H. C. 2014. Dietary assessment methods in epidemiologic
studies. *Epidemiol Health*, 36, e2014009. DOI: 10.4178/epih/e2014009. PMID:
25078382

Stamler, J., Elliott, P., Kesteloot, H., Nichols, R., Claeys, G., Dyer, A. R. & Stamler, R.
1996. Inverse relation of dietary protein markers with blood pressure. Findings for 10,020
men and women in the INTERSALT Study. INTERSALT Cooperative Research Group.
INTERNational study of SALT and blood pressure. *Circulation*, 94, 1629-34. DOI:
10.1161/01.cir.94.7.1629. PMID: 8840854

Statistics Canada. 2015. Measured body mass index, Canadian Community Health
Survey-Nutrition 2015. Available in CANSIM: tables 105-2023 to 105-2026.
<https://www150.statcan.gc.ca/n1/en/type/data?text=105-2023..105-2026> [Online].
[Accessed September,10 2021].

Statistics Canada. 2016a. Immigration and Ethnocultural Diversity Highlight Tables, 2016 Census. <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hltfst/imm/index-eng.cfm> [Online]. [Accessed September 1, 2021].

Statistics Canada. 2016b. Age and Sex Highlight Tables, 2016 Census. <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hltfst/as/index-eng.cfm> [Online]. [Accessed September, 1 2021].

Stunkard, A. J. & Messick, S. 1985. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J Psychosom Res*, 29, 71-83. DOI: 10.1016/0022-3999(85)90010-8. PMID: 3981480

Subar, A. F., Kipnis, V., Troiano, R. P., Midthune, D., Schoeller, D. A., Bingham, S., Sharbaugh, C. O., Trabulsi, J., Runswick, S., Ballard-Barbash, R., Sunshine, J. & Schatzkin, A. 2003. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol*, 158, 1-13. DOI: 10.1093/aje/kwg092. PMID: 12835280

Subar, A. F., Thompson, F. E., Potischman, N., Forsyth, B. H., Buday, R., Richards, D., McNutt, S., Hull, S. G., Guenther, P. M., Schatzkin, A. & Baranowski, T. 2007. Formative research of a quick list for an automated self-administered 24-hour dietary

recall. J Am Diet Assoc, 107, 1002-7. DOI: 10.1016/j.jada.2007.03.007. PMID: 17524721

Tasevska, N., Runswick, S. A. & Bingham, S. A. 2006. Urinary potassium is as reliable as urinary nitrogen for use as a recovery biomarker in dietary studies of free living individuals. J Nutr, 136, 1334-40. DOI: 10.1093/jn/136.5.1334, PMID: 16614426

Thompson, F. E., Dixit-Joshi, S., Potischman, N., Dodd, K. W., Kirkpatrick, S. I., Kushi, L. H., Alexander, G. L., Coleman, L. A., Zimmerman, T. P., Sundaram, M. E., Clancy, H. A., Groesbeck, M., Douglass, D., George, S. M., Schap, T. E. & Subar, A. F. 2015. Comparison of Interviewer-Administered and Automated Self-Administered 24-Hour Dietary Recalls in 3 Diverse Integrated Health Systems. Am J Epidemiol, 181, 970-8. DOI: 10.1093/aje/kwu467. PMID: 25964261

Thorsen, T., Shriver, T., Racine, N., Richman, B. A. & Schoeller, D. A. 2011. Doubly labeled water analysis using cavity ring-down spectroscopy. Rapid Commun Mass Spectrom, 25, 3-8. DOI: 10.1002/rcm.4795. PMID: 21154648

Tielemans, S. M., Geleijnse, J. M., Van Baak, M. A., Engberink, M. F., Brink, E. J., De Jong, P. E., Gansevoort, R. T., Bakker, S. J. & Group, P. S. 2013. Twenty-four hour

urinary urea excretion and 9-year risk of hypertension: the PREVEND study. *J*

Hypertens, 31, 1564-9. DOI: 10.1097/HJH.0b013e328362213b. PMID: 23751964

Timon, C. M., Van Den Barg, R., Blain, R. J., Kehoe, L., Evans, K., Walton, J., Flynn, A. & Gibney, E. R. 2016. A review of the design and validation of web- and computer-based 24-h dietary recall tools. *Nutr Res Rev*, 29, 268-280. DOI: 10.1017/S0954422416000172. PMID: 27955721

