The American Journal of Clinical Nutrition 117 (2023) 1278–1287

The American Journal of CLINICAL NUTRITION

journal homepage: https://ajcn.nutrition.org/

Original Research Article

Validation of the smartphone-based dietary assessment tool "Traqq" for assessing actual dietary intake by repeated 2-h recalls in adults: comparison with 24-h recalls and urinary biomarkers

Desiree A. Lucassen [1,*], Elske M. Brouwer-Brolsma [1], Hendriek C. Boshuizen [1,2], Monica Mars [1], Johan de Vogel-Van den Bosch [3], Edith JM. Feskens [1]

[1] *Division of Human Nutrition and Health, Wageningen University and Research, Wageningen, The Netherlands;* [2] *National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands;* [3] *Danone Nutricia Research, Utrecht, The Netherlands*

A B S T R A C T

Background: Conventional dietary assessment methods are affected by measurement errors. We developed a smartphone-based 2-h recall (2hR) methodology to reduce participant burden and memory-related bias.
Objective: Assessing the validity of the 2hR method against traditional 24-h recalls (24hRs) and objective biomarkers.
Methods: Dietary intake was assessed in 215 Dutch adults on 6 randomly selected nonconsecutive days (i.e., 3 2hR-days and 3 24hRs) during a 4-wk period. Sixty-three participants provided 4 24-h urine samples, to assess urinary nitrogen and potassium concentrations.
Results: Intake estimates of energy (2052±503 kcal vs. 1976±483 kcal) and nutrients (e.g., protein: 78±23 g vs. 71±19 g; fat: 84±30 g vs. 79±26 g; carbohydrates: 220±60 g vs. 216±60 g) were slightly higher with 2hR-days than with 24hRs. Comparing self-reported protein and potassium intake to urinary nitrogen and potassium concentrations indicated a slightly higher accuracy of 2hR-days than 24hRs (protein: −14% vs. −18%; potassium: −11% vs. −16%). Correlation coefficients between methods ranged from 0.41 to 0.75 for energy and macronutrients and from 0.41 to 0.62 for micronutrients. Generally, regularly consumed food groups showed small differences in intake (<10%) and good correlations (>0.60). Intake of energy, nutrients, and food groups showed similar reproducibility (intraclass correlation coefficient) for 2hR-days and 24hRs.
Conclusions: Comparing 2hR-days with 24hRs showed a relatively similar group-level bias for energy, most nutrients, and food groups. Differences were mostly due to higher intake estimates by 2hR-days. Biomarker comparisons showed less underestimation by 2hR-days as compared with 24hRs, suggesting that 2hR-days are a valid approach to assess the intake of energy, nutrients, and food groups.
This trial was registered at the Dutch Central Committee on Research Involving Human Subjects (CCMO) registry as ABR. No. NL69065.081.19.

Keywords: dietary assessment, technology, innovation, ecological momentary assessment, recall, validation, nutritional biomarkers, healthy adults

Introduction

Research on the role of nutrition in health and disease prevention mostly relies on self-reported dietary intake data, i.e., 24-h recalls (24hRs), food frequency questionnaires (FFQs), or food records. Although these methods are the mainstay of dietary assessment, they have several drawbacks [1,2]. FFQs and 24hRs are retrospective and prone to memory-related bias. In contrast, food records are prospective and prone to reactivity bias, i.e., a user may alter their food intake because they are aware that they are observed or to simplify the recording task. More importantly, irrespective of the method, the researcher and participant burden is high [3].

The recent implementation of new technologies has resulted in the development of multiple web- and smartphone-based dietary assessment tools and substantially improved the quality of dietary assessment (see Eldridge et al. [4] for an overview). Compared with conventional methods, web-based tools have many advantages such as the integration of a fixed food consumption database. This facilitates automatic coding of reported food items, which reduces measurement error, improves accuracy, increases user-friendliness, lowers participant and

Abbreviations: 2hR, 2-h recall; 24hR, 24-h recall; app, application; BMR, basal metabolic rate; DIASS Study, DIetary ASSessment Study; DNFCS, Dutch National Food Consumption Survey; EMA, ecological momentary assessment; ICC, Intraclass Correlation Coefficients; K, potassium; N, nitrogen; PABA, para-aminobenzoic acid; PAL, physical activity level.

* Corresponding author.
E-mail address: desiree.lucassen@wur.nl (D.A. Lucassen).

https://doi.org/10.1016/j.ajcnut.2023.04.008
Received 27 October 2022; Received in revised form 7 February 2023; Accepted 10 April 2023
Available online 11 April 2023

researcher burden, and reduces costs [1,5]. Smartphone-based tools (apps) can even further advance the field as they are perceived as easier to complete, more flexible (i.e., no computer needed), and less burdensome [6]. Moreover, apps have the major advantage of enabling (near) real-time data collection [1,3,7]. This concept is widely used in behavioral and social sciences, where it is referred to as ecological momentary assessment (EMA); repeated real-time assessment of an individual's behavior in their own environment, where the ecological aspect focuses on the individual's "real-world" and the momentary aspect on the individual's current or very recent state [8].

Currently, all available research and commercial dietary assessment apps are based on the food record approach, and are still prone to socially desirable answers and reactivity bias [2,3]. Moreover, there are only a limited number of fully automated (i.e., no manual coding) and validated dietary assessment apps that are appropriate for use in nutrition research. When validated, apps are validated only against traditional self-report methods and not against objective measures such as doubly labeled water or urinary recovery markers (i.e., nitrogen for protein intake, potassium) [9–13].

To further improve the quality of dietary assessment, we recently developed an innovative smartphone-based tool called "Traqq" as described elsewhere [14]. In short, Traqq is a flexible dietary assessment application (app; iOS/Android) that can be tailored to different research questions, e.g., food list, portion size estimation, sampling schemes. In contrast to existing apps, Traqq can be used as both a food record and a recall method. Moreover, the recall module is flexible in terms of recall/reporting period, which enables shorter reporting periods and thus offering the opportunity to deviate from traditional 24hRs to shorter recall periods (e.g., 2 or 4 h) according to the EMA principle. This enables the collection of (near) real-time dietary intake data, which reduces the reliance on memory, takes less time to complete, and consequently should have a lower burden for the respondent, thus increasing the accuracy of the reports.

In this study, we validated the accuracy of the collected dietary intake data using the EMA principle. We compared the use of repeated, consecutive 2hRs on one day for (near) real-time assessment of actual food intake, i.e., energy, macro/micronutrients, and food groups, to traditional 24hRs and urinary recovery biomarkers.

Methods

Participants

An extensive validation study, "The DIetary ASSessment (DIASS) Study", was conducted between June 2019 and May 2020 and included 215 participants who were 18–70 y of age. Participants were eligible for participation if they were able to speak and read Dutch, in possession of a smartphone with Internet plan, metabolically stable (i.e., gained or lost $\leq$3 kg in the past 3 mo), and willing to maintain their dietary habits for the duration of the study. The DIASS study had a crossover design with 2 study periods: 1 study period focused on actual intake (i.e., 2hR-days vs. 24hRs) and 1 on habitual intake (i.e., random 2hRs vs. FFQ). More details on the DIASS study can be found elsewhere [15].

The present study describes the data of the actual study period including participants who completed 3 2hR-days and 3 24hRs (n = 162; Supplemental Figure 1), and 4 24-h urine samples (n = 65; subsample). The DIASS study was approved by the ethics committee of Wageningen University and Research (ABR No.: NL69065.081.19) and conducted according to the guidelines laid down in the Declaration of Helsinki. Written informed consent was obtained from all the participants.

Study design

Food intake was assessed on randomly selected nonconsecutive days over a 4-wk study period. Participants completed 3 2hR-days and 3 24hRs (i.e., either web-based or interviewer administered). Recall days were randomly selected and scheduled over the 4-wk study period using the statistical analyses system (SAS) version 9.4 (SAS Institute Inc), except when in combination with urine collections. Urine collections were matched to the recall days (i.e., 2× to 2hR-day and 2× to 24hR-day), where recall days were randomly scheduled and then preannounced to facilitate the 24-h urine collection on the recall days (i.e., the days to be recalled). In the case of nonresponse, a new day was randomly selected and scheduled.

