Traqq:

Development and evaluation of an ecological momentary dietary assessment smartphone app

Desiree A. Lucassen

Propositions

1. Ecological momentary dietary assessment is a valid method to assess dietary intake.
   (this thesis)

2. Integrating all respondents' wishes in dietary assessment tools decreases quality of reports.
   (this thesis)

3. Interdisciplinary collaboration is hindered by conflicting research interests.

4. Current (academic) funding structure hampers implementation of effective health-promoting interventions.

5. The 'quantified self' movement makes being healthy a math problem.

6. Chocolate is an essential part of a healthy diet.

Propositions belonging to the thesis, entitled:

Traqq
Development and evaluation of an ecological momentary dietary assessment smartphone app

Desiree A. Lucassen
Wageningen, 20 December 2022

Traqq

Development and evaluation of an ecological momentary dietary assessment smartphone app

Desiree A. Lucassen

Thesis committee

Promotor

Prof. Dr. Ir. Edith J.M. Feskens
Professor of Global Nutrition
Wageningen University and Research

Co-promotors

Dr. Elske M. Brouwer-Brolsma
Assistant Professor, Division of Human Nutrition and Health
Wageningen University and Research

Dr. Ir. Monica Mars
Associate Professor, Division of Human Nutrition and Health
Wageningen University and Research

Other members

Prof. Dr. Ir. Lisette C.P.G.M. de Groot, Wageningen University and Research
Prof. Dr. Ir. Joline Beulens, Amsterdam University Medical Centers
Prof. Dr. Dirk K.J. Heylen, University of Twente
Dr. Sharon Kirkpatrick, University of Waterloo

This thesis was conducted under auspices of the Graduate School VLAG (Advanced Studies in Food Technology, Agrobiotechnology, Nutrition and Health Studies).

Traqq

Development and evaluation of an ecological momentary dietary assessment smartphone app

Desiree A. Lucassen

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 20 December 2022
at 4 p.m. in the Aula.

Desiree A. Lucassen

Traqq: Development and evaluation of an ecological momentary dietary assessment smartphone app, 232 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2022)
With references, with summary in English

ISBN: 978-94-6447-398-8

DOI: 10.18174/576747

Table of content

Chapter 1

General introduction

Adapted from:

Brouwer-Brolsma, E. M., Lucassen, D. A., de Rijk, M. G., Slotegraaf, A., Perenboom, C., Borgonjen, K., Siebelink, E., Feskens, E. J. M. & de Vries, J. H. (2020, February). Dietary intake assessment: From traditional paper-pencil questionnaires to technology-based tools. In *International Symposium on Environmental Software Systems* (pp. 7-23). Springer, Cham.

ASSESSING DIETARY INTAKE: WHY?

Scurvy was a major cause of disability and mortality among long-distance sailors for decades [1]. In 1497, the Portuguese explorer Vasco da Gama led an expedition to India and reported that crew members with scurvy recovered days after eating fresh oranges. Yet, another 150 years passed before scurvy was finally acknowledged as being caused by malnutrition [1].

Fortunately, our understanding of how diet influences the human body evolved more rapidly during the past decades, and nutrient-related diseases considerably decreased. Accurate dietary assessment played an important role in these developments by generating quantitative information on the intake of foods, energy, and/or nutrients. The demand for quantitative information on dietary intake is still high, now more and more focusing on the exploration of diet-related determinants of today's challenges such as obesity and (age-related) non-communicable diseases (NCD) [2, 3]. For example, the world's population is faced with a large prevalence of the cardiometabolic risk factors overweight [4], hypertension [5], hypercholesterolemia [6], all in the range of 40%, and/or hyperglycaemia, respectively [7], ~10%, indicating the size of the current public health burden.

Consequently, many studies nowadays focus on the identification of modifiable dietary factors affecting the development of obesity and NCD risk. Nutritional epidemiologists, for instance, focus on potential associations between dairy consumption and body weight development or diabetes risk using data from large observational cohort studies [8, 9]. Dietary intake assessment is also an important component of dietary intervention studies. By modifying the consumption of a nutrient, food or diet in a controlled way and monitoring the potential impact on a selected health parameter, intervention studies are key to providing evidence on the actual causality of the role of a dietary component [10]. Dietary assessment is also performed by various national organisations to monitor the food and nutrient intake of the general population, which serves the formulation and evaluation of food policy [11]. Finally, a very important non-research-related application of dietary assessment is the healthcare setting, where it is used to prevent or treat diseases caused by malnutrition or disease-related malnutrition. Dietary assessment allows the healthcare professional to diagnose and provide feedback on nutritional status to the patient and to educate the patient to improve dietary habits.

ASSESSING DIETARY INTAKE: HOW?

Currently, self-report methods are the most commonly used dietary assessment methods in nutrition research. These can be generally divided in methods of recall and methods of real-time recording, and are described more in detail below.

METHODS OF RECALL

In nutrition research 24-hour recalls and food frequency questionnaires are the most commonly used methods of recall. Methods of recall rely on an individual to recall, i.e., remember, all foods and drinks consumed during a previous time period.

24-hour recall

The 24-hour recall (24hR) is an open-ended method to generate detailed information on all foods and drinks consumed during the previous 24 hours, i.e., *actual* intake, usually starting with breakfast on the previous day. On the individual level, data of two to three non-consecutive 24hRs can be used to gain insight into the *habitual* intake of commonly consumed foods. More than three days are needed to sufficiently capture the day-to-day variation of a variety of nutrients and foods that are episodically consumed, such as vitamin A, vitamin C, cholesterol, and fish [12]. The required observation period for individuals with a stable food pattern is usually shorter than the required observation period for individuals with a varied food pattern due to less day-to-day variation. Traditionally, 24hRs were mostly carried out by trained dietitians, either face-to-face or by telephone. In general, the interview can be completed in approximately 30 minutes, whereas food coding by the dietitian requires another 30-60 minutes. In research, dietitians perform the 24hRs according to the multiple-pass method; a validated five-step method developed to systematically conduct 24hRs and provides standardized questions and response options [13, 14]. Due to the workload related to this method, it is expensive to use face-to-face or phone-based 24hRs, which limits its use to small-scale studies. Fortunately, recent technological innovations led to the development of various self-administered 24hR tools. The development process of these 24hR tools is extensively described in literature [15-20] Obviously, it is of key importance that these new tools are just as accurate as the dietitian-guided recalls. So far, validation studies of self-administered recalls show promising results, but also clues for further improvement [21-28].

At the division of Human Nutrition and Health of Wageningen University, we also developed a self-administered web-based dietary 24hR tool, Compl-eat™ [28]. Contrary to the traditional method, this web-based tool is not guided by a research dietitian but self-completed by the respondent. Yet, comparable to the traditional 24hR, the web-based tool is based on the multiple-pass method, ensuring proper guidance while reporting the consumed foods [13, 14]. Identical to the traditional 24hR, respondents are requested to report their dietary intake of the previous day, starting in the morning after waking up till the next morning. At 6.00 AM, the respondent receives an invite via e-mail to complete the recall; the invite remains effective until midnight that same day. Compl-eat™ contains an extensive food list based on the Dutch food composition table [29], including commonly used synonyms as well as previously entered foods and standard recipes. This food list is flexible and can be easily modified in order to be tailored to specific research questions or updated to include new food items. Consumed amounts can be reported in food item-specific commonly used household measures, standard portion sizes, or in grams/litres. Compl-eat™ also comprises a recipe module, which facilitates the reporting of a complete dish by selecting or modifying a standard recipe from the food list. In addition, the respondent has the option to enter all ingredients of an original recipe in combination with the consumption amount of the meal. Yield and retention factors, i.e., retained weight and nutrients after cooking, are automatically taken into account. Respondents also have the possibility to include notes to clarify their input. After each eating occasion, respondents receive prompts to report on commonly omitted foods, such as sugar and/or milk in coffee/tea, oils and fats used in the preparation of dishes, snacks/candies and fruits) [28]. Generally, all web-based 24hRs are checked by research dietitians for completeness, unusual portion sizes, and notes entered by the respondent. Identified errors and notes are processed according to a standardised protocol, using standard portion sizes and recipes. Respondents are not contacted for clarifications. Examples of errors include the report of 125 cups of coffee instead of one cup of 125 g. Notes may relate to a product consumed, but could not be identified in the food list. The computation module of Compl-eat™ subsequently calculates food groups, energy and nutrient intakes. Respondents require on average 40-45 minutes to complete the web-based recall, which is 10-15 minutes more compared to the traditional 24hR method. However, the dietitians can process the recalls in 5-10 minutes, whereas approximately 60-90 minutes are needed to complete the interview and coding according to the traditional method.

Food frequency questionnaire

A Food Frequency Questionnaire (FFQ) is a fixed-food list, with or without portion size descriptions, inquiring about the consumption frequency of foods and beverages over the past month, past three months, or past year, i.e., focussing on *habitual* rather than *actual* intake. FFQs can be interviewer-based and self-administered. In general, an extensive FFQ that addresses macronutrients and the majority of micronutrients can be completed in approximately 45 minutes. FFQs are primarily designed to rank respondents according to their intakes and not to estimate absolute intakes. Nevertheless, in the case of nutrients or foods with large day-to-day variability, such as fish or alcohol, an FFQ may be more accurate than other methods also in terms of absolute intake. Important benefit of an FFQ is the efficiency of administration and processing. Several thousands of FFQs can easily be processed at once, and the output is relatively easy to convert to computer-ready data, making the FFQ a very practical method for use in large-scale studies, such as in cohort studies. However, in contrast to a recall or food record, the use of validated FFQs requires intensive preparation before it can be sent to the respondents. Food items that contribute most to the relevant energy and nutrient intakes in the target population need to be identified, and the FFQ needs to be validated (i.e., measures it what it is intended to measure), ideally using objective markers. However, due to the limited availability of such markers, as will be discussed later, and the associated costs, validation studies are often conducted using other self-reported dietary assessment methods as the reference method (e.g., 24hRs, food records) [30]. This is not optimal, as part of the measurement error is then correlated, which may lead to overestimation of the validity. All in all, the development of FFQs is a skilled task, time-consuming, and expensive. Fortunately, also this research area substantially developed in terms of automatization during the past years [31-35].

Together with partners we developed the Dutch FFQ-TOOL™; a data-driven web-based computer system developed to generate and process tailored FFQs (i.e., for nutrients of interest and population under study) by standardized, reproducible, relatively fast and flexible procedures [35]. The FFQ-TOOL™ has three main functionalities, i.e., 'selection of food items', 'question generation', and 'nutrient and food calculations'. The selection of food items is a semi-automated process. The FFQ-TOOL™ uses data from the Dutch National Food Consumption Survey (DNFCS) [36] to tailor the FFQ to the nutrients and (Dutch) population of interest, which is comparable to the procedure used to develop paper-based FFQs. Generally, researchers aim to cover about 80% of the absolute intake level and 80% of the between-person variability of each nutrient under study [37, 38]. The FFQ-TOOL™ indicates to which extent an item contributes to the total intake or the variation in intake for the nutrient(s) of interest for each aggregation level. Depending on the research question, the researcher subsequently selects the most suitable aggregation level and related food items. Thereafter, the selected food items are automatically translated to standard questions. Once the FFQ is completed, individual food(group) intake, and energy and nutrient intake is computed through the computation module of the FFQ-TOOL™, which is facilitated by attached (Dutch) food composition tables.