Tooze, J. A., Subar, A. F., Thompson, F. E., Troiano, R., Schatzkin, A. & Kipnis, V. 2004. Psychosocial predictors of energy underreporting in a large doubly labeled water study. *Am J Clin Nutr*, 79, 795-804. DOI: 10.1093/ajcn/79.5.795. PMID: 15113717

Trabulsi, J. & Schoeller, D. A. 2001. Evaluation of dietary assessment instruments against doubly labeled water, a biomarker of habitual energy intake. *Am J Physiol Endocrinol Metab*, 281, E891-9. DOI: 10.1152/ajpendo.2001.281.5.E891. PMID: 11595643

Trafimow, D. 2015. The attenuation of correlation coefficients: a statistical literacy issue. *Teach Stat*, 38, 25-28. DOI: 10.1111/test.12087

Wark, P. A., Hardie, L. J., Frost, G. S., Alwan, N. A., Carter, M., Elliott, P., Ford, H. E., Hancock, N., Morris, M. A., Mulla, U. Z., Noorwali, E. A., Petropoulou, K., Murphy, D., Potter, G. D. M., Riboli, E., Greenwood, D. C. & Cade, J. E. 2018. Validity of an online 24-h recall tool (myfood24) for dietary assessment in population studies: comparison with biomarkers and standard interviews. *BMC Med*, 16, 136. DOI: 10.1186/s12916-018-1113-8. PMID: 30089491

Westerterp, K. R. 2017. Doubly labelled water assessment of energy expenditure: principle, practice, and promise. *Eur J Appl Physiol*, 117, 1277-1285. DOI: 10.1007/s00421-017-3641-x. PMID: 28508113

Wu, W., Yang, D., Tiselius, H. G., Ou, L., Mai, Z., Chen, K., Zhu, H., Xu, S., Zhao, Z. & Zeng, G. 2015. Collection and storage of urine specimens for measurement of urolithiasis risk factors. *Urology*, 85, 299-303. DOI: 10.1016/j.urology.2014.10.030. PMID: 25623670

Zimmerman, T. P., Hull, S. G., McNutt, S., Mittl, B., Islam, N., Guenther, P. M., Thompson, F. E., Potischman, N. A. & Subar, A. F. 2009. Challenges in converting an interviewer-administered food probe database to self-administration in the National Cancer Institute Automated Self-administered 24-Hour Recall (ASA24). *J Food Compos Anal*, 22, S48-S51. DOI: 10.1016/j.jfca.2009.02.003. PMID: 20161418

Figure captions

Figure 1: Bland-Altman plot showing the difference between the self-reported and the biomarker-derived value vs. the average of the two values for sodium intake (mg per day), Bias = -76 mg, $r=0.005$, $p=0.95$, limits of agreements = -2918 to 2766 mg.

Figure 2: Bland-Altman plot showing the difference between the self-reported and the biomarker-derived value vs. the average of the two values for potassium intake (mg per day), Bias = -70 mg, $r=-0.06$, $p=0.48$, limits of agreements = -2050 to 1909 mg.

Figure 3: Bland-Altman plots showing the difference between the self-reported and the biomarker-derived value vs. the average of the two values for protein intake (g per day), bias = -5.2 g, $r=0.17$, $p=0.06$, limits of agreements = -60.7 to 50.2 g.

Table 1. Subjects' characteristics

Variables	Total sample (n=130)		Men (n= 61)		Women (n=69)	
	Mean or n	SD or (%)	Mean or n	SD or (%)	Mean or n	SD or (%)
Age (years)	40.5	13.8	39.5	14.3	41.4	13.3
18-34	49	(37.7)	25	(41.0)	24	(34.8)
35-49	43	(33.1)	18	(29.5)	25	(36.2)
50-65	38	(29.2)	18	(29.5)	20	(29.1)
Body mass index (kg/m ²)	26.7	4.6	26.3	3.6	27.1	5.3
Normal weight (<25.0)	52	(40.0)	25	(41.0)	27	(39.1)
Overweight (25.0-29.9)	49	(37.7)	25	(41.0)	24	(34.8)
Obese (≥ 30.0)	29	(22.3)	11	(18.0)	18	(26.1)
Waist circumference (cm)	90.5	12.9	92.5	11.7	88.7	13.8
Ethnicity						
Caucasian	118	(90.8)	53	(86.9)	65	(94.2)
Latin American	8	(6.2)	5	(8.2)	3	(4.4)
Others	4	(3.1)	3	(4.9)	1	(1.5)
Highest level of education						
High School	9	(6.9)	6	(9.8)	3	(4.4)
CEGEP*	27	(20.8)	9	(14.8)	18	(26.1)
University	94	(72.3)	46	(75.4)	48	(69.6)
Household income, CAD\$**						
< 30,000	26	(20.0)	13	(21.3)	13	(18.8)
30,000 to < 60,000	27	(20.8)	14	(22.9)	13	(18.8)

60,000 to < 90,000	28	(21.5)	11	(18.0)	17	(24.6)
≥ 90,000	38	(29.2)	18	(29.5)	20	(29.0)

*In the Quebec education system, CEGEP is the first level of postsecondary studies and precedes university studies. It includes pre university programs and technical programs.