Methods of dietary assessment

2-h recalls

The smartphone app Traqq was used for the 2hR-days. On 3 random recall days, all participants (n = 162, 100%) received an invitation to report their food intake every 2 h. On average, participants received 8 consecutive 2hR invitations on a recall day, see Figure 1 for an example scheme for a 2hR-day. Notifications were sent at the end of each 2-h interval with a reporting window of 1 h (e.g., interval 06:00–08:00; notification at 08:00; reporting deadline at 09:00). The morning after the recording day, another invitation was sent to report on potential nighttime food intake (e.g., nighttime interval 22:00–06:00; notification at 08:00; reporting deadline at 09:00). The 2hR-day sampling scheme was individualized based on the participant's sleeping pattern, as inquired via the baseline questionnaire, to minimize risk that participants were disturbed while sleeping. To illustrate, if a participant indicated to wake up at 09:00, the first notification was sent at 10:00 instead of 08:00. For all participants, no invitations were sent after 22:00. Participants report their food intake by clicking on the notification or opening the app. Thereafter, the search screen opens and food items can be selected from an extensive food list based on the Dutch Food Composition Database [16]. Subsequently, participants are prompted to report quantity and eating occasions, i.e., breakfast, lunch, dinner, and snack. Quantity can be reported in household measures (e.g., spoon, cup), standard portion sizes (e.g., small, large), or amount in grams. Traqq also contains a "My Dishes" feature where participants can enter all ingredients of a recipe and the amount consumed of the dish, with yield and retention factors automatically being taken into account. The "My Dishes" feature can also be used to create frequently consumed product combinations (e.g., daily breakfast products), which simplifies reporting these items and decreases (mis)calculation errors.

24-h recalls

Participants also completed 3 random nonconsecutive web-based 24hRs (n = 128, 79%) or 3 random nonconsecutive interviewer-administered 24hRs (n = 34, 21%). Web-based 24hRs were administered via Compl-eat, a self-administered web-based dietary 24hR-tool developed by our department based on the automated multiple-pass method, a 5-step method to assist the participant in recalling food intake of the previous 24 h [17,18]. With this method, participants first complete a quick list of consumed foods and subsequently provide

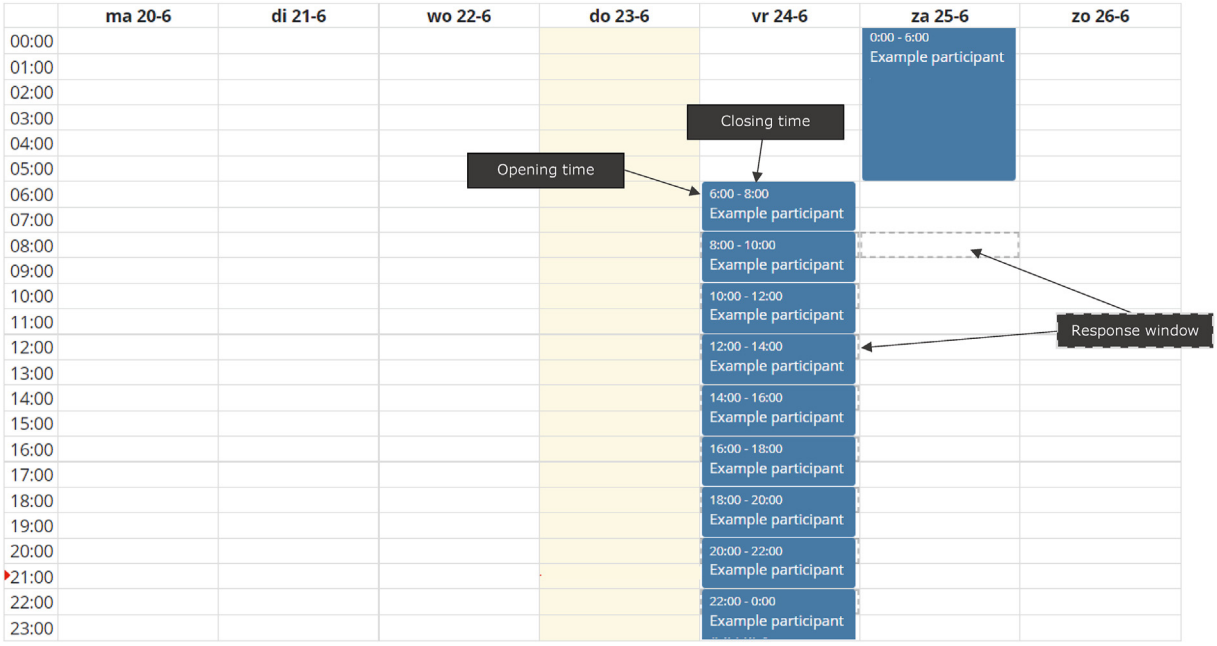

FIGURE 1. An example of a 2hR-day sampling scheme.

detailed information about the type of foods, consumed quantities, and eating occasions [19]. The reporting method in Compl-eat is similar to the reporting method in Traqq. Foods are identified in a food list, and portion sizes are reported in household measures, standard portion sizes, or in grams [17]. Additionally, Compl-eat contains a recipe module similar to the "My Dishes" function of Traqq. Invitations for the web-based 24hRs were sent via e-mail at 06:00 in the morning after the recall day. The 24hR could be completed until midnight the same day.

The interviewer-administered 24hRs were administered via telephone and conducted by trained dieticians using the multiple-pass approach [19]. Methods of portion size estimation included household measures, standard portions, or in grams. The inter-viewer-administered 24hRs were coded by the trained dietician and entered in Compl-eat using the Dutch Food Composition Database [16]. Although the interviewer-administered 24hRs are seen to be the most accurate version of the 24hR method and included to ensure the accuracy of Compl-eat, no major differences in reported intake were found between results of 24hRs administered via Compl-eat and by telephone (unpublished results). Therefore, the reported intakes were combined in the current analyses.

Computation of dietary intake data

Data from both 2hR-days and 24hRs were entered in the computation module of Compl-eat [17]. Total intakes of energy and nutrients were calculated using the Dutch Food Composition Database [16]. Data were thoroughly checked by well-trained dieticians according to a standardized protocol, particularly focusing on reported amounts. Unusual amounts were corrected using standard portion sizes and recipes (e.g., 35 slices of bread was corrected to 1 slice of 35 g).

Urine collection and biochemical analysis of nutritional biomarkers

The urine collection (24h) was performed according to a standardized protocol. Participants received 3-L containers containing the preservative lithium dihydrogenphosphate (25 g), 3 100 mg para-aminobenzoic acid (PABA) tablets (KAL Vitamins), and a questionnaire for each 24h-urine collection. Urine collection started with the second voiding after waking up and was completed with the first voiding after waking up the next day. To verify the completeness of the 24h-urine samples, participants were instructed to ingest 1 PABA tablet with each main meal (i.e., breakfast, lunch, diner), and they were informed that this process was to check the completeness of the collection [20]. Simultaneously, participants were instructed to record the beginning and end time of the 24h-urine collection, time of ingestion of PABA tablets, and any possible deviations from the protocol (e.g., missed urine collection). Urine samples were handed in at the study center where they were weighed, mixed, aliquoted into 5 mL samples and stored at −80ºC until further analysis.

Urinary creatinine was used to assess the completeness of the urine sample. Urinary creatinine concentrations were measured at 520 nm on the Synchrony LX20 by the modified Jaffé procedure using a commercial kit. The 24h-urine collections were classified as complete if they met all of the following criteria: *1*) collection time of 22–26 h, *2*) sample volume $\geq$ 500 mL, *3*) no more than 1 reported missed void, *4*) estimated missed volume $\leq$5% of the total volume, and *5*) creatinine levels of >10 mg/kg for females and >15 mg/kg for males [21]. Of the 259 collected 24h-urine samples, 177 (68%) were classified as complete; only complete samples were used for data analyses.

Urinary 24h-nitrogen (N) was used to estimate protein intake; 24h-N excretion was determined by the Kjeldahl technique (Foss KjeltecTM 2300 analyzer; Foss Analytical). Assuming that approximately 81% of N is excreted via 24h-urine (i.e., 19% fecal and skin losses), and that protein contains 16% of N [22], dietary protein intake was calculated with the following formula:

$$\text{Protein (g/d)} = \text{urinary N (mol/L)} \times \text{volume 24h-urine (L)} \times 14\ \text{(g/mol)} \times 6{\cdot}25 \,/\, 0{\cdot}81$$

Finally, the urinary potassium (K) concentration was used to assess potassium intake. Urinary potassium was measured with an ion-

selective electrode on a Roche 917 analyzer (Roche Diagnostics). Assuming that approximately 77% of potassium is excreted via 24h-urine, 24-h potassium intake was calculated with the following formula:

$$\text{K (mg/d)} = \text{urinary K (mol/L)} \times \text{volume 24h-urine (L)} \times 39\ \text{(g/mol)} \times 1000 / 0.77$$

Other variables

General participant characteristics (i.e., age, gender, educational level, daytime activities, sleeping pattern, intention to maintain current body weight) were acquired using a questionnaire. Height was measured without shoes using a stadiometer (SECA 213; SECO Corp.) and weight was assessed without shoes, heavy clothing, and empty pockets on a digital scale (SECA 877; SECA Corp.). BMI was calculated as weight/height^2.