In addition, we also developed the Eetscore™; a self-administered web-based screener to assess habitual diet quality during the previous month. In contrast to above-described method and tool, the Eetscore™ is a relatively short FFQ specifically developed to fulfil the demand for a shorter and less burdensome questionnaire. The Eetscore™ its primary aim is not to obtain quantitative food or nutrient intakes, but to assess the individual's diet quality. First a dietary index was developed that indicated diet quality, and in the second step a short FFQ was designed that covered the items of this index. The most recent version of the Eetscore™ is based on the Dutch Healthy Diet 2015-index (DHD15-index) [39]. The DHD15-index includes fifteen components, including vegetables, fruit, whole-grain products, legumes, nuts, dairy, fish, tea, fats and oils, coffee, red meat, processed meat, sweetened beverages and fruit juices, alcohol, and salt. For each component a respondent can score from 0 to 10. In addition to these 15 components of the DHD15-Index, the Eetscore™ comprises one additional component, i.e., the unhealthy choices component. This component was added based on the guidelines of the Netherlands Nutrition Centre aiming to get insight in dietary intake beyond the Dutch dietary guidelines [40]. The Eetscore™ can be completed in approximately 10-15 minutes and therefore respondent and researcher burden as well as the associated costs are relatively low. The Eetscore™ also has an unique addition, the possibility to provide immediate personal dietary advice after submission of the questionnaire. However, the Eetscore™ can also be administered without the advice module, and solely used as screener.

METHOD OF REAL-TIME MONITORING

The food record is the most commonly used method of real-time monitoring in nutrition research. Individuals record all foods and drinks consumed throughout one or more days.

Food record

Food records are open-ended and generate detailed information, i.e., amount and type, on all foods and drinks consumed during the recording period. Similar to the 24hR, a one-day food record provides information on *actual* food and nutrient intake; two to three non-consecutive day food records provide information on the *habitual* intake of commonly consumed foods on the individual level. More days are needed to cover the nutrients and foods that are less commonly consumed, similar as for the 24HR [12]. The completion time of a one-day food record is approximately 30 minutes distributed over the day. In theory, multiple (i.e., 7 day) weighed food records are the most accurate self-reported dietary assessment method; the so-called "gold standard" [41]. In case of weighed food records, the respondent is instructed to weigh all foods and drinks consumed, ideally using scales with an accuracy up to 1 gram. Following a demonstration on the weighing and reporting of consumed foods (i.e., food type such as white bread vs. whole-wheat bread, food brands, recipe details) the respondent receives a simple notebook. A disadvantage of dietary records is that they are prone to reactivity bias, very intrusive for respondents, and also time-consuming and labour-intensive for dietitians due to the food coding of the notebooks. Weighed food records can be very useful in dietary studies, but not feasible for use in large-scale studies. The non-weighed food record largely follows the same procedure, but is less intrusive as food quantity is estimated, using e.g., standard portion-sizes and household measures. Obviously, this procedure requires more of the dietitian in terms of the interpretation of the portion size estimates and is thus less precise compared to the weighed food record. Fortunately, also for the food record, technological inventions have led to promising innovations, including the use of mobile devices. Whereas the more basic apps still collect dietary intake data through descriptive text [42], other apps are also exploring the potential of before and after photography, which provides additional information on consumed portion sizes and potentially undocumented foods [43].

TRUE VS. MEASURED DIET: SOURCES OF MEASUREMENT ERROR

Studies exploring diet–disease associations often show inconsistent findings [44, 45]; varying from null associations, beneficial associations to adverse associations. Inconsistencies may relate to various factors, including study population (e.g., healthy vs. health-compromised population), variation in the exposure (e.g., population with high intakes of a certain food or nutrient vs. population with a low intake) or outcome (e.g., low vs. high prevalence of a certain disease) under study, the covariates considered (e.g. inadequate vs. satisfactory correction for confounders) or the applied statistical approach (which may affect statistical power to detect potential associations). Methodological issues related to the assessment of the exposure (dietary factors) are also commonly discussed. Indeed, it is indisputable that the above-described methods have their limitations and introduce measurement error. This error can be both "intake-related", reflecting the correlation between the error and true intake, and "person-specific", indicating errors related to the respondent's personal characteristics [46]. Besides, errors can be systematic or random [46]. To be more specific, a shared factor for all methods of recall, such as 24hR and FFQ, is its sensitivity to memory-related bias (Figure 1). Other shared sources of errors for these methods as well as the non-weighed food record include the inaccurate estimation of portion sizes and errors in food composition tables. An additional source of measurement error for FFQs is the large supply of newly available foods, which cannot be fully reflected in a fixed-food list. Additionally, reporting obtained through food records may be influenced by the fact that respondents are made aware of their habits while recording/collecting. To limit this specific source of error it is therefore important to emphasize that respondents should not change their food intake at the time of recording/collecting (i.e., reactivity bias).

As all dietary assessment methods contain measurement error, it is important to understand the impact of these errors to correctly interpret nutritional outcomes. Random errors (e.g., day-to-day variation) decrease the precision of the assessment, they do not influence mean dietary intake but increase the variation in the population and may weaken (i.e., attenuate) the strength of the diet-disease association (although this also depends on the measurement error in the covariates) [12]. In contrast, systematic errors, such as systematic under- or overreporting, decrease the accuracy of the method resulting in an inaccurate mean intake. Yet, in contrast to random errors, the diet-disease association remains unaffected [12]. Validation studies are of extreme importance to gain insight into present measurement errors, especially after development of a new dietary assessment method and/or tool. The selection of an appropriate reference method is of special importance, where independent measures, such as recovery markers (discussed later), are preferred. A variety of statistical methods can then be used to (partly) correct for measurement errors. Measurement error models can be used to calculate, for instance, validity coefficients and attenuation factors. Validity coefficients can be used to assess the loss of statistical power to detect a diet-disease association and to assess how well a method is able to rank participants according to their unknown true dietary intake. Validity coefficients of <0.20 are classified as poor, 0.20-0.49 as acceptable, and ≥0.50 as good [47]. Attenuation factors can be used to give insight in the extent to which diet-disease relations are attenuated by measurement error, e.g., using self-report food intake data instead of true intake. In addition, attenuation factors can also be used to correct attenuated diet-disease associations assessed with that specific method. An attenuation factor closer to 1 means less attenuation, with 1 representing no attenuation at all [48].

INNOVATIONS

It may be clear that each dietary assessment tool has its strengths and its weaknesses. Up to 15-20 years ago, above presented methods were completely paper-pencil based, which shifted more and more towards web-based throughout the past decade. The current pace of technological development is very valuable to improve our methods, i.e., reduce sources of error, increase user-friendliness, and decrease workload of dietitians and/or researchers. Moreover, due to the absence of a dietitian/researcher, data obtained through web-based tools are for instance expected to be less biased by socially desirable answers. Web-based tools are also assumed to be less burdensome for the respondents as they can complete the dietary assessment at a time that is convenient for them.

DIETARY ASSESSMENT APPS

In addition to the web-based tools, many smartphone-based dietary assessment tools (apps) are being developed, enabled by the universal adoption of smartphones in the population. To illustrate, in 2019, 83% of Dutch citizens over 12 years used a smartphone with mobile internet outside the home [49], showing that apps provide the opportunity to enable real-time data collection at any location, at any time. Furthermore, research shows that respondents prefer to use dietary assessment apps over conventional dietary assessment methods [50]. Respondents are often familiar which such apps due to the high availability and popularity of commercial diet tracking apps (e.g., MyFitnessPal, Lifesum, Lose it!). However, publicly available apps have several limitations that make them unfit for research. The most important one being their unreliable food composition databases. The compilation procedures of these databases are non-standardized, non-documented, and users are often allowed to add foods to the database (i.e., user-compiled database), without any form of quality control [51], resulting in inaccurate nutrition calculations [52, 53]. Another disadvantage of publicly available apps is their purpose of weight management, i.e., they provide feedback on recorded intake which could result in adaptations of a user’s diet. However, in research, you often do not want participants to alter their diet as this could hamper estimates of habitual dietary intake and negatively affect results in e.g., cohort studies on diet-disease relationships [53, 54].

Still, apps have undeniable potential for use in dietary assessment and many research groups seized this opportunity which resulted in a variety of different dietary assessment apps. These apps can be roughly divided into text-based apps, i.e., traditional intake recording via fixed food list, and image-based apps, i.e., intake recording via images/pictures. However, to date, image-based apps still require manual coding of food items, as they remain unable to correctly identify all food items and/or consumed amounts from images [50, 55]. Hence, text-based apps are currently predominantly used in nutrition research. There is only limited availability of fully automated and validated dietary assessment apps without feedback [42, 56, 57]. For an overview of currently available and validated apps that can be used for nutrition research see Table 1.

Table 1. Overview of fully automated and validated dietary assessment research apps, without feedback, for adults

App	Country	Method	Population	Reference method	First author, year [ref]
e-12HR	Spain	Recall[1]	University students and staff (>18y)	FFQ, food record	Béjar, 2018 [58] Béjar, 2019 [59]
e-CA	Switzerland	Food record	Adults (20-60y)	24hRs	Bucher Della Torre, 2017 [60]
e-DIA	Australia	Food record	University students (19-24y)	24hRs	Rangan, 2015 [61] Rangan, 2016 [62]
Eat and Track (EaT)	Australia	Food record	Young adults (18-30y)	24hRs	Wellard-Cole, 2019 [63]
PIQNIQ	USA	Food record[2]	Adults (18-65y)	24hRs	Blanchard, 2021 [64]
Research Food Diary (RFD)	Australia	Food record	University students and staff (>18y)	24hRs	Ambrosini, 2018 [65]

[1] e-12HR is only able to capture a selection of food groups.

[2] PIQNIQ can also be used as an image-assisted recall, i.e., respondent takes pictures of food intake throughout the day and reports intake the following day in the app. Yet, this is not an official 24hR but more a combination of a record and a recall.

Strikingly, most validated research apps are based on the food record method. Although the food record method is not new, the reporting of food intake via an app is. Therefore, it is important that new dietary assessment apps are extensively validated, preferably against independent markers, before being applied in research. However, even after extensive validation, method-related errors remain such as reactivity bias [66, 67]. This is not the case for recalls, which emphasizes the need for validated recall-based apps [48]. Yet, no recall-based apps exist except for one: the electronic 12-hour dietary recall app “e-12HR”. Although the e-12HR is well-validated, it is only able to capture the intake of certain food groups and not an entire diet [58, 59].

ECOLOGICAL MOMENTARY DIETARY ASSESSMENT

The 12hR is a variant of the 24hR. Although 12hRs are already less reliant on memory, 12-hours is still a long time period and memory-related bias remains. So why not go a step further? Apps have the major advantage of enabling (near) real-time assessment of dietary intake data [66, 68-70]. This method of real-time assessment is often used in behavioural and social sciences, and referred to as ecological momentary assessment (EMA); repeated real-time assessment of individual's behaviour in their natural 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 [70].