** A total of 11 participants (5 men and 6 women) did not want to answer the question on income.

Draft

Table 2. Mean daily intakes of sodium, potassium and proteins and correlation coefficients between values derived from three 24-hour recalls using the R24W and from recovery biomarkers (n=126).

	R24W		Biomarkers		% difference	Raw correlation	Deattenuated correlation
	Mean intakes	SD	Mean intakes	SD			
Sodium (mg)	3466	1335	3542	1329	-2.2 %	0.41*	0.48
Potassium (mg)	3303	957	3373	1012	-2.1 %	0.47*	0.56
Proteins (g)	99.3	34.0	104.6	29.8	-5.0 %*	0.61*	0.68

*Student's t test or Pearson correlation with a P value of < 0.05

% difference = percentage of difference between reported intake from R24W and the biomarker-derived value (negative values indicate an underestimation in self-reported intakes and positive values indicate overestimation in self-reported intakes as compared with the reference method).

Table 3. Cross-classification for daily sodium, potassium and protein intakes into quartiles of the distribution using either reported intakes from three 24-hour recalls with the R24W or values derived from recovery biomarkers.

	Same quartile (%)	Adjacent quartiles (%)	± 1 quartile apart (%)	Misclassified (1 st v. 4 th quartile) (%)	K _w
Sodium	39.7	38.9	78.6	4.8	0.30
Potassium	42.9	38.1	81.0	3.2	0.36
Proteins	42.1	43.7	85.7	0.8	0.41

K_w = weighted Kappa

Table 4

Correlation coefficients between body esteem related to appearance and misreporting in sodium, potassium and proteins (n=126).

Misreporting*	Body esteem related to appearance					
	Total sample	Sex		BMI		
		Men	Women	< 25.0	25.0-29.9	≥30.0
	(n=126)	(n=59)	(n=67)	(n=51)	(n=47)	(n=28)
Sodium		-0.13	0.33†	-0.09	-0.09	0.36
Potassium	-0.12					
Proteins	0.04					

* The misreporting value was obtained by subtracting the “true” intake as determined by the biomarker from the reported intake as evaluated by the R24W

†Pearson correlation with a P value < 0.05

Table 5

Statistical test outcomes and interpretation for sodium, potassium and protein intakes

	Validity at a group level				Validity at the individual level		
	Agreement	Agreement	Presence, direction and extent of bias	Association	Agreement	Agreement	
Tests	% difference	Student's <i>t</i> test	Bland-Altman (slope of the regression)	Deattenuated correlation coefficient	Cross-classification	K _w	Total of poor outcomes (/7)
Criteria for a good outcome (G)		P>0.05	P>0.05	≥0.50	≥50% in the same quartile; <10% in opposite quartiles	≥0.61	
Criteria for an acceptable outcome (A)	0.0-10.0%			0.20-0.49		0.20-0.60	
Criteria for a poor outcome (P)	>10.0%	P≤0.05	P≤0.05	<0.20	<50% in the same quartile; ≥10% in opposite quartiles	<0.20	
Sodium	A	G	G	A	P-G	A	1
Potassium	A	G	G	G	P-G	A	1
Proteins	A	P	G	G	P-G	A	2

K_w= weighted Kappa

Figure 1 Bland-Altman plot showing the difference between the self-reported and the biomarker-derived value vs. the average of the two values for sodium intake (mg per day), Bias = -76 mg, $r=0.005$, $p=0.95$, limits of agreements = -2918 to 2766 mg.

Figure 2 Bland-Altman plot showing the difference between the self-reported and the biomarker-derived value vs. the average of the two values for potassium intake (mg per day), Bias = -70 mg, $r=-0.06$, $p=0.48$, limits of agreements = -2050 to 1909 mg.

Figure 3 Bland-Altman plots showing the difference between the self-reported and the biomarker-derived value vs. the average of the two values for protein intake (g per day), bias = -5.2 g, $r=0.17$, $p=0.06$, limits of agreements = -60.7 to 50.2 g.