Physical activity levels (PALs) were assessed over a 7-d period by means of the ActiGraph wGT3X-BT (ActiGraph LLC). The ActiGraph was not worn during showering, bathing, swimming, or contact sports [23]. The accelerometer data were used to determine the participant's percentage of time spent in sedentary, light, moderate, vigorous, and very vigorous activity using the Troiano algorithm [24]. The daytime activity percentages were multiplied with the corresponding PAL according to the guidelines set by the WHO. The WHO guidelines describe a mean PAL, based on factorial calculations of the time spent on activities during the day and the energy cost of those activities (i.e., sedentary: 1.4, light activity: 1.55, moderate: 1.7, vigorous: 1.8, very vigorous: 2.2) [25]. This resulted in an individual PAL for each participant.

At the end of the study period, participants were asked to indicate which dietary assessment method they preferred (i.e., 2hR-days or 24hRs).

Measurement error models

Measurement error models were used to compare the results of 2hR-day assessment with 24hRs and urinary recovery biomarkers (i.e., protein and potassium). Dietary intakes estimated with multiple 24hRs as well as protein and potassium intakes estimated from urinary analysis were assumed to be the best method to approximate true intake [26]. Our measurement error model assumed a linear relationship between the 2hR-days, 24hRs, and the true (unknown) intake. For the 2hR-days and the 24hRs (when not used as reference measurement), intake-related bias, person-specific bias, and a constant bias were assumed. Reference measurements were assumed to be an unbiased measurement. To evaluate the comparability of the 2hR-days and 24hR days with the biomarkers as the reference method, the following measurement error models were used:

$$\text{Reference method } X \text{ (Biomarker)} : X_{ij} = \overline{T}_i + \Delta T_{ij} + \varepsilon_{Xij} \quad (1)$$

$$\text{2hR days or 24hRs (R)}: R_{ij} = \alpha_R + \beta_R\left(\overline{T}_i + \Delta T_{ij}\right) + w_{Ri} + \varepsilon_{Rij} \quad (2)$$

where i is the person, j the occasion, α the constant bias, and β the proportional scaling bias (i.e., intake-related bias). The average (habitual) true intake of person i is $\overline{T}_i$, whereas the true intake on day j is given by $\overline{T}_i + \Delta T_{ij}$. The person-specific bias of the method is given by w_{Ri} and the random error by ε_{Rij}. $\overline{T}_i, \Delta T_{ij}$, w_{Ri}, and ε_{Rij} are each assumed to follow mutually independent normal distribution with variance $\mathrm{var}T$, $\mathrm{var}\Delta T, \mathrm{var}w_{Xi}$, and $\mathrm{var}\varepsilon_{Xij}$. respectively. In this model, the assumptions of negligible error correlation between the reference method and the 2hR-days, and between replicates of the reference method (within the same person), and the absence of proportional scaling bias in the reference method ($\beta_X = 1$) were made to enable the estimation of the model parameters. Note that the assumption of unbiasedness of the reference method is probably not satisfied when using 24hRs as the reference method, but it is reasonable for urinary nitrogen and potassium [27,28]. To evaluate the 2hR-days with the 24hRs as the reference method, we used the same model but without ΔT_{ij} as 2hR and 24hR measurements took place on different days.

Statistical analysis

Results are presented as means with standard deviations (mean ± SD) and frequencies with percentages (n (%)). Under- and over-reporters were identified and excluded based on the Goldberg cutoffs for both methods (n = 16; all underreporters). Participants were identified as dietary under- or overreporters if their ratio of average daily total energy intake to basal metabolic rate (EI:BMR) fell outside an individualized cutoff. BMR was calculated using the Harris and Benedict equation, taking into account gender, age, weight, and height [29]. Individual cutoffs were estimated using the method recommended by Black [30]. For this, the PAL as determined by the accelerometer was used.

To evaluate the 2hR-days against the 24hRs for intake of energy, nutrients, and food groups, multiple analyses were performed [31]. First, absolute intake differences between methods were calculated and expressed as group-level bias ((mean intake 2hR-days) / (mean intake 24hRs) × 100 – 100). A group-level bias of ≤10% was classified as acceptable (i.e., indication of a relatively similar mean intake) [31]. Second, absolute differences between the 2hR-days and the 24hRs were evaluated using paired t-tests. Third, the Spearman correlation coefficients were calculated to assess the strength and direction of the association between the methods. Correlation coefficients of <0.20 were classified as poor, 0.20–0.49 as acceptable, and ≥0.50 as good [31]. Mean ± SD intakes of protein and K, assessed with both 2hR-days and 24hRs, were also compared against the matched 24h-urine samples. Again, group-level bias and paired t-tests were used to evaluate absolute differences, Spearman correlations were calculated to examine the association between the methods, and Bland–Altman plots were created to examine the level of agreement.

Validity coefficients and attenuation factors were calculated using the estimates of the measurement error models [32]. Validity coefficients were estimated to assess the ability of the 2hR-days to rank participants according to their intake and assess the loss of statistical power in the case of 2hR-days would be used to detect a diet–disease association. Validity coefficients of <0.20 were classified as poor, 0.20–0.49 as acceptable, and ≥0.50 as good. Attenuation factors provide information about the extent to which diet–health associations are affected by measurement error, e.g., using the 2hR data instead of true intake. The provided attenuation factors can be used to correct for measurement errors in future studies on diet–disease associations that use 2hR-days to assess dietary intake. An attenuation factor closer to 1 means less attenuation (with 1 representing no attenuation at all). The following equations were used:

$$\text{Validity coefficient} : \rho_{XT} = \sqrt{\frac{\beta_X^2 \mathrm{var}T}{\beta_X^2 \mathrm{var}T + \beta_X^2 \mathrm{var}\Delta T/k + \mathrm{var}\varepsilon_{Xij}/k + \mathrm{var}w_{Xi}}} \quad (3)$$

$$\text{Attenuation factor} : \lambda_X = \frac{\rho_{XT}^2}{\beta_X} \quad (4)$$

where var$T$ is the variance of the habitual true intake, varε_{Xij} the variance of the random within-person error, varw_{Xi} is the variance of the person-specific bias, varΔT is the variance of the day-to-day variation in true intake (not present when 24hR is the reference) and k is the number of replicates of the 2hR-days (k=3 for comparison with 24hR ; k=1 for comparison with biomarker).

The reproducibility was evaluated using the intraclass correlation coefficients (ICC) between the 3 2hR-days and between the 3 24hRs.

$$\text{ICC} : \frac{\text{Between} - \text{person variance}}{(\text{Between} - \text{person variance} + \text{Within} - \text{person variance})} \quad (5)$$

All analyses were performed using IBM SPSS Statistics version 25.0 (SPSS Inc.) and SAS Software version 9.4 (SAS Institute Inc).

Results

Participant characteristics

Participants had a mean ± SD age of 40.4±18.8 y, 73% were female and 62% were highly educated. Overall, participants had a sedentary or lightly active lifestyle, 73% had a healthy BMI (<25 kg/m^2), and 71% did not follow a diet regimen. The majority of the participants preferred the use of 2hR-days over traditional 24hRs (87%) (Table 1).

Accuracy of energy and nutrients reported with 2hR-days compared with 24hRs

Estimated mean intakes of energy and most nutrients were higher with 2hR-days than with 24hRs, as supported by statistically significant paired t-tests, except for carbohydrates (en%), alcohol (en% and g), β-carotene, and vitamin D (Table 2). Overall, differences between the estimated intakes with 2hR-days and 24hRs were small (group-level bias ≤10%). For macronutrients, only reported intakes of animal protein (g) and alcohol (en% and g) had a group-level bias exceeding 10%. For micronutrients, group-level bias exceeded 10% for β-carotene, riboflavin, vitamin B6, vitamin B12, and vitamin D.