First attempts to integrate EMA and dietary assessment approaches have resulted in two EMDA (i.e., ecological momentary dietary assessment) approaches: event-contingent EMDA and signal-contingent EMDA [71, 72]. Event-contingent EMDA is basically a mobile food record, where an eating occasion triggers the recording of what is being consumed. In contrast, with signal-contingent EMDA, individuals are prompted to report recent dietary intake. Although individuals are prompted to report dietary intake, signal-contingent EMDA does not allow for real-time dietary intake assessment. Yet, it does allow near real-time recording by using short recall intervals (e.g., past hour). The sampling scheme is determined by the researcher and can occur either at fixed times or at random times within a fixed time period; current signal-contingent EMDA methods all use a brief survey focussing on only part of an individual's diet such as intake of specific foods or food groups (e.g., snacks, alcoholic beverages, fruits and vegetables) [71, 72].

The Dutch 'Snackimpuls app' is currently the only signal-contingent EMDA method that is directly linked to a national food consumption database and is able to provide data on intake of energy, carbohydrates, fat, and protein. However, as the name already reveals, the Snackimpuls app only assesses intake of snacks (i.e., non-main meals) [73]. The other existing methods only assess frequency of consumption of certain foods [71]. Still, signal-contingent EMDA is very promising due to the possibility of (unannounced) sampling, short recall intervals (i.e., low reliance on memory), and limited reporting burden [72]. To optimize this method, some aspects warrant attention. First, the sampling scheme needs to be optimized to ensure coverage of a full day (i.e., *actual* intake) or three or more full days (i.e., *habitual* intake). Second, the surveying method to ensure individuals can record their entire dietary intake. Therefore, it is also important to ensure direct linkage to a trustworthy food composition database. Finally, extensive validation is essential to ensure the new method assesses what it is intended to assess, preferably against independent biomarkers.

Biomarkers

Biological markers for dietary intake are more objective than self-reported dietary intake methods, as they are not affected by memory, social desirability and/or errors in food composition tables [74, 75]. There are several well-validated nutritional biomarkers that can be used to assess the validity and accuracy of (new) dietary assessment methods. In general, two classes of biomarkers can be discerned: recovery biomarkers and concentration biomarkers.

Recovery biomarkers are based on the recovery of certain food compounds in urine, i.e., doubly labelled water (for metabolic rate and total energy expenditure), nitrogen (for total protein), potassium, and sodium. Recovery biomarkers are founded on the concept of the metabolic balance between intake and excretion over a fixed time period and are able to estimate absolute nutrient intakes [74]. Thus, excretion levels are highly correlated with intake (>0.80) [74, 76]. Although urinary sodium can be used to validate self-reported sodium intake, it is often not included in validation studies. Self-assessment of sodium intake is extremely difficult as intake is not only determined by sodium present in foods. Sodium is also added while cooking and while eating in very small amounts, making estimation of added sodium (or salt) nearly impossible [76].

Concentration biomarkers are based on the correlation of certain food compounds or metabolites in e.g., the circulation and intake of corresponding foods or nutrients. Examples are carotenoids and polyunsaturated fatty acids (PUFA). These blood concentration levels show often lower correlations with intake (<0.60) than recovery markers, and can, therefore, not be used to estimate absolute levels of intake [74, 76]. Still, serum carotenoids and n-3 PUFA can be used to assess the relative validity (i.e., ranking) for the intake of fruit/vegetables and fish, respectively [77-79].
Unfortunately, currently not many biomarkers are available that can be used to validate self-reported dietary intake. However, this field is developing rapidly and metabolomic techniques now provide a unique opportunity to measure up to thousands of metabolites at once, hopefully providing valuable information on the food metabolome using a variety of body tissues in the near future [76, 80].

RATIONALE FOR THIS THESIS

Due to innovations in technology the methods to assess dietary intake have improved regarding cost- and time-effectiveness, labour-intensiveness and respondent and researcher burden. However, novel tools still share various methodological problems with the traditional self-report methods. Some of these issues can be traced back to the original dietary assessment methodology, important examples being memory-related bias, social desirability bias, and reactivity bias (Figure 1). Although the implementation of technology decreased these sources of error, major issues remain. Therefore, it is important to not only innovate current methodologies but also explore new approaches.

Memory-related bias	Respondents struggle to accurately remember their food intake, i.e., they forget to report certain foods
Reactivity bias	Respondents may alter their food intake due to the awareness that they are being observed
Social desirability bias	Respondents reports higher intakes of healthy foods and a lower intake of unhealthy foods, or even omits unhealthy foods

Figure 1. Most common sources of bias in dietary assessment

The overarching objective of this thesis is, therefore, to develop a flexible smartphone-based EMDA tool that can be tailored to specific research objectives and to further explore the use of signal-contingent EMDA to collect detailed dietary intake data. A 2-hour reporting period would minimize the reliance on memory compared to the traditional 24hRs. The 2-hour reporting window was considered an optimum, avoiding 'I did not consume anything' responses when using shorter time-windows, and avoiding the higher memory-related and reporting burden using 24-hour time windows. Several steps were taken to reach this objective:

1. To evaluate the accuracy of different portion size estimation aids (Chapter 2)

2. To develop a smartphone-based dietary assessment tool 'Traqq®' (Chapter 3)

3. Designing an extensive evaluation study that allows assessment of the validity, usability and perceived burden of the 2hR methodology (Chapter 4)

4. Validation of repeated 2hRs during one day to assess *actual* intake on specific days (Chapter 5)

5. Evaluation of random 2hRs over a longer time period to assess *habitual* intake (Chapter 6)

6. Exploration of technological innovations with the potential to further improve current dietary assessment efforts (Chapter 7)

The first two objectives relate to the development of the methodology and the smartphone-based tool to apply the methodology. Portion size estimation is a vital part of dietary assessment as it allows quantification of reported food intake. The current dietary assessment tools make use of textual portion size descriptions (i.e., household measures, standard portion sizes) and estimation in grams. In **Chapter 2**, a pilot study was conducted to explore whether the use of image-based portion size estimation would result in more accurate portion size assessments. **Chapter 3** describes the protocol that was used for the development of the smartphone-based tool. The objectives 3 to 5 relate to the evaluation and validation of the new EMA-based 2hR methodology. **Chapter 4** describes the design of the extensive study that was conducted to thoroughly evaluate the new methodology. Results of this evaluation study are described chapters 5 and 6. **Chapter 5** describes the results of the validation of repeated 2hRs during one day for assessing *actual* intake, while **Chapter 6** describes the results of the evaluation of random 2hRs over a longer time period for assessing *habitual* intake. After developing and evaluating both the tool and the new methodology, **Chapter 7** explores technological innovations that could be applied to further improve current dietary assessment efforts, including the new tool. This chapter also discusses opportunities to further develop the tool(s) to not only measure dietary behaviours, but also positively influence them. Finally, in **Chapter 8** the main findings of the different chapters are summarized and discussed in the context of existing literature. This chapter also includes lessons learned in combination with recommendations for future research and implications for research and practice.

REFERENCES

1. Carpenter, K.J., The discovery of vitamin C. Ann Nutr Metab, 2012. 61(3): p. 259-64.
2. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. Lancet, 2017. 390(10113): p. 2627-2642.
3. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet, 2016. 388(10053): p. 1545-1602.
4. WHO. Global health observatory data: obesity, situations and trends. secondary global health observatory data: obesity, situations and trends. 2016; Available from: https://www.who.int/gho/ncd/risk_factors/overweight_obesity/obesity_adults/en/.
5. WHO. Global health observatory data: raised blood pressure, situation and trends. secondary global health observatory data: raised blood pressure, situation and trends. 2008; Available from: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en/.
6. WHO. Global health observatory data: raised cholesterol, situation and trends secondary global health observatory data: raised cholesterol, situation and trends. 2008; Available from: http://www.who.int/gho/ncd/risk_factors/cholesterol_text/en/
7. WHO. Global health observatory data: raised fasting blood glucose, situation and trends. secondary global health observatory data: raised fasting blood glucose, situation and trends. 2014; Available from: http://www.who.int/gho/ncd/risk_factors/blood_glucose_text/en/
8. Brouwer-Brolsma, E.M., et al., Dairy shows different associations with abdominal and BMI-defined overweight: Cross-sectional analyses exploring a variety of dairy products. Nutr Metab Cardiovasc Dis, 2018. 28(5): p. 451-460.
9. Brouwer-Brolsma, E.M., et al., Dairy product consumption is associated with pre-diabetes and newly diagnosed type 2 diabetes in the Lifelines Cohort Study. Br J Nutr, 2018. 119(4): p. 442-455.
10. Staudacher, H.M., et al., The challenges of control groups, placebos and blinding in clinical trials of dietary interventions. Proc Nutr Soc, 2017. 76(3): p. 203-212.
11. Huybrechts, I., et al., Global comparison of national individual food consumption surveys as a basis for health research and integration in national health surveillance programmes. Proc Nutr Soc, 2017. 76(4): p. 549-567.
12. Willett, W.C., Nutritional Epidemiology. 2013, New York: Oxford University Press Inc

13. Moshfegh, A.J., et al., The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. Am J Clin Nutr, 2008. 88(2): p. 324-32.
14. Blanton, C.A., et al., The USDA Automated Multiple-Pass Method accurately estimates group total energy and nutrient intake. J Nutr, 2006. 136(10): p. 2594-9.
15. Carter, M.C., et al., Development of a UK Online 24-h Dietary Assessment Tool: myfood24. Nutrients, 2015. 7(6): p. 4016-32.
16. Carter, M.C., et al., Development of a New Branded UK Food Composition Database for an Online Dietary Assessment Tool. Nutrients, 2016. 8(8).
17. Evans, K., et al., Development and evaluation of a concise food list for use in a web-based 24-h dietary recall tool. J Nutr Sci, 2017. 6: p. e46.
18. Jacques, S., et al., Development of a Web-Based 24-h Dietary Recall for a French-Canadian Population. Nutrients, 2016. 8(11): p. 724.
19. Simpson, E., et al., Iterative Development of an Online Dietary Recall Tool: INTAKE24. Nutrients, 2017. 9(2): p. 118.
20. Timon, C.M., et al., The development, validation, and user evaluation of Foodbook24: a web-based dietary assessment tool developed for the Irish adult population. Journal of medical Internet research, 2017. 19(5): p. e6407.
21. Touvier, M., et al., Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. Br J Nutr, 2011. 105(7): p. 1055-64.
22. Subar, A.F., et al., The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. J Acad Nutr Diet, 2012. 112(8): p. 1134-7.
23. Foster, E., et al., Validity and reliability of an online self-report 24-h dietary recall method (Intake24): a doubly labelled water study and repeated-measures analysis. J Nutr Sci, 2019. 8: p. e29.
24. Arab, L., et al., Eight self-administered 24-hour dietary recalls using the Internet are feasible in African Americans and Whites: the energetics study. J Am Diet Assoc, 2010. 110(6): p. 857-64.
25. Greenwood, D.C., et al., Validation of the Oxford WebQ Online 24-Hour Dietary Questionnaire Using Biomarkers. Am J Epidemiol, 2019. 188(10): p. 1858-1867.
26. Wark, P.A., et al., Validity of an online 24-h recall tool (myfood24) for dietary assessment in population studies: comparison with biomarkers and standard interviews. BMC Med, 2018. 16(1): p. 136.
27. Timon, C.M., et al., Comparison of a Web-Based 24-h Dietary Recall Tool (Foodbook24) to an Interviewer-Led 24-h Dietary Recall. Nutrients, 2017. 9(5).