Spearman correlation coefficients between 2hR-days and 24hR were acceptable to good for energy and macronutrients, ranging from 0.41 for total protein (en%) to 0.75 for plant-based protein (g). Similarly, for the micronutrients, correlations ranged between 0.41 for β-carotene and 0.62 for potassium.

Validity coefficients for energy and macronutrients were all judged as good (range 0.57–0.86). A similar trend was observed for micronutrients except for vitamin B12 and vitamin D, which were acceptable (0.33 and 0.41, respectively).

Attenuation factors for energy and macronutrients ranged between 0.20 for total protein (en%) and 0.47 for fiber. They varied somewhat more for micronutrients, ranging from 0.04 for vitamin B12 to 0.42 for β-carotene and potassium.

TABLE 1
General characteristics of the participants included in this validation study

	Males (39)	Females (107)	Total (146)
Mean age, y (SD)	46.8 (18.8)	38.1 (18.3)	40.4 (18.8)
The age category (n, (%))			
<25 y	10 (26)	45 (42)	55 (38)
25–50 y	7 (18)	27 (25)	34 (23)
>50 y	22 (56)	35 (33)	57 (39)
Mean BMI, kg/m^2 (SD)	24.8 (4.2)	23.2 (3.5)	23.6 (3.8)
BMI category (n, %)			
<25 kg/m^2	22 (57)	85 (79)	107 (73)
25.0–29.9 kg/m^2	13 (33)	17 (16)	30 (21)
≥30 kg/m^2	4 (10)	5 (5)	9 (6)
Mean estimated BMR, kcal/d (SD)[1]	1780 (144)	1439 (127)	1530 (200)
Mean PAL (SD)	1.46 (0.02)	1.46 (0.01)	1.46 (0.01)
Educational level (n, (%))			
Low (i.e., primary or lower education)	0 (0)	3 (3)	3 (2)
Intermediate (i.e., secondary or higher vocational education)	16 (41)	37 (34)	53 (36)
High (i.e., college or university)	23 (59)	67 (63)	90 (62)
Diet regimen (n, (%))			
Yes, always	3 (8)	22 (21)	25 (17)
Yes, sometimes	4 (10)	14 (13)	18 (12)
Never	32 (82)	71 (66)	103 (71)
Preferred method (n, (%))			
2hR-days	35 (90)	91 (85)	126 (87)
24hRs	4 (10)	11 (10)	15 (10)
Not answered	0 (0)	5 (5)	5 (3)

[1] Based on the Harris and Benedict equation, for males: BMR=(66.4730+(13.7516*weight))+(5.0033*height)−(6.7750*age), and for females: BMR = (655.0955+(9.5634*weight))+(1.8496*height)−(4.6756*age) [29].

Accuracy of food groups reported with 2hR-days compared with 24hRs

Statistically significant paired t-tests were found only for "alcoholic beverages," "grains and cereals," "nonalcoholic beverages," and "nuts, seeds, and snacks" (Table 3). Group-level bias was relatively small for "bread," "cheese," "dairy," "eggs," "fish," "fruit," "meats and poultry," "pastry, cake, and biscuits," "potatoes," "sugar and confectionery," and "vegetables." Group-level bias was large (>10%) for "composite dishes," "nonalcoholic beverages," "nuts, seeds, and snacks," "savory sandwich fillings," and "vegetarian products," where food intake estimates were higher with the 2hR-days than with 24hRs. In contrast, "alcoholic beverages," "fats, oils, and savory sauces," "grains and cereals," "legumes," and "soups" showed a large group-level bias but with higher food intakes estimated with the 24hRs than with 2hR-days.

Accordingly, Spearman correlation coefficients between 2hR-days and 24hRs varied across food groups as well with the majority being higher than 0.52. Remaining food groups were classified being acceptable, except for "composite dishes," "fish," and "soups" that were classified as poor (0.16, 0.14, and 0.09, respectively).

In agreement, the validity coefficients and attenuation factors were classified as good for most food groups. Validity coefficients for "composite dishes"'(0.15) and "fish" (0.11) were low. Attenuation factors ranged from 0.02 for "composite dishes" and "fish" to 0.54 for "nonalcoholic beverages."

Comparison of self-reported intake with urinary recovery biomarkers

Compared with urinary recovery biomarkers, 2hR-days showed slightly lower underestimation than 24hRs for both protein intake (−13.7% vs. −17.9%) and potassium intake (−11.0% vs. −16.0%) (Table 4).

For protein intake, the correlations between the self-report measures and urinary nitrogen assessed on the same day were similar and classified as good (0.59 for 2hR-days vs. biomarker; 0.57 for 24hRs vs. biomarker). For potassium, a good correlation was found for 2hR-days compared with urinary potassium (0.62), which was lower and acceptable for 24hRs (0.45).

TABLE 2
Mean energy and nutrient intakes were assessed by 3 2hR-days and 3 24hRs with corresponding group-level bias, paired *t*-tests, Spearman correlation coefficients between the 2hR-days and 24hRs, and validity coefficients and attenuation factors of the 2hR-days with the 24hRs as the reference method ($n = 146$).

	2hRs	24hRs	Group-level bias (%)[1]	P [2]	Correlation coefficient [3] (95% CI)	Validity coefficient (95% CI)	Attenuation factor (95% CI)
	Mean ± SD	Mean ± SD					
Energy (kcal)	2052 ± 503	1976 ± 483	3.9	0.04	0.63 (0.50, 0.72)	0.72 (0.62, 0.83)	0.36 (0.28, 0.45)
Protein (en%)	16 ± 3	15 ± 3	4.6	<0.01	0.41 (0.26, 0.54)	0.69 (0.48, 0.89)	0.20 (0.13, 0.27)
Protein (g)	78 ± 23	71 ± 19	9.1	<0.001	0.62 (0.49, 0.72)	0.72 (0.61, 0.84)	0.29 (0.22, 0.36)
Plant-based proteins (g)	36 ± 11	34 ± 11	3.7	0.04	0.75 (0.66, 0.82)	0.84 (0.79, 0.89)	0.44 (0.36, 0.53)
Animal protein (g)	42 ± 21	37 ± 17	14.4	<0.001	0.64 (0.53, 0.74)	0.75 (0.63, 0.86)	0.27 (0.20, 0.33)
Fat (en%)	35 ± 7	35 ± 6	2.0	0.17	0.45 (0.31, 0.58)	0.78 (0.61, 0.95)	0.31 (0.22, 0.39)
Fat (g)	84 ± 30	79 ± 26	6.5	0.03	0.54 (0.40, 0.65)	0.66 (0.52, 0.79)	0.29 (0.21, 0.37)
SFA (g)	30 ± 12	29 ± 11	4.9	0.12	0.65 (0.53, 0.74)	0.68 (0.54, 0.82)	0.28 (0.20, 0.36)
MUFA (g)	30 ± 12	28 ± 11	7.8	0.02	0.50 (0.36, 0.62)	0.78 (0.64, 0.91)	0.32 (0.24, 0.40)
PUFA (g)	16 ± 7	15 ± 6	6.4	0.09	0.45 (0.30, 0.57)	0.57 (0.41, 0.73)	0.22 (0.14, 0.30)
Cholesterol (mg)	202 ± 113	194 ± 108	4.3	0.35	0.54 (0.40, 0.65)	0.73 (0.64, 0.81)	0.26 (0.18, 0.33)
Carbohydrates (en%)	44 ± 7	45 ± 7	−1.8	0.09	0.66 (0.55, 0.75)	0.86 (0.81, 0.90)	0.42 (0.34, 0.51)
Carbohydrate (g)	220 ± 60	216 ± 60	1.6	0.31	0.70 (0.59, 0.78)	0.84 (0.80, 0.89)	0.46 (0.38, 0.54)
Mono/disaccharides (g)	94 ± 34	90 ± 34	4.2	0.09	0.67 (0.55, 0.75)	0.78 (0.68, 0.88)	0.43 (0.34, 0.52)
Polysaccharides (g)	126 ± 41	126 ± 38	0.02	0.99	0.72 (0.62, 0.80)	0.83 (0.78, 0.88)	0.39 (0.31, 0.47)
Fiber (g)	24 ± 7	23 ± 7	0.1	0.95	0.66 (0.55, 0.75)	0.85 (0.77, 0.93)	0.47 (0.38, 0.55)
Alcohol (en%)	2.5 ± 3.6	3.0 ± 4.1	−17.3	<0.001	0.55 (0.41, 0.66)	0.76 (0.69, 0.83)	0.38 (0.29, 0.47)
Alcohol (g)	7 ± 11	9 ± 12	−15.7	0.12	0.54 (0.41, 0.65)	0.76 (0.68, 0.83)	0.35 (0.26, 0.45)
Ca (mg)	989 ± 384	906 ± 296	9.2	<0.01	0.56 (0.43, 0.67)	0.67 (0.53, 0.80)	0.23 (0.16, 0.29)
Fe (mg)	11 ± 4	10 ± 3	5.7	0.03	0.57 (0.44, 0.68)	0.65 (0.50, 0.79)	0.27 (0.19, 0.35)
K (mg)	3114 ± 820	2994 ± 778	4.0	0.03	0.62 (0.50, 0.72)	0.78 (0.68, 0.88)	0.42 (0.33, 0.50)
β-Carotene (μg)	2922 ± 3582	3299 ± 4470	−11.4	0.23	0.41 (0.26, 0.54)	0.73 (−1.03, 2.48)	0.42 (−0.27, 1.12)
Thiamin (mg)	0.97 ± 0.32	0.92 ± 0.28	5.3	0.07	0.48 (0.33, 0.60)	0.66 (0.55, 0.77)	0.18 (0.12, 0.24)
Riboflavin (mg)	1.39 ± 0.49	1.25 ± 0.41	10.7	<0.001	0.61 (0.48, 0.71)	0.75 (0.63, 0.86)	0.32 (0.24, 0.39)
Vitamin B6 (mg)	1.40 ± 0.43	1.18 ± 0.53	18.7	<0.001	0.47 (0.32, 0.59)	0.71 (0.54, 0.87)	0.25 (0.17, 0.33)
Vitamin B12 (μg)	3.83 ± 4.44	3.40 ± 2.08	12.9	0.23	0.50 (0.36, 0.62)	0.33 (0.14, 0.53)	0.04 (0.01, 0.07)
Vitamin C (mg)	92 ± 49	88 ± 45	4.7	0.32	0.48 (0.34, 0.60)	0.66 (0.46, 0.89)	0.20 (0.13, 0.28)
Vitamin D (μg)	2.15 ± 1.29	2.44 ± 1.60	−11.9	0.04	0.45 (0.30, 0.57)	0.41 (0.20, 0.62)	0.16 (0.07, 0.26)
Vitamin E (mg)	12 ± 5	12 ± 4	3.5	0.28	0.42 (0.27, 0.55)	0.68 (0.50, 0.87)	0.25 (0.17, 0.33)
Folate (μg)	259 ± 77	258 ± 80	0.3	0.91	0.58 (0.45, 0.68)	0.74 (0.61, 0.87)	0.33 (0.24, 0.41)

[1] Group-level bias = (mean 2hR-days) / (mean 24hRs) × 100 − 100.
[2] Paired *t*-test between mean intake assessed with 2hRs and 24hRs.
[3] Spearman correlation between the mean of 3 2hR-days and 3 24hRs.

Validity coefficients for protein intake (0.41 for 2hR-days vs. biomarker; 0.44 for 24hRs vs. biomarker) were acceptable. For potassium, a good validity coefficient was found for 2hR-days compared with urinary potassium (0.54), and an acceptable correlation for 24hRs (0.49).

Attenuation factors were also relatively similar for both methods (protein: 0.27 for 2hR-days vs. biomarker; 0.38 for 24hRs vs. biomarker, and potassium: 0.52 for 2hR-days vs. biomarker; 0.55 for 24hRs).

The Bland–Altman plots showed relatively similar patterns when comparing intakes of protein and potassium for both 2hR-days and 24hR against urinary biomarkers (Figure 2). The regression line of differences was insignificant for both urinary protein comparisons (2hR-days $\beta = 0.19$, $P = 0.14$; 24hRs $\beta = 0.01$, $P = 0.91$) and for the urinary potassium and 24hR comparison ($\beta = -0.24$, $P = 0.12$), whereas the regression line of differences for the comparison with 2hR-days was significant ($\beta = -0.28$, $P = 0.01$). Yet, both urinary potassium comparisons showed a similar pattern and indicated that differences between the methods decreased, whereas the intake increased.

Reproducibility of the 2hR-days compared with the 24hRs

The ICC for repeated 2hR-days showed acceptable reproducibility for energy and macronutrients (range 0.27–0.49), which is similar to the reproducibility observed for the repeated 24hRs (range 0.21–0.51) (Supplemental Table S1). For micronutrients, the variation in ICC was larger, with acceptable-to-good reproducibility between 2hR-days, except for vitamin B12 (0.04) and vitamin D (0.19). The 24hRs also showed an acceptable reproducibility for all micronutrients (range 0.21-0.46), except for β-Carotene that had a good reproducibility (0.80).

The reproducibility for food groups was similar for 2hR-days and 24hR (Supplemental Table S2), except for the group "fats, oils, and savory sauces" (0.31 and 0.14, respectively), "grains and cereals" (0.33 and 0.18,

TABLE 3
Mean intake of food groups (g/d) was assessed by 3 2hR-days and 3 24hRs with corresponding group-level bias, paired *t*-tests, Spearman correlation coefficients between the 2hR-days and 24hRs, and validity coefficients and attenuation factors of the 2hR-days with the 24hRs as the reference method ($n = 146$).

	2hRs	24hRs	Group-level bias (%)[1]	P [2]	Correlation coefficient [3] (95% CI)	Validity coefficient (95% CI)	Attenuation factor (95% CI)
	Mean ± SD	Mean ± SD					
Alcoholic beverages	109 ± 177	141 ± 224	−22.8	<0.05	0.52 (0.38, 0.63)	0.67 (0.53, 0.80)	0.35 (0.25, 0.45)
Bread	120 ± 60	122 ± 60	−2.1	0.44	0.70 (0.59, 0.78)	0.83 (0.79, 0.88)	0.42 (0.33, 0.50)
Cheese	31 ± 33	30 ± 23	5.7	0.52	0.49 (0.34, 0.61)	0.66 (0.55, 0.77)	0.14 (0.09, 0.19)
Composite dishes	58 ± 77	45 ± 74	27.8	0.14	0.16 (−0.01, 0.32)	0.15 (−0.20, 0.51)	0.02 (−0.03, 0.06)
Dairy	248 ± 195	235 ± 173	5.4	0.24	0.75 (0.66, 0.82)	0.87 (0.83, 0.91)	0.45 (0.38, 0.53)
Eggs	17 ± 22	17 ± 22	0.8	0.95	0.36 (0.20, 0.49)	0.58 (0.45, 0.72)	0.14 (0.08, 0.21)
Fats, oils, and savory sauces	28 ± 24	32 ± 30	−12.7	0.12	0.43 (0.29, 0.56)	0.55 (0.27, 0.83)	0.21 (0.10, 0.32)
Fish	15 ± 28	14 ± 29	8.8	0.71	0.14 (−0.03, 0.29)	0.11 (−0.16, 0.39)	0.02 (−0.03, 0.08)
Fruit	179 ± 136	173 ± 119	3.1	0.56	0.62 (0.50, 0.72)	0.80 (0.68, 0.92)	0.33 (0.25, 0.41)
Grains and cereals	58 ± 58	71 ± 65	−17.7	0.02	0.53 (0.39, 0.64)	0.72 (0.50, 0.94)	0.28 (0.18, 0.37)
Legumes	7 ± 18	9 ± 21	−29.6	0.18	0.21 (0.05, 0.37)	0.34 (−0.02, 0.70)	0.07 (−0.01, 0.15)
Meat and poultry	58 ± 62	60 ± 55	−4.1	0.60	0.62 (0.50, 0.72)	0.73 (0.57, 0.90)	0.25 (0.17, 0.32)
Nonalcoholic beverages	1670 ± 788	1405 ± 644	18.8	<0.001	0.77 (0.69, 0.84)	0.87 (0.81, 0.92)	0.54 (0.47, 0.61)
Nuts, seeds, and snacks	31 ± 36	25 ± 26	26.4	0.04	0.29 (0.13, 0.43)	0.49 (0.29, 0.69)	0.06 (0.02, 0.10)
Pastry, cake, and biscuits	42 ± 37	42 ± 37	−0.1	0.99	0.53 (0.39, 0.64)	0.69 (0.59, 0.79)	0.22 (0.14, 0.29)
Potatoes	47 ± 49	49 ± 51	−3.4	0.74	0.26 (0.10, 0.41)	0.48 (0.30, 0.66)	0.08 (0.03, 0.14)
Savory sandwich fillings	12 ± 16	11 ± 16	13.1	0.23	0.53 (0.40, 0.65)	0.69 (0.60, 0.78)	0.23 (0.16, 0.30)
Soups	31 ± 61	43 ± 77	−27.9	0.11	0.09 (−0.07, 0.25)	0.24 (−0.02, 0.50)	0.08 (−0.01, 0.16)
Sugar and confectionery	27 ± 25	26 ± 26	4.1	0.58	0.65 (0.53, 0.74)	0.77 (0.64, 0.90)	0.35 (0.26, 0.44)
Vegetables	175 ± 139	181 ± 126	−3.1	0.52	0.62 (0.49, 0.72)	0.84 (0.79, 0.89)	0.37 (0.29, 0.44)
Vegetarian products	32 ± 61	29 ± 58	13.3	0.27	0.58 (0.45, 0.68)	0.88 (0.84, 0.92)	0.49 (0.41, 0.57)

[1] Group-level bias = (mean 2hR-days) / (mean 24hRs) × 100 − 100.
[2] Paired *t*-test between mean intake assessed with 2hRs and 24hRs.
[3] Spearman correlation between the mean of 3 2hR-days and 3 24hRs.