28. Meijboom, S., et al., Evaluation of dietary intake 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, 2017. 6: p. e49.
29. RIVM. NEVO-online (2016/5.0). 2016 [cited 2018 August 15].
30. Cade, J., et al., Development, validation and utilisation of food-frequency questionnaires - a review. Public Health Nutr, 2002. 5(4): p. 567-87.
31. Kristal, A.R., et al., Evaluation of web-based, self-administered, graphical food frequency questionnaire. J Acad Nutr Diet, 2014. 114(4): p. 613-21.
32. Fallaize, R., et al., Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. J Med Internet Res, 2014. 16(8): p. e190.
33. Labonte, M.E., et al., Validity and reproducibility of a web-based, self-administered food frequency questionnaire. Eur J Clin Nutr, 2012. 66(2): p. 166-73.
34. Wise, A. and N.M. Birrell, Design and analysis of food frequency questionnaires--review and novel method. Int J Food Sci Nutr, 2002. 53(3): p. 273-9.
35. Molag, M., Towards Transparent Development of Food Frequency Questionnaires. Scientific basis of the Dutch FFQ-TOOLTM: a computer system to generate, apply and process FFQs, in Division of Human Nutrition and Health. 2010, Wageningen University: Wageningen.
36. van Rossum, C.T.M., et al., The diet of the Dutch : Results of the first two years of the Dutch National Food Consumption Survey 2012-2016. 2016, Dutch Institute for Public Health and the Environment: Bilthoven.
37. Eussen, S.J., et al., A national FFQ for the Netherlands (the FFQ-NL1.0): development and compatibility with existing Dutch FFQs. Public Health Nutr, 2018. 21(12): p. 2221-2229.
38. Sluik, D., et al., A national FFQ for the Netherlands (the FFQ-NL 1.0): validation of a comprehensive FFQ for adults. Br J Nutr, 2016. 116(5): p. 913-23.
39. Looman, M., et al., Development and evaluation of the Dutch Healthy Diet index 2015. Public Health Nutr, 2017. 20(13): p. 2289-2299.
40. The Netherlands Nutrition Centre. Wat staat er niet in de Schijf van Vijf? . [cited 2019 17-10-2019]; Available from: https://www.voedingscentrum.nl/nl/gezond-eten-met-de-schijf-van-vijf/omgaan-met-producten-buiten-de-schijf-van-vijf.aspx.
41. Cameron, M.E. and W.A. van Staveren, Manual on Methodology for Food Consumption Studies. Methods for Data Collection at an Individual Level. 1988, Oxford: Oxford University Press.
42. Eldridge, A.L., et al., Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients, 2018. 11(1).

43. Boushey, C.J., et al., New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proc Nutr Soc, 2017. 76(3): p. 283-294.
44. Gijsbers, L., et al., Consumption of dairy foods and diabetes incidence: a dose-response meta-analysis of observational studies. Am J Clin Nutr, 2016. 103(4): p. 1111-24.
45. Schwingshackl, L., et al., Food Groups and Risk of Hypertension: A Systematic Review and Dose-Response Meta-Analysis of Prospective Studies. Adv Nutr, 2017. 8(6): p. 793-803.
46. Kipnis, V., et al., Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol, 2003. 158(1): p. 14-21; discussion 22-6.
47. Lombard, M.J., et al., Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutrition journal, 2015. 14(1): p. 1-11.
48. Kipnis, V., et al., Bias in dietary-report instruments and its implications for nutritional epidemiology. Public health nutrition, 2002. 5(6a): p. 915-923.
49. CBS. Internettoegang en internetactiviteiten; persoonskenmerken. 2022 15 October 2021 [cited 2022 22 July]; Available from: https://opendata.cbs.nl/#/CBS/nl/dataset/84888NED/table?ts=1658476779452.
50. Sharp, D.B. and M. Allman-Farinelli, Feasibility and validity of mobile phones to assess dietary intake. Nutrition, 2014. 30(11-12): p. 1257-66.
51. Maringer, M., et al., User-documented food consumption data from publicly available apps: an analysis of opportunities and challenges for nutrition research. Nutr J, 2018. 17(1): p. 59.
52. Griffiths, C., L. Harnack, and M.A. Pereira, Assessment of the accuracy of nutrient calculations of five popular nutrition tracking applications. Public health nutrition, 2018. 21(8): p. 1495-1502.
53. Khazen, W., et al., Rethinking the Use of Mobile Apps for Dietary Assessment in Medical Research. J Med Internet Res, 2020. 22(6): p. e15619.
54. Livingstone, M.B.E. and A.E. Black, Markers of the validity of reported energy intake. The Journal of nutrition, 2003. 133(3): p. 895S-920S.
55. Rollo, M.E., et al., What are they really eating? A review on new approaches to dietary intake assessment and validation. Current nutrition reports, 2016. 5(4): p. 307-314.
56. Zhang, L., et al., A systematic review and meta-analysis of validation studies performed on dietary record apps. Advances in Nutrition, 2021. 12(6): p. 2321-2332.
57. König, L.M., et al., Characteristics of smartphone-based dietary assessment tools: A systematic review. Health Psychology Review, 2021: p. 1-25.
58. Béjar, L.M., Ó.A. Reyes, and M.D. García-Perea, Electronic 12-hour dietary recall (e-12HR): comparison of a mobile phone app for dietary intake assessment with a food

frequency questionnaire and four dietary records. JMIR mHealth and uHealth, 2018. 6(6): p. e10409.
59. Béjar, L.M., et al., Relative validity of a method based on a smartphone app (Electronic 12-hour Dietary Recall) to estimate habitual dietary intake in adults. JMIR mHealth and uHealth, 2019. 7(4): p. e11531.
60. Bucher Della Torre, S., et al., Development and evaluation of e-CA, an electronic mobile-based food record. Nutrients, 2017. 9(1): p. 76.
61. Rangan, A.M., 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 and uHealth, 2015. 3(4): p. e4613.
62. Rangan, A.M., et al., Electronic Dietary Intake Assessment (e-DIA): relative validity of a mobile phone application to measure intake of food groups. British Journal of Nutrition, 2016. 115(12): p. 2219-2226.
63. Wellard-Cole, L., 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, 2019. 11(3): p. 621.
64. Blanchard, C.M., et al., Evaluation of PIQNIQ, a novel mobile application for capturing dietary intake. The Journal of nutrition, 2021. 151(5): p. 1347-1356.
65. Ambrosini, G.L., et al., Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutrition journal, 2018. 17(1): p. 1-10.
66. Thompson, F.E. and A.F. Subar, Dietary assessment methodology. Nutrition in the Prevention and Treatment of Disease, 2017: p. 5-48.
67. Subar, A.F., et al., Addressing current criticism regarding the value of self-report dietary data. The Journal of nutrition, 2015. 145(12): p. 2639-2645.
68. Shim, J.-S., K. Oh, and H.C. Kim, Dietary assessment methods in epidemiologic studies. Epidemiology and health, 2014. 36.
69. Stone, A., et al., The science of real-time data capture: Self-reports in health research. 2007: Oxford University Press.
70. Shiffman, S., A.A. Stone, and M.R. Hufford, Ecological momentary assessment. Annu Rev Clin Psychol, 2008. 4(1): p. 1-32.
71. Schembre, S.M., et al., Mobile ecological momentary diet assessment methods for behavioral research: systematic review. JMIR mHealth and uHealth, 2018. 6(11): p. e11170.
72. Maugeri, A. and M. Barchitta, A Systematic Review of Ecological Momentary Assessment of Diet: Implications and Perspectives for Nutritional Epidemiology. Nutrients, 2019. 11(11).

73. Wouters, S., et al., Assessing energy intake in daily life: Signal-contingent smartphone application versus event-contingent paper and pencil estimated diet diary. Psychologica Belgica, 2016. 56(4): p. 357.
74. Jenab, M., et al., Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet, 2009. 125(5-6): p. 507-25.
75. Gao, Q., et al., A scheme for a flexible classification of dietary and health biomarkers. Genes & nutrition, 2017. 12(1): p. 1-15.
76. Mattes, R.D., et al., Valuing the Diversity of Research Methods to Advance Nutrition Science. Adv Nutr, 2022.
77. Brevik, A., et al., Six carotenoids in plasma used to assess recommended intake of fruits and vegetables in a controlled feeding study. Eur J Clin Nutr, 2004. 58(8): p. 1166-73.
78. Al-Delaimy, W.K., et al., Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). Eur J Clin Nutr, 2005. 59(12): p. 1387-96.
79. Saadatian-Elahi, M., et al., Plasma phospholipid fatty acid profiles and their association with food intakes: results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. Am J Clin Nutr, 2009. 89(1): p. 331-46.
80. Brouwer-Brolsma, E.M., et al., Combining traditional dietary assessment methods with novel metabolomics techniques: present efforts by the Food Biomarker Alliance. Proc Nutr Soc, 2017. 76(4): p. 619-627.

Chapter 2

The accuracy of portion size estimation using food images and textual descriptions of portion sizes: an evaluation study

Desiree A. Lucassen
Romy F. Willemsen
Anouk Geelen
Corine Perenboom
Edith J.M. Feskens

Journal of Human Nutrition and Dietetics 2021; 34(6): 945-952

ABSTRACT

BACKGROUND

Inaccurate self-report of portion sizes is a major cause of measurement error in dietary assessment. To reduce this error, different portion size estimation aids (PSEAs) have been developed, including food images (image based, IB-PSE) and textual descriptions of portion sizes (text-based, TB-PSE). We assessed the accuracy of portion size estimation by IB-PSE and TB-PSE.

METHODS

True intake of one lunch was ascertained in forty participants. Self-reported portion sizes were assessed after 2 and 24 hours by means of TB-PSE and IB-PSE, in random order. Wilcoxon's tests were used to compare mean true intakes to reported intakes. Moreover, proportions of reported portion sizes within 10% and 25% of true intake were assessed. An adapted Bland-Altman approach was used to assess agreement between true and reported portion sizes. Analyses were conducted for all foods and drinks combined and for predetermined food types.

RESULTS

No significant differences were observed between reported portion sizes at 2 and 24 hours after lunch. Combining median relative errors of all foods items resulted in an overall 0% error rate for TB-PSE and 6% error rate for IB-PSE. Comparing reported portion sizes within 10% (31% vs. 13%) and 25% (50% vs. 35%) of the true intake showed a better performance for TB-PSE compared to IP-PSE, respectively. Bland-Altman plots indicated a higher agreement between reported and true intake for TB-PSE compared to IB-PSE.