TABLE 4
Self-reported intake of protein and potassium as compared with their urinary recovery biomarker, with corresponding group-level bias, paired-t-tests, Spearman correlation coefficient, validity coefficients and attenuation factors of the 2hR-days or the 24hRs with the urinary recovery biomarker as the reference method.

	Self-reported intake	Urinary biomarker	Group-level bias (%)[1]	P [2]	Correlation coefficient [3] (95% CI)	Validity coefficient (95% CI)	Attenuation factor (95% CI)
	Mean ± SD	Mean ± SD					
Protein (g/d)							
2hR-days ($n = 87$)	80.1 ± 31.7	92.8 ± 27.5	−13.7	<0.001	0.59 (0.42, 0.72)	0.41 (0.18, 0.63)	0.27 (0.09, 0.46)
24hRs ($n = 75$)	71.6 ± 24.2	87.2 ± 23.9	−17.9	<0.001	0.57 (0.37, 0.71)	0.44 (0.25, 0.62)	0.38 (0.17, 0.58)
Potassium (mg/d)							
2hR-days ($n = 87$)	3429 ± 1161	3852 ± 1453	−11.0	<0.01	0.62 (0.46, 0.74)	0.54 (0.42, 0.67)	0.52 (0.32, 0.73)
24hRs ($n = 75$)	3060 ± 1085	3645 ± 1285	−16.0	<0.001	0.45 (0.24, 0.37)	0.49 (0.32, 0.66)	0.55 (0.26, 0.84)

[1] Group-level bias = (mean self-reported intake) / (mean biomarker) × 100 − 100.
[2] Paired *t*-test between mean intake assessed with 2hRs and 24hRs.
[3] Spearman correlation between individual measurements on the same day.

respectively), and "vegetarian products" (0.53 and 0.48, respectively), where the ICC was higher for the 2hR-days than for the 24hRs.

Discussion

We developed a new smartphone-based 2hR methodology by combining traditional dietary assessment approaches and EMA principles with the assumption that 2hR time-windows are less sensitive to memory-related errors and less obtrusive than traditional approaches. We showed that 2hR-days provide higher intake estimates for energy, most nutrients, and most food groups compared with validated 24hRs. Validation against objective urinary biomarkers for protein and potassium intake further substantiated these findings by showing that 2hR-days intake estimates were also more accurate, i.e., slightly closer to the "true intake" than 24hRs intake estimates. Finally, most participants preferred 2hR-days over traditional 24hRs.

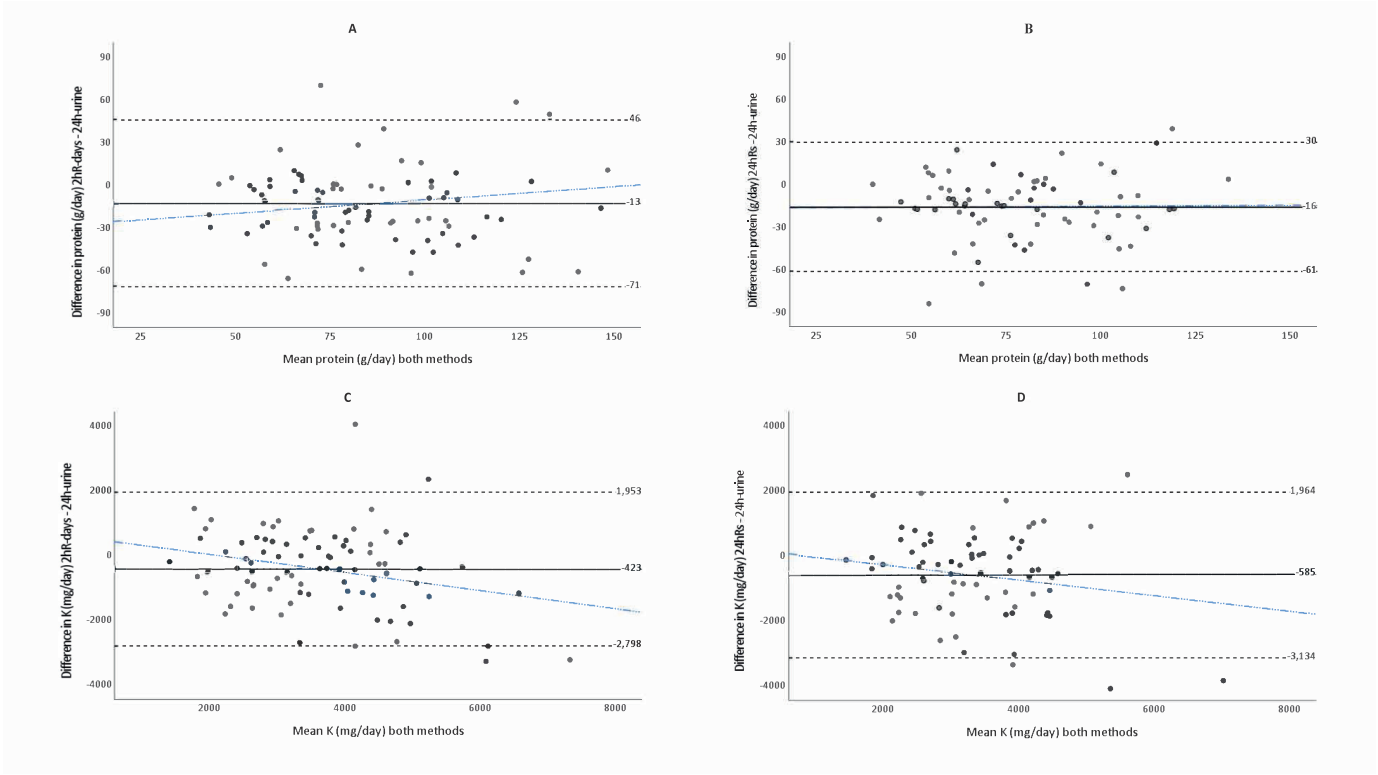

FIGURE 2. Bland–Altman plots of the differences in intake estimated with the self-report measure and the biomarker, plotted against the mean of both methods (g/d). Mean difference (solid line), 95% limits of agreement (1.96 × SD of mean difference; dashed line), and linear regression line (blue dashed line) are included. Protein intake differences are plotted for (A) 2hR-days vs. biomarkers and (B) 24hRs vs. biomarker; potassium intake differences are plotted in (C) 2hR-days vs. biomarkers, and (D) 24hRs vs. biomarker.

Our results showed low group-level bias for energy and most macronutrients (≤10%), except for animal protein and alcohol (14%–17%). In terms of protein, the Dutch National Food Consumption (DNFCS) data showed a mean animal protein intake of 51 g/d (95% CI: 51, 51 g/d) of the average Dutch population, which is closer to our 2hR-days (i.e., 42±21 g/d) than our 24hRs (37±17 g/d) estimates [33]. In addition, the total protein intake estimate by the 2hR-days was closer to the total protein intake estimate based on urinary nitrogen excretion than 24hRs (−14% vs. −18%, respectively) and had smaller limits of agreement; this suggests that 2hR-days may provide a more precise and accurate estimate of protein intake than 24hRs. In terms of alcohol, DNFCS data showed a 11 g/d (95% CI: 10, 12 g/d) mean intake estimate, which is close to our 24hR estimate (9±12g/d) but higher than our 2hR-day intake estimate (7±11 g/d) [33]. Accordingly, a similar difference between 2hR-days and 24hR was observed for the food group "alcoholic beverages" (109 g/d vs. 141 g/d). As it is well known that alcohol consumption varies highly across days [34], it is difficult to determine the exact origin and direction of this difference between 2hR-days and 24hR [35]. A possible explanation could be the short reporting deadline of the nighttime recall (i.e., 1 h), which is easily missed after a late night (e.g., a party). In contrast, a 24hR remains open for an entire day, giving participants more time to respond after a night out. However, this is an assumption and more research is needed to determine an optimal sampling scheme to ensure that we capture episodically consumed foods such as alcoholic beverages.

Differences in absolute micronutrient intakes were relatively small; only β-carotene, riboflavin, vitamin B6, vitamin B12, and vitamin D group-level bias slightly exceeded 10%. For vitamin B12 and vitamin D, the ICCs showed a larger variation in the reported intake between the 2hR-days (0.04 and 0.19, respectively) and 24hRs (0.26 and 0.25, respectively). Therefore, these differences could be caused by day-to-day variation in, for instance, reported "fish," and do not necessarily imply that a method performs badly. In terms of correlations (range 0.41–0.61), results are well within the acceptable range suggested by Willet and colleagues (0.4–0.7) [34].

Group-level bias was low for the majority of regularly consumed food groups (≥5 d/wk according to the DNFCS), i.e., "bread," "dairy," "fruit," "meats and poultry," and "vegetables" [33], suggesting at least similar accuracy for 2hR-days and 24hRs [4,17], which is further underlined by good correlations for these food groups (≥0.6). Larger differences were observed for regularly consumed food groups "fats, oils, and savory sauces," "grains and cereals," and "nonalcoholic beverages." Intake estimates were lower for "fats, oils, and savory sauces" and "grains and cereals" with 2hR-days than 24hRs, yet the ICCs were higher for the 2hR-days, where higher ICCs suggest better recollection and thus reporting with shorter recall periods. In contrast, intake estimates for nonalcoholic beverages were higher for 2hR-days than 24hRs, which may be explained by the fact that nonalcoholic beverages are often consumed throughout the day, not always linked to specific eating occasions, and thus more difficult to recall with a 24hRs than 2hRs.

The results of the attenuation factors for nutrients and food groups were in line with the correlation coefficients, with an attenuation factor of 0.54 for "nonalcoholic beverages" being the highest. Relatively similar attenuation factors were observed for protein (2hR-days: 0.27 vs. 24hR: 0.38) and potassium (2hR-days: 0.52 vs. 24hRs: 0.55), with the urinary biomarker as reference. To illustrate, Freedman and colleagues observed a similar range of attenuation factors for protein in

their biomarker analyses (0.14–0.54) [36]). The attenuation factors for self-reported protein and potassium intake (i.e., using the 24hRs as the reference method) were 0.29 and 0.42, respectively. Although attenuation factors for self-reported protein are similar to the biomarker comparisons, larger differences are seen for self-reported potassium, this is possibly due to day-to-day variation.

As far as we know, Traqq is the first recall-based dietary assessment app with a 2hR approach. However, there are several validation studies of food record-based apps against 24hRs. Although validation studies using objective markers are lacking [37], evaluation studies mostly show lower intake estimates of energy and macronutrients by food record apps compared with 24hRs [9–12,38,39]. In contrast, our results show mostly higher intake estimates of the 2hR approach, which may relate to the fact that our approach minimizes reactivity bias while limiting memory-related bias owing to the relatively short reporting window of the recall method. Specifically, with the 2hR, participants register their food intake every 2 h of the day and immediately send it to an external server after which data are not visible to the participant. With regular food records, food intake reports remain visible throughout the day, which increases the likelihood of introducing reactivity bias. Finally, these data may suggest that our smartphone-based 2hR approach is able to provide a more accurate (near) real-time assessment of dietary intake compared with food record-based apps.

Although the design of this validation study is well thoughtout, there are still some methodological issues that warrant discussion. First, we used a validated 24hR method as well as objective urinary biomarkers as a reference method to validate the 2hR-day approach. As both 2hR and 24hR rely on memory, the same food composition tables, and similar portion size suggestions, differences in these data may be inflated by correlated errors. However, validations of 2hR data against urinary recovery biomarkers for protein and potassium show similar trends and thus confirm the differences between 2hR and 24hR. Second, the majority of our sample consisted of highly educated females, which may have affected the generalizability of the validation results. Therefore, additional validation with a more diverse population may be needed. Nevertheless, considering that 2hRs have lower reliance on memory makes it a promising approach for use in populations with decreasing cognitive abilities. The strength of the current study is that we used multiple tests to assess the validity of 2hR-days, which has been suggested as the most optimal approach to assess the validity of a dietary assessment method [31,35].

In conclusion, the comparison of 2hR-days with 24hRs showed relatively similar group-level bias for energy, most nutrients, and food groups. Differences between the 2 methods were mostly due to higher intake estimates by 2hR-days, which was also the preferred method by the majority of the participants. Finally, comparisons with biomarkers showed less underestimation of protein and potassium intake by 2hR-days compared with 24hRs, suggesting that using 2hR-days is a valid approach to assess the intake of energy, nutrients, and food groups.

Acknowledgments

We thank all the participants for their valuable contribution to this study. We also thank all the research staff and students that were involved in the execution of this study.

Author contributions

The authors' responsibilities were as follows – DAL, EMBB, and EJMF designed research; DAL conducted research; DAL and HCB analyzed data; DAL wrote the manuscript; DAL had primary responsibility for the final content. All authors read, contributed, and approved the final manuscript.

Author disclosures

The authors declare no conflicts of interest.

Funding

The development of Traqq, the 2-h recall methodology, and the DIASS study was executed by Wageningen University and Research and partly funded by the Ministry of Agriculture, Nature and Food Quality and industry, in the context of TKI Agri&Food PPS – project Smart Food Intake (AF16096) and by the 4 Dutch Technical Universities, 4TU – Pride and Prejudice program.

Data availability

Data described in the manuscript, code book, and analytic code will be made available upon request pending application and approval.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajcnut.2023.04.008.

References

[1] J.S. Shim, K. Oh, H.C. Kim, Dietary assessment methods in epidemiologic studies, Epidemiol Health 36 (2014), e2014009, https://doi.org/10.4178/epih/e2014009.

[2] A.F. Subar, L.S. Freedman, J.A. Tooze, S.I. Kirkpatrick, C. Boushey, M.L. Neuhouser, et al., Addressing current criticism regarding the value of self-report dietary data, J. Nutr. 145(12) 2639–2645. doi: 10.3945/jn.115.219634.

[3] F.E. Thompson, A.F. Subar, Dietary assessment methodology, in: A.M. Coulston, C.J. Boushey, M.G. Ferruzi, L.M. Delahanty (Eds.), Nutrition in the prevention and treatment of disease, 4th ed, Elsevier Inc., Amsterdam, The Netherlands, 2017, pp. 5–48.

[4] A.L. Eldridge, C. Piernas, A.K. Illner, M.J. Gibney, M.A. Gurinovic, J.H.M. de Vries, et al., Evaluation of new technology-based tools for dietary intake assessment-an ILSI Europe dietary intake and exposure task force evaluation, Nutrients 11 (1) (2018) 55, https://doi.org/10.3390/nu11010055.

[5] E.M. Brouwer-Brolsma, D. Lucassen, M.G. De Rijk, A. Slotegraaf, C. Perenboom, K. Borgonjen, et al., Dietary intake assessment: from traditional paper–pencil questionnaires to technology-based tools. Environmental software systems data science in action, in: Proceedings: 13th IFIP WG 511 International Symposium, ISESS 2020 13, Springer, Wageningen, The Netherlands, 2020, pp. 7–23. Feb 5–7, 2020.

[6] D.B. Sharp, M. Allman-Farinelli, Feasibility and validity of mobile phones to assess dietary intake, Nutrition 30 (11–12) (2014) 1257–1266.

[7] A. Stone, S. Shiffman, A. Atienza, L. Nebeling, The science of real-time data capture: self-reports in health research, Oxford University Press, New York, 2007.

[8] S. Shiffman, A.A. Stone, M.R. Hufford, Ecological momentary assessment, Annu. Rev. Clin. Psychol. 4 (1) (2008) 1–32, https://doi.org/10.1146/annurev.clinpsy.3.022806.091415.