CONCLUSIONS

Although the use of TB-PSE still results in measurement error, our results suggest a more accurate dietary intake assessment with TB-PSE than IB-PSE.

INTRODUCTION

Accurate dietary assessment is essential in nutrition research. Although dietary intake is still often assessed using paper-pencil tools, i.e. food frequency questionnaires (FFQs), food records (FRs) and 24-hour recalls (24hRs), dietary assessment techniques have advanced rapidly in recent years. The last decade numerous valuable computer-based and web-based tools, mostly based on 24hRs and FFQs, have been developed [1-3]. More recently also different smartphone applications (i.e. apps), mostly based on FRs, have been developed to collect real-time dietary intake data [3, 4]. Important benefits of these new tools include that they are assumed to lower burden on both participant and researcher compared to traditional techniques [3-6].

A fundamental aspect of accurate dietary assessment is portion size estimation [6-8]. However, assessment of portion sizes is challenging and a major cause of error in dietary assessment [6, 9-11]. Difficulties occur while reporting previously consumed foods as well as when judging displayed foods [4, 11, 12]. The accuracy of portion size estimation is affected by various factors, including type of food and serving size [6, 10, 13]. Generally, single-unit foods (e.g. sliced bread, fruits) are more likely to be reported correctly compared to liquids or amorphous foods (e.g. pasta, lettuce) [4, 12, 14]. Another issue in portion size estimation is that large portions tend to be underestimated and small portions tend to be overestimated, which is also known as the ‘flat-slope phenomenon’ [11]. In addition, foods consumed in small portions (e.g. spreads) are likely to be estimated more accurately than large portions of foods [13].

Portion size estimation aids (PSEAs) (e.g. images, referent objects, portion size suggestions) have been suggested to result in more accurate portion sizes estimates [15-17]. However, research indicates that these PSEAs still result in measurement error and that further optimization of PSEAs is needed [17], especially with respect to PSEAs that may be implemented in web-based and smartphone-based dietary assessment tools. The most commonly used PSEAs in web-based and smartphone-based tools are portion size suggestions (i.e. standard portion sizes and household measures), food images, and free entry of weight in grams [1]. As individuals fail to recognize the metric quantities of portion sizes, estimations in grams are usually inaccurate [18]. For this reason, participants tend to prefer the use of household measures rather than estimation in grams [17, 18]. Yet, inconsistent or vague descriptions of household measures may still result in measurement error, especially among individuals that are not frequently involved in meal preparation [18, 19]. Therefore, clear descriptions of the portion sizes are crucial [20].

To facilitate the estimation of portion sizes, several dietary assessment tools have included food images as visual aids, where individuals are requested to select the most comparable image with respect to the portion size consumed or displayed (i.e. image-based portion size assessment or IB-PSE). Previous research indicates that IB-PSE is particularly influenced by three main elements, namely perception, conceptualization and memory [13]. Despite these elements of potential error, IB-PSE is suggested to be a useful aid to estimate portion sizes [14, 21-24]. However, there is only limited evidence on the reliability of IB-PSE in real-life situations [14, 19]. Up to now, the reliability of IB-PSE has mainly been examined by exposing participants to foods and food images simultaneously while focussing on perception and not conceptualization and memory [22-24]. More specifically, the majority of previous research only compared PSEAs to weighed portion sizes as a reference technique [12, 19, 21-24]. To the best of our knowledge, none of the previous studies examined the accuracy of portion size estimation using a combination of textual descriptions of household measures (e.g. spoons, cups, glasses), standard portion sizes (e.g. small, medium, large) and estimation in grams (i.e. for the purpose of this study referred to as text-based portion size estimation or TB-PSE) and IB-PSE.

Therefore, the current study aimed to compare the accuracy of TB-PSE and IB-PSE. As we hypothesize that accuracy varies over different food types, accuracy of both PSEAs was examined for all foods and drinks combined and for specific food types. In addition, to gain a first insight in the effect of memory on the accuracy of the PSEAs, the portion sizes were reported after either 2 hours or 24 hours.

METHODS

PARTICIPANTS

Participants were recruited through a convenience sampling method using a database of research volunteers of the division of Human Nutrition and Health of Wageningen University and Research (WUR), social media accounts of the division (i.e. Facebook and Twitter), and through posters. Eligible participants were Dutch speaking, not visually impaired, not participating in another dietary intervention study, not an employee of the division, and not having any formal training in the field of nutrition. In total, 40 participants aged 20-70 years old were included in this study that was conducted during a 2-week period in February 2018. Participants were stratified by sex and age to ensure equal distribution of these characteristics and randomly assigned to two groups. Participants were informed that the study focused on different digital methods to assess food intake. The true study purpose was not disclosed until the end of the study. Written informed consent was obtained from all participants.

OVERALL STUDY DESIGN

Participants were invited for one lunch at the study centre as part of the cross-over study and asked to complete two dietary questionnaires on a tablet or computer; 2 and 24 hours after lunch. The first group reported their food intake 2 hours after lunch by means of TB-PSE and 24 hours after lunch by means of IB-PSE. The second group reported their intake with the two PSEAs in the opposite order. As previous studies suggest that the potential difficulty to accurately estimate portion size depends on the type of food, we offered a variety of commonly consumed food types in the Netherlands [7, 12-14] (Table 1).

Table 1. Food items offered, by food type.

<u>Amorphous</u>
- Cheese
- Crunchy muesli
- Fruit salad
- Scrambled eggs
- Yogurt
<u>Liquids</u>
- Milk
- Orange juice
- Water
<u>Single-units</u>
- Bread slices
- Bread rolls
<u>Spreads</u>
- Jam
- Margarine

Each participant was provided with pre-weighed, ad libitum amounts of the food items. Each item was offered in a container without indication of the content. To minimize the effect of tableware on portion size estimation [25], the participants received a variety of tableware. After lunch, plate waste was weighed to assess true intake of each food item. Weights were taken with 'Sartorius Signum 1' calibrated weighing scales. True intake was calculated by the following formula:

True intake (g) = Pre-weighed food item (g) – Plate waste food item (g)

PORTION SIZE ASSESSMENT

For the purpose of this study, a TB-PSE and IB-PSE questionnaire was developed in Qualtrics (Qualtrics, Provo, UT, USA). The question formulation and portion size estimation within the TB-PSE questionnaire were based on Compl-eat™; a self-administered web-based dietary 24hR-tool developed by WUR [20]. Portion sizes described in Compl-eat™ are a combination of estimation in grams/millilitres, standard portion sizes and household measures, which are based on the 'Food portion sizes and coding instructions' [26]. The question formulation within the IB-PSE questionnaire was also based on Compl-eat™, thus ensuring that observed differences were solely due to the different PSEAs and not due to differences in question formulation. For the IB-PSE questionnaire, the portion size images from the Automated Self-Administered 24-hour dietary recall (ASA24) picture book, developed by the National Cancer Institute, Bethesda, MD [27], were used. This picture book contains 3 to 8 portion size images per food item. To the best of our knowledge, this is the only freely available picture book portraying food images with known amounts (g) for research purposes [28]. Questionnaires started with questions whether or not a type of food was consumed, which was followed by questions on the amount of food consumed by means of one of the PSEAs. An example question from each questionnaire can be found in Supplement A.

ADDITIONAL MEASUREMENTS

On the study day, participants completed a short questionnaire about basic characteristics (i.e. age, sex, educational level). In addition, weight and height were measured to calculate participants' BMI (kg/m2). Participants were characterized in three educational levels (low: primary or lower education, intermediate: secondary or higher vocational education, high: college or university) and four age groups (18-28, 29-45, 46-55, 56-70 years).

STATISTICAL ANALYSIS

Normally distributed data is displayed as means (M) and standard deviations (SDs) in case of continuous variables, or frequencies in case of categorical variables; non-normally distributed data as medians and interquartile ranges (IQRs). Significant differences between true and reported intake, and between 2 and 24 hours, were assessed for each PSEA. To allow comparison between PSEAs across different food types, relative differences were calculated. As previous research indicated that accuracy of portion size estimation varies over food types, all analyses were conducted for all foods and drinks combined and for predetermined food types individually (i.e. “all foods excluding liquids”, “amorphous foods”, “liquids”, “single-units”, “spreads”; Table 1). As there are no guidelines on the acceptable level of accuracy [7, 14, 29], the proportion of the reported intake that fell within 10% and 25% of true intake were assessed, which is in line with comparable studies in this research area [14]. Proportions within 10% of true intake will be deemed acceptably accurate, whereas proportions within 25% of true intake will be used to get further insight in the levels of accuracy [30]. To determine agreement between reported and true intake for both PSEAs, Bland-Altman plots with 95% limits of agreement (LMO) were plotted. Usually the Bland-Altman method is applied for assessing agreement between two imperfect measures. Since true intake was assessed an adapted Bland-Altman method was used to plot the differences between reported and true intake against true intake [14, 31]. However, when true intake increased, the absolute error increased. Therefore, we plotted the log-transformed ratio of reported and true intake against log-transformed true intake. Middle line indicates the mean and the upper and lower lines indicate borders based on mean ± 1.96 SD. Since the variables were not normally distributed, Wilcoxon signed rank test was used to test within group and the Wilcoxon rank sum test was used to test for between group differences. All analyses were conducted with SAS software, version 9.4 (SAS Institute Inc., Cary, NC, USA). Statistical significance was set at $p < 0.05$.

RESULTS

A total of 40 participants took part in this study. Participants had a mean ± SD age of 46.9 ± 19.2 years (range 20.7–69.4 years), BMI 24.9 ± 3.8 kg/m2, 47.5% was men and the majority of the population was highly educated (62.5%). Participant characteristics did not significantly differ between group 1 (2hR: TB; 24hR: IB) and group 2 (2hR: IB; 24hR: TB) (Table 2). Furthermore, no significant differences were observed between reported at 2 and at 24 hours after lunch, for each PSEA. Therefore, the results are only shown per PSEA and are not subdivided per time point.

Table 2. Characteristics of the participants

	Total (*n* 40)			Group 1† (*n* = 20)			Group 2‡ (*n* = 20)		
	Mean	SD	%	Mean	SD	%	Mean	SD	%
Men			47.5			50.0			45.0
Age (years)	46.9	19.2		48.7	19.8		45.0	18.9	
BMI (kg/m^2)	24.9	3.8		25.9	4.1		24.0	3.3	
Educational level			0.0			0.0			0.0
Low			37.5			35.0			40.0
Intermediate			62.5			65.0			60.0
High									

† Group 1: 2hR = TB-PSE; 24hR = IB-PSE

‡ Group 2: 2hR = IB-PSE; 24hR = TB-PSE

No significant differences were found between groups.

Median true intake for “all foods and drinks combined” was 94 g (IQR: 128 g), while median reported intake was 75 g (IQR: 120 g) for TB-PSE and 88 g (IQR: 164 g) for IB-PSE. Comparing the true intake with the reported intake, as assessed with TB-PSE, pointed towards significant differences for “all foods excluding liquids”, “amorphous foods”, “liquids” and “spreads” (Table 3). For IB-PSE, significant differences with the true intake were observed for “all foods and drinks combined”, “liquids”, “single-units” and “spreads”. For “all foods and drinks combined” the median relative difference was 0% (IQR: 44%) as assessed by TB-PSE, and 6% (IQR: 115%) as assessed by IB-PSE (Table 3).

Significantly higher relative errors were shown for IB-PSE than for TB-PSE for “all foods and drinks combined”, “all foods excluding liquids”, “amorphous foods” and “liquids”. For “all foods and drinks combined” the proportion of reported intakes within 10% of true intake was 31% for TB-PSE and 13% for IB-PSE, the proportion within 25% of true intake was 50% for TB-PSE and 35% for IB-PSE. For TB-PSE, the lowest proportion within 10% and 25% of true intake was observed for “spreads”, whereas for IB-PSE, the lowest proportion was observed for "liquids". The highest proportion of reported intake that fell within 10% and 25% of true intake was, for both PSEAs, observed for the food type "single-units" (Table 3).

The log-transformed Bland-Altman plot of "all foods and drinks combined" showed a higher level of agreement for TB-PSE (M: 0.04; LOA: -1.11-1.03) than for IB-PSE, as shown by more widely scattered estimates and wider limits of agreement for IB-PSE (Supplement B). Excluding liquids did not substantially alter these findings; agreement for TB-PSE (M: -0.10; LOA: -1.22-1.00) remained higher compared to IB-PSE (M:0.03; LOA: -1.37-1.43). The same trend was observed for the other food types (Supplement B). The highest level of agreement was observed for "single-units" (TB-PSE M: -0.02; LOA: -0.30-0.25 vs. IB-PSE M: -0.09; LOA: -0.84-0.66), whereas the lowest level of agreement was observed for "amorphous foods" (TB-PSE M: -0.13; LOA: -1.43-1.15 vs. IB-PSE M: 0.17; LOA: -1.38-1.71).

Table 3. Median true intake, median reported intake for both PSEAs, and reported intakes within 10% and 25% of true intake for both PSEAs for all foods and per food type.

			TB-PSE					IB-PSE				
	Obs (n)	Median true intake (g, IQR)	Median rep. intake (g, IQR)	Median diff. (g, IQR)†	Median rel. diff. (%, IQR)‡	<10% of true (%)	<25% of true (%)	Median rep. intake (g, IQR)	Median diff. (g, IQR)†	Median rel. diff. (%, IQR)‡	<10% of true (%)	<25% of true (%)
All food items	326	94 (128)	75 (120)	0 (25)	0 (44)	31	50	88 (164)	3 (74)***	6 (115)***	13	35
All excl. liquids	263	66 (86)	50 (101)	0 (22)***	0 (44)	32	49	61 (95)	-1 (41)	-5 (74)**	15	41
Amorphous	150	84 (124)	66 (107)	-7 (45)**	-12 (56)	12	25	104 (119)	3 (70)	4 (109)***	13	31
Liquids	63	194 (125)	220 (150)	29 (80)***	15 (39)	30	56	355 (356)	211 (373)***	118 (122)***	5	11
Single-units	59	100 (50)	100 (50)	0 (0)	0 (0)	95	95	72 (61)	-7 (25)*	-14 (18)	29	80
Spreads	54	17 (15)	15 (14)	-3 (14)*	-23 (63)	9	24	14 (12)	-3 (10)*	-22 (63)	9	30

All food items., all foods and drinks combined; obs., observations; rep., reported; diff., difference; rel., relative; excl., excluding.

† Calculated as reported intake minus true intake. Thus, positive differences represent overestimations and negative differences represent underestimations. Significant differences between reported and true intake was assessed with a Wilcoxon signed-rank test. Significant differences are indicated by * for *P*<0.05, ** for *P*<0.01, *** for *P*<0.0001.

‡ Relative differences (%) = (reported intake (g) –true intake (g)) / true intake (g) * 100. Significant differences between intake reported with TB-PSE and IB-PSE were assessed with a Wilcoxon signed-rank test. Significant differences are indicated by * for *P*<0.05, ** for *P*<0.01, *** for *P*<0.0001 in the IB-PSE “Mean rel. diff. column”.

DISCUSSION

In this study, the reported intake and its estimation error for "all foods and drinks combined" using IB-PSE significantly differed from true intake while no statistically significant difference was observed between the reported intake and its estimation error from true intake using TB-PSE. However, as indicated by the proportion of reported intakes within 10% and 25% of true intake, being 31% and 50% using TB-PSE compared to 13% and 35% using IB-PSE, meaning that for both PSEA's only the minority of estimations lies within the acceptable range, further improvements to increase the accuracy of portion size estimation are needed.

Before discussing our findings, the strengths and limitations of our study will be discussed. First, despite the fact that participants consumed their lunch in a controlled setting, we strived to mimic a real-life situation. Specifically, in contrast to most other studies, participants could choose from a selection of food items and actually consumed the selected items [19, 24]. Furthermore, participants had the opportunity to choose between different sizes of tableware [25] and had ad libitum access to the foods provided [32]. Moreover, all products were served in bowls, jugs and plates without indication of content. Second, as the accuracy of two PSEAs was assessed separately, accuracy of both methods could be studied independently. Moreover, due to the study's cross-over design the accuracy of both PSEAs was assessed in each participant. Third, to our knowledge, this is the first study comparing the two PSEAs, while keeping all other factors in the questionnaire identical. Finally, to avoid extra focus on portion sizes, participants were not informed on the goal of the study and did not see the weighing of the foods. A limitation of our study is that we used the ASA24 picture book in a Dutch population. The ASA24 is the only freely available photo database for research with known portion size weights. However, the ASA24 photographs are based on the 5th and 95th percentile of intake per product in the US and as such tailored for usage in the US [14, 33, 34]. It is known that portion sizes in the US are larger than in the Netherlands [35, 36]. To illustrate, the glasses in the study of Donders-Engelen, Van der Heijden [26] range between 100 g and 220 g whereas the glasses in ASA24 range between 177 g and 473 g. As ASA24 does not contain pictures of the smallest portion sizes consumed in the Netherlands, this may explain the overestimated intakes by IB-PSE estimates in our study (e.g. 118% for "liquids"). However, we have to note that the portion size database that currently is being used in the Netherlands dates from 2003. It is known that plate sizes have increased in the past decades [36], which on its turn may have led to an underestimation of TB-PSEs.

A more general limitation of the ASA24 food images is the usage of cutlery as reference, which is meant to help participants estimate the real-life size of a portion. However, as cutlery can vary in size, it might not be the best reference and as such explain the more scattered points observed in the Bland-Altman plot of IB-PSE compared to TB-PSE. Finally, in view of generalisability it needs to be mentioned that our participants were relatively old and highly educated. However, several previous studies concluded that age and education level did not affect the participants ability to estimate portion sizes [19, 22, 23, 37]. In addition, we only tested a limited number of food items, and as such our findings are only applicable to these tested food items.

As hypothesized, the accuracy of reported intake with both PSEAs varied between the different food types. Both PSEAs overestimated the median reported intake of “liquids” whereas the intake of “all foods excluding liquids” and “spreads” were (slightly) underestimated. In addition, for TB-PSE, the reported median intake of “amorphous foods” was underestimated, while for IB-PSE the intake was overestimated. Previous research showed both under- and overestimations of portion size estimations [7, 14]. Moreover, the accuracy of food intake estimates varied depending on the food types [12, 13, 38]. Both PSEAs showed the highest estimation errors for “liquids, which is not in line with similar studies showing the highest estimation errors for “amorphous foods” [12-14, 37]. In contrast to previous studies, which mostly provided liquids in containers that were identical to containers portrayed on the images, we aimed to resemble the real-life situation and therefore studied commonly-used PSEA descriptions and used glasses that did not necessarily match with the glasses on the images. As conceptualization plays a major role in the accurateness of portion size estimation [13], it is easier to estimate portion sizes when the portion sizes are similar to the portions portrayed on the images [23, 39] or the textual descriptions [18, 20]. For instance, the description “lemonade glass” lacks detail and can easily result in misclassification. In agreement with our study, Hernandez, Wilder [7] also studied the intake of liquids in containers that were not identical to the containers on the images and also observed the highest estimation errors for liquids, which underlines the influence of conceptualization.

As illustrated by small errors for “single-units” and “spreads” and larg(er) errors for “amorphous foods” and “liquids” for both PSEAs, our findings clearly indicate that foods consumed in small or defined units are more accurately estimated than foods consumed in larger amounts. These findings are in line with previous studies [23, 37, 39]. Generally, the accuracy for the food types “amorphous foods”, “liquids” and “single-units” was higher for TB-PSE than for IB-PSE estimates, except for “spreads” which were more accurately estimated with IB-PSE. The latter may relate to the fact that textual description of the size of spoons and spread on bread is open to interpretation, whereas a picture may provide a better impression of the portion size estimate [13]. Moreover, the fact that we used images of spoons, instead of images of spread on bread, to estimate the amount of “spreads” consumed, may have resulted in more accurate estimates for this food type [12]. The size of the bread might influence the perception of the portion size and thereby lead to errors in estimations [21].

We found no significant differences in accuracy between reporting after 2 hours and 24 hours for each of the PSEAs. Based on this, we concluded that memory did not influence the accuracy of portion size estimations within this timeframe. Therefore, only the combined results per PSEA were used for further analysis. However, after dividing the participants per PSEA over the two time points, the sample size per group was very small (i.e. ~10 participants) and therefore we had less power to detect significant differences. Previous research has shown that errors increase after 1-2 hours, compared to immediate estimations [24]. However, our first time point was after two hours and in line with our results, De Keyzer, Huybrechts [21] found no increase in estimation errors after 1-2 days compared to after 4 days [21]. To truly understand the effect of memory on accuracy of portion size estimation more research is needed with a larger sample size.

Due to lack of consensus on the minimal required level of accuracy for PSEAs no strong conclusion can be drawn on that matter. However, the accuracy of the reported intake by TB-PSE was higher than by IB-PSE for all food types except for “spreads”, which was higher with IB-PSE. Overall, TB-PSE provided more accurate portion size estimations than IB-PSE. As discussed, these findings are different from previous studies [14, 21-24]. However, in contrast to these studies we incorporated all elements that influence IB-PSE (i.e. perception, conceptualization, memory), instead of focusing on one or two of these elements [22-24], in an attempt to mimic a real-life situation. Therefore, our findings in combination with previous studies may indicate that IB-PSE is a useful PSEA, but only when judging displayed foods and not for retrospective portion size estimation.

TB-PSE and IB-PSE were selected due to their applicability for implementation in web-based and smartphone-base dietary assessment tools. However, there are other PSEAs which would be applicable for implementation in web-based or smartphone-based dietary assessment tools (e.g. remote food photography method, body-worn monitors) [8, 40]. These innovative tools also have a range of drawbacks, for instance, it is known that they are unable to detect all aspects of the food consumed (e.g. no difference detected between spinach vs. spinach a la crème) [41]. Furthermore, individuals might feel uncomfortable wearing the device, especially long-term, and it is difficult to guarantee the privacy of bystanders [40]. Moreover, even though these devises have been proven to be up to 90% accurate [40], such devices are expensive and therefore not suited for large-scale studies. Selecting a PSE-tool needs to be considered carefully while taking into account study design, methods and target group [8]. Therefore, even though there are new, more innovative PSE-tools being developed, it is still valuable to further improve both TB-PSE and IB-PSE. These PSEAs are easy to implement in web-based and smartphone-based tools, relatively inexpensive, well-known and therefore easy to use with limited training.

To conclude, in our study TB-PSE is shown to be more accurate than IB-PSE. Country-specific pictures with a clear reference are needed to improve the accuracy of IB-PSE. Next to this, we can conclude that TB-PSE seems to be an accurate PSEA for “single-units”, as 95% of the reported intake fell within 10% of true intake. However, for the other food types, only 32% or less of the reported intakes fell within 10% of truth. Therefore, in line with Bucher, Rollo [42], we conclude that the accuracy of portion size estimations with TB-PSE needs to be improved further and therefore standardized terminology is needed to avoid ambiguity with regard to textual descriptions of portion sizes. Finally, the use of a combination of PSEAs might be valuable to increase accuracy of portion size estimation.

REFERENCES

1. Timon, C.M., et al., A review of the design and validation of web- and computer-based 24-h dietary recall tools. Nutr Res Rev, 2016. 29(2): p. 268-280.
2. Ngo, J., et al., A review of the use of information and communication technologies for dietary assessment. Br J Nutr, 2009. 101 Suppl 2: p. S102-12.
3. Shim, J.S., K. Oh, and H.C. Kim, Dietary assessment methods in epidemiologic studies. Epidemiol Health, 2014. 36: p. e2014009.
4. Thompson, F.E. and A.F. Subar, Dietary Assessment Methodology, in Nutrition in the Prevention and Treatment of Disease. 2017, Elsevier Inc. p. 5-48.
5. Naska, A., A. Lagiou, and P. Lagiou, Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Res, 2017. 6: p. 926.
6. Walton, J., Dietary Assessment Methodology for Nutritional Assessment. Topics in Clinical Nutrition, 2015. 30(1): p. 33-46.
7. Hernandez, T., et al., Portion size estimation and expectation of accuracy. Journal of Food Composition and Analysis, 2006. 19: p. S14-S21.
8. Sharma, V. and R. Chadha, Effectiveness of food portion size estimation aids for diet assessment: A systematic review. International Journal of Food Sciences and Nutrition, 2017. 2(5): p. 106-112.
9. Subar, A.F., et al., Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. J Nutr, 2015. 145(12): p. 2639-45.
10. Vuckovic, N., et al., A qualitative study of participants' experiences with dietary assessment. J Am Diet Assoc, 2000. 100(9): p. 1023-8.
11. Young, L.R. and M.S. Nestle, Portion sizes in dietary assessment: issues and policy implications. Nutr Rev, 1995. 53(6): p. 149-58.
12. Subar, A.F., et al., Assessment of the accuracy of portion size reports using computer-based food photographs aids in the development of an automated self-administered 24-hour recall. J Am Diet Assoc, 2010. 110(1): p. 55-64.
13. Nelson, M., M. Atkinson, and S. Darbyshire, Food photography. I: The perception of food portion size from photographs. Br J Nutr, 1994. 72(5): p. 649-63.
14. Kirkpatrick, S.I., et al., 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, 2016. 146(12): p. 2567-2573.
15. Byrd-Bredbenner, C. and J. Schwartz, The effect of practical portion size measurement aids on the accuracy of portion size estimates made by young adults. J Hum Nutr Diet, 2004. 17(4): p. 351-7.

16. Faggiano, F., et al., Validation of a method for the estimation of food portion size. Epidemiology, 1992. 3(4): p. 379-82.
17. Faulkner, G.P., et al., An evaluation of portion size estimation aids: precision, ease of use and likelihood of future use. Public Health Nutr, 2016. 19(13): p. 2377-87.
18. Spence, M., et al., A qualitative study of psychological, social and behavioral barriers to appropriate food portion size control. Int J Behav Nutr Phys Act, 2013. 10(92): p. 92.
19. Biltoft-Jensen, A., et al., Accuracy of food photographs for quantifying food servings in a lunch meal setting among Danish children and adults. J Hum Nutr Diet, 2018. 31(1): p. 131-140.
20. Meijboom, S., et al., Evaluation of dietary intake assessed by the Dutch self-administered web-based dietary 24-h recall tool (Compl-eat™) against interviewer-administered telephone-based 24-h recalls. Journal of Nutritional Science, 2017. 6.
21. De Keyzer, W., et al., Food photographs in nutritional surveillance: errors in portion size estimation using drawings of bread and photographs of margarine and beverages consumption. Br J Nutr, 2011. 105(7): p. 1073-83.
22. Naska, A., et al., Evaluation of a digital food photography atlas used as portion size measurement aid in dietary surveys in Greece. Public Health Nutr, 2016. 19(13): p. 2369-76.
23. Ovaskainen, M.L., et al., Accuracy in the estimation of food servings against the portions in food photographs. Eur J Clin Nutr, 2008. 62(5): p. 674-81.
24. Szenczi-Cseh, J., Z. Horvath, and A. Ambrus, Validation of a food quantification picture book and portion sizes estimation applying perception and memory methods. Int J Food Sci Nutr, 2017. 68(8): p. 960-972.
25. Van Ittersum, K. and B. Wansink, Plate size and color suggestibility: the Delboeuf Illusion's bias on serving and eating behavior. Journal of Consumer Research, 2011. 39(2): p. 215-228.
26. Donders-Engelen, M.R., L.J.M. Van der Heijden, and K.F.A.M. Hulshof, Maten, Gewichten en Codenummers 2003. Food Portion Sizes and Coding Instructions. 2003, Zeist: Wageningen University: Division of Human Nutrition and TNO Nutrition,.
27. Subar, A.F., et al., The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. J Acad Nutr Diet, 2012. 112(8): p. 1134-7.
28. ASA24. ASA24: Automated Self-Administrated 24-hour dietary assessment Tool. 2018 06-02-2018 30-04-2018]; Available from: https://epi.grants.cancer.gov/asa24/respondent/citations.html.
29. European Food Safety Authority, Use of the EFSA Comprehensive European Food Consumption Database in Exposure Assessment. EFSA Journal, 2011. 9(3).

30. Lombard, M.J., et al., Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J, 2015. 14: p. 40.
31. Krouwer, J.S., Why Bland–Altman plots should use X, not (Y+ X)/2 when X is a reference method. Statistics in Medicine, 2008. 27(5): p. 778-780.
32. Clancy, A.K., et al., Food Composition Database Format and Structure: A User Focused Approach. PLoS One, 2015. 10(11): p. e0142137.
33. Kirkpatrick, S.I., et al., 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, 2014. 100(1): p. 233-40.
34. Kirkpatrick, S.I., et al., Validity of the National Cancer Institute's Automated Self-Administered 24-hour Recall (ASA24): results of a feeding study. The FASEB Journal, 2013. 27(1 Supplement): p. 230.4-230.4.
35. Rozin, P., et al., The ecology of eating: smaller portion sizes in France Than in the United States help explain the French paradox. Psychol Sci, 2003. 14(5): p. 450-4.
36. Steenhuis, I.H., F.H. Leeuwis, and W.M. Vermeer, Small, medium, large or supersize: trends in food portion sizes in The Netherlands. Public Health Nutr, 2010. 13(6): p. 852-7.
37. Bucher, T., et al., The international food unit: a new measurement aid that can improve portion size estimation. Int J Behav Nutr Phys Act, 2017. 14(1): p. 124.
38. Robson, P.J. and M.B.E. Livingstone, An evaluation of food photographs as a tool for quantifying food and nutrient intakes. Public Health Nutrition, 2000. 3(2): p. 183-192.
39. Vereecken, C.A., et al., Young adolescents' nutrition assessment on computer (YANA-C). Eur J Clin Nutr, 2005. 59(5): p. 658-67.
40. McClung, H.L., et al., Dietary Intake and Physical Activity Assessment: Current Tools, Techniques, and Technologies for Use in Adult Populations. Am J Prev Med, 2018. 55(4): p. e93-e104.
41. Zhu, F., et al., The Use of Mobile Devices in Aiding Dietary Assessment and Evaluation. IEEE J Sel Top Signal Process, 2010. 4(4): p. 756-766.
42. Bucher, T., et al., Position paper on the need for portion-size education and a standardised unit of measurement. Health Promot J Austr, 2017. 28(3): p. 260-263.

SUPPLEMENTARY MATERIAL

Supplement A

Example question of TB-PSE questionnaire

Scrambled eggs: How much did you eat (first select a serving unit and then insert the number of servings)?

Grams (1 gram) ▾

Grams (1 gram)
Pieces (50 grams)
Tablespoon (15 grams)

Insert the number of consumed servings for **scrambled eggs**:

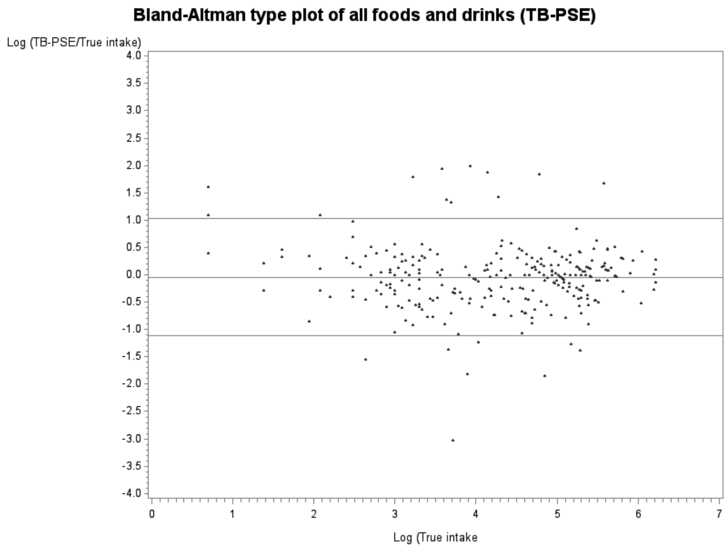
Example question of IB-PSE questionnaire

Scrambled eggs: How much did you eat (first select a serving unit and then insert the number of servings)? (4 images)

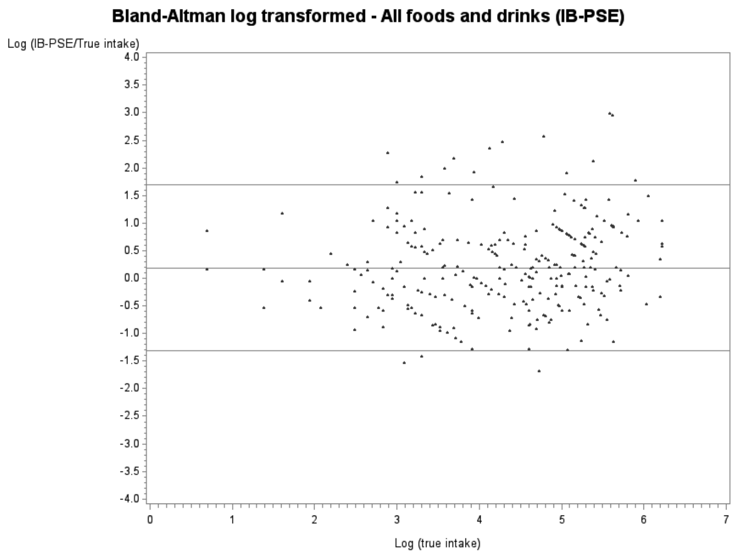
Insert the number of consumed servings for **scrambled eggs**:

SUPPLEMENT B

All foods and drinks

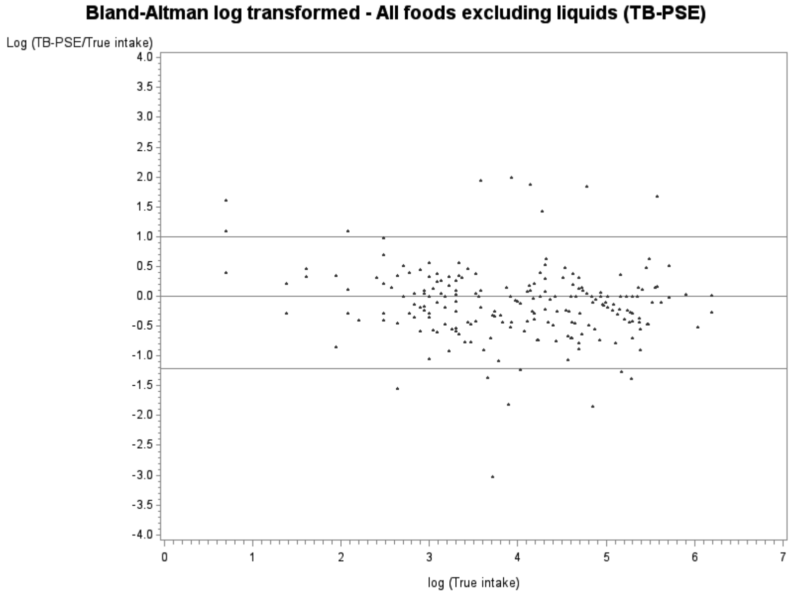


sFigure 1. Bland-Altman plot of log transformed proportion of TB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for all foods and drinks.

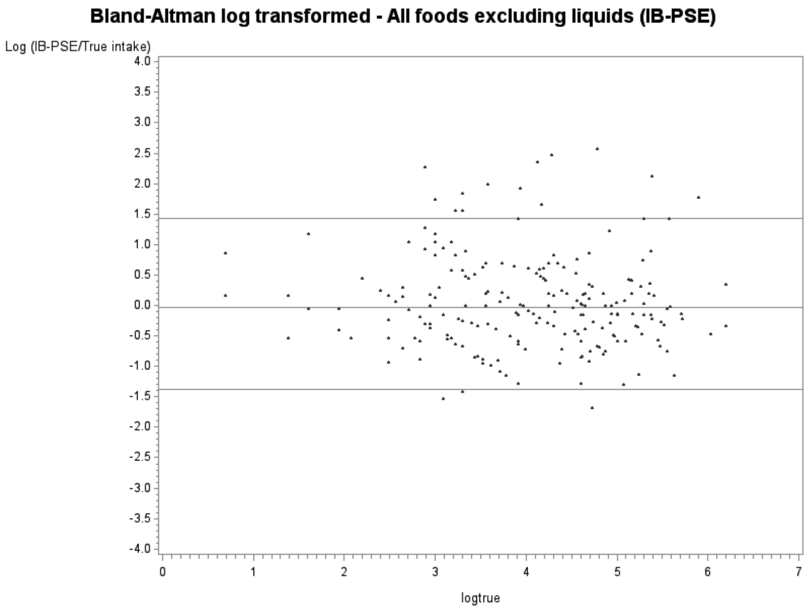


sFigure 2. Bland-Altman plots of log transformed proportion of IB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for all foods and drinks.

All foods excluding liquids

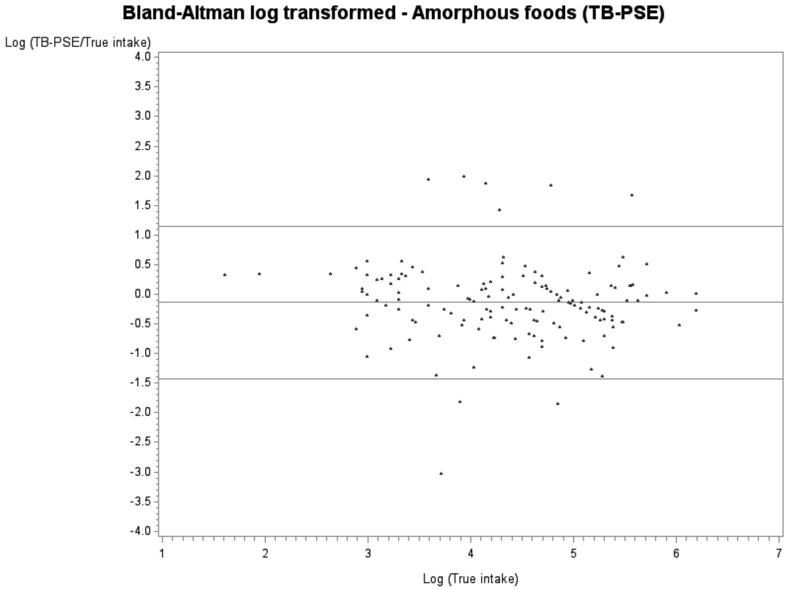


sFigure 3. Bland-Altman plot of log transformed proportion of TB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for all foods excluding liquids.

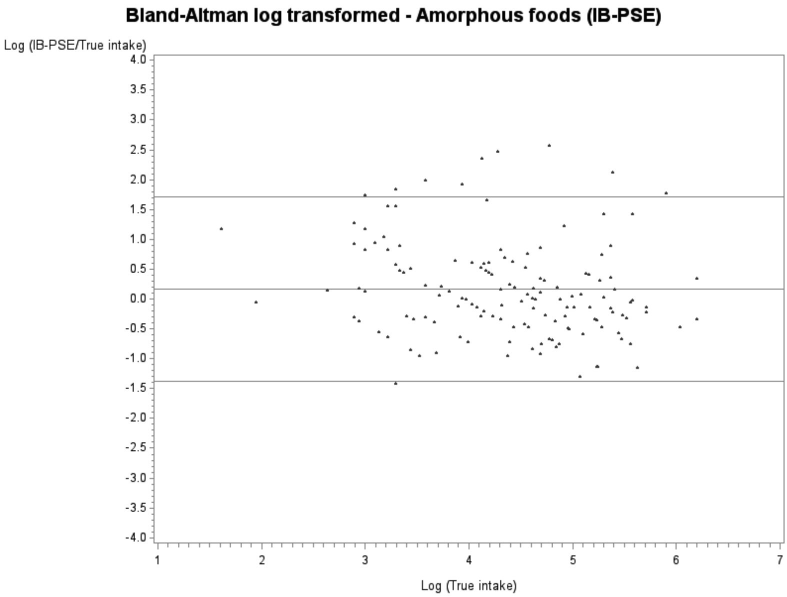


sFigure 4. Bland-Altman plots of log transformed proportion of IB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for all foods excluding liquids.

Amorphous foods

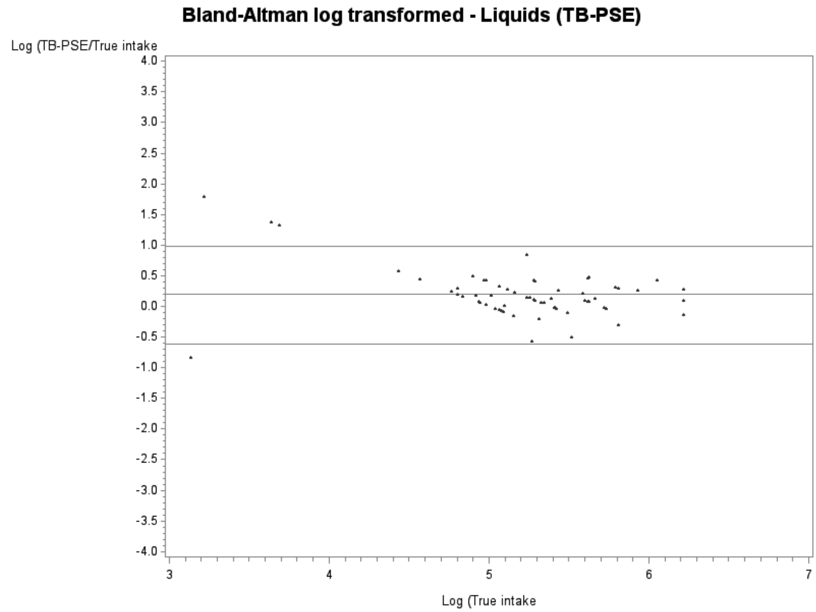


sFigure 5. Bland-Altman plot of log transformed proportion of TB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for amorphous foods.

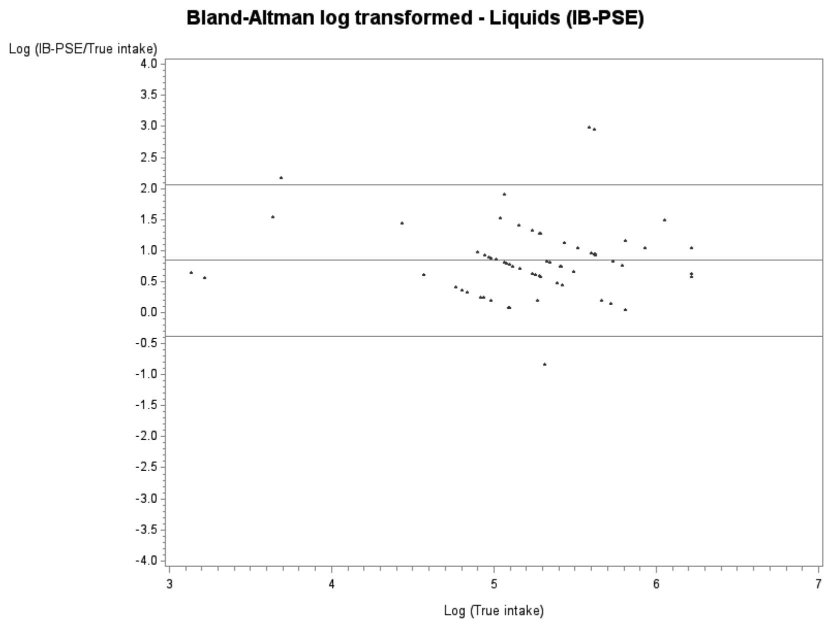


sFigure 6. Bland-Altman plots of log transformed proportion of IB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for amorphous foods.

Liquids