[9] S. Bucher Della Torre, I. Carrard, E. Farina, B. Danuser, M. Kruseman, Development and evaluation of e-CA, an electronic mobile-based food record, Nutrients 9 (1) (2017) 76, https://doi.org/10.3390/nu9010076.

[10] M.C. Carter, V.J. Burley, C. Nykjaer, J.E. Cade, My Meal Mate" (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss, Br. J. Nutr. 109 (3) (2013) 539–546, https://doi.org/10.1017/S0007114512001353.

[11] L. Wellard-Cole, J. Chen, A. Davies, A. Wong, S. Huynh, A. Rangan, et al., Relative validity of the eat and track (EaT) smartphone app for collection of dietary intake data in 18-to-30-year olds, Nutrients 11 (3) (2019) 621, https://doi.org/10.3390/nu11030621.

[12] A.M. Rangan, S. O'Connor, V. Giannelli, M.L. Yap, L.M. Tang, R. Roy, et al., Electronic dietary intake assessment (e-DIA): comparison of a mobile phone digital entry app for dietary data collection with 24-hour dietary recalls, JMIR Mhealth Uhealth 3 (4) (2015) e98, https://doi.org/10.2196/mhealth.4613.

[13] G.L. Ambrosini, M. Hurworth, R. Giglia, G. Trapp, P. Strauss, Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls, Nutr. J. 17 (1) (2018) 5, https://doi.org/10.1186/s12937-018-0315-4.
[14] D.A. Lucassen, E.M. Brouwer-Brolsma, A.M. van de Wiel, E. Siebelink, E.J.M. Feskens, Iterative development of an innovative smartphone-based dietary assessment tool: Traqq, JoVE 169 (2019), e62032, https://doi.org/10.3791/62032.
[15] D.A. Lucassen, E.M. Brouwer-Brolsma, A.I. Slotegraaf, E. Kok, E.J. Feskens, DIetary ASSessment (DIASS) study: design of an evaluation study to assess validity, usability and perceived burden of an innovative dietary assessment methodology, Nutrients 14 (6) (2022) 1156, https://doi.org/10.3390/nu14061156.
[16] RIVM N-online [Internet]. Version 2016/5.0 [cited August 15, 2018]. Available from: https://nevo-online.rivm.nl/.
[17] S. Meijboom, M.T. van Houts-Streppel, C. Perenboom, E. Siebelink, A.M. van de Wiel, A. Geelen, et al., Evaluation of dietary intake was assessed by the Dutch self-administered web-based dietary 24-h recall tool (Compl-eat™) against interviewer-administered telephone-based 24-h recalls, J. Nutr. Sci. 6 (2017) e49, https://doi.org/10.1017/jns.2017.45.
[18] J.E. Cade, Measuring diet in the 21st century: use of new technologies, Proc. Nutr. Soc. 76 (3) (2017) 276–282, https://doi.org/10.1017/S0029665116002883.
[19] J.M. Conway, L.A. Ingwersen, B.T. Vinyard, A.J. Moshfegh, The effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and non-obese women, Am. J. Clin. Nutr. 77 (5) (2003) 1171–1178, https://doi.org/10.1093/ajcn/77.5.1171.
[20] A.F. Subar, D. Midthune, N. Tasevska, V. Kipnis, L.S. Freedman, Checking for completeness of 24-h urine collection using para-amino benzoic acid not necessary in the observing protein and energy nutrition study, Eur. J. Clin. Nutr. 67 (8) (2013) 863–867, https://doi.org/10.1038/ejcn.2013.62.
[21] C. Lassalle, K. Castetbon, F. Laporte, G.M. Camilleri, V. Deschamps, M. Vernay, et al., Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers, Br. J. Nutr. 113 (6) (2015) 953–962, https://doi.org/10.1017/S0007114515000057.
[22] S.A. Bingham, Urine nitrogen as a biomarker for the validation of dietary protein intake, J. Nutr. 133 (3) (2003) 921S–924S, https://doi.org/10.1093/jn/133.3.921S.
[23] A.P. Hills, N. Mokhtar, N.M. Byrne, Assessment of physical activity and energy expenditure: an overview of objective measures, Front. Nutr. 1 (2014) 5, https://doi.org/10.3389/fnut.2014.00005.
[24] R.P. Troiano, D. Berrigan, K.W. Dodd, L.C. Masse, T. Tilert, M. McDowell, Physical activity in the United States was measured by an accelerometer, Med Sci Sports Exerc 40 (1) (2008) 181–188, https://doi.org/10.1249/mss.0b013e31815a51b3.
[25] Joint FAO/WHO/UNU expert consultation. Energy and protein requirements: a report of a Joint FAO/WHO/UNU Expert Consultation [held in Rome from 5 to 17 October 1981], in: World Health Organization, Technical Report Series, World Health Organization, Geneva, 1985.
[26] M. Jenab, N. Slimani, M. Bictash, P. Ferrari, S.A. Bingham, Biomarkers in nutritional epidemiology: applications, needs and new horizons, Hum. Genet. 125 (5–6) (2009) 507–525, https://doi.org/10.1007/s00439-009-0662-5.
[27] N. Tasevska, S.A. Runswick, S.A. Bingham, Urinary potassium is as reliable as urinary nitrogen for use as a recovery biomarker in dietary studies of free living individuals, J. Nutr. 136 (5) (2006) 1334–1340, https://doi.org/10.1093/jn/136.5.1334.
[28] S.A. Bingham, J.H. Cummings, Urine nitrogen as an independent validatory measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet, Am. J. Clin. Nutr. 42 (6) (1985) 1276–1289, https://doi.org/10.1093/ajcn/42.6.1276.
[29] C.J.K. Henry, Basal metabolic rate studies in humans: measurement and development of new equations, Public Health Nutr 8 (7a) (2005) 1133–1152, https://doi.org/10.1079/phn2005801.
[30] A.E. Black, Critical evaluation of energy intake using the Goldberg cut-off for energy intake: basal metabolic rate. A practical guide to its calculation, use and limitations, Int. J. Obesity 24 (9) (2000) 1119–1130, https://doi.org/10.1038/sj.ijo.0801376.
[31] M.J. Lombard, N.P. Steyn, K.E. Charlton, M. Senekal, Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods, Nutr. J. 14 (2015) 40, https://doi.org/10.1186/s12937-015-0027-y.
[32] R.J. Kaaks, Biochemical markers as additional measurements in studies of the accuracy of dietary questionnaire measurements: conceptual issues, Am. J. Clin. Nutr. 65 (4) (1997) 1232S–1239S, https://doi.org/10.1093/ajcn/65.4.1232S.
[33] C. Van Rossum, E. Buurma-Rethans, C. Dinnissen, M. Beukers, H. Brants, M. Ocké, The diet of the Dutch: results of the Dutch National Food Consumption survey 2012–2016. [Bilthoven, rijksinstituut voor volksgezondheid en milieu RIVM, The Netherlands, 2020. https://rivm.openrepository.com/handle/10029/624455.
[34] W. Willett, Nutritional epidemiology, 3rd ed., Oxford University Press, New York, NY, 2013.
[35] S.I. Kirkpatrick, T. Baranowski, A.F. Subar, J.A. Tooze, E.A. Frongillo, Best practices for conducting and interpreting studies to validate self-report dietary assessment methods, J. Acad. Nutr. Diet. 119 (11) (2019) 1801–1816, https://doi.org/10.1016/j.jand.2019.06.010.
[36] L.S. Freedman, J.M. Commins, J.E. Moler, L. Arab, D.J. Baer, V. Kipnis, et al., Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake, Am. J. Epidemiol. 180 (2) (2014) 172–188, https://doi.org/10.1093/aje/kwu116.
[37] L. Zhang, A. Misir, H. Boshuizen, M. Ocké, A systematic review and meta-analysis of validation studies performed on dietary record apps, Adv. Nutr. 12 (6) (2021) 2321–2332, https://doi.org/10.1093/advances/nmab058.
[38] M. Ocké, C. Dinnissen, A. Stafleu, J. de Vries, C. van Rossum, Relative validity of MijnEetmeter: a food diary app for self-monitoring of dietary intake, Nutrients 13 (4) (2021) 1135, https://doi.org/10.3390/nu13041135.
[39] J.L. Lemacks, K. Adams, A. Lovetere, Dietary intake reporting accuracy of the bridge2U mobile application food log compared to control meal and dietary recall methods, Nutrients 11 (1) (2019) 199, https://doi.org/10.3390/nu11010199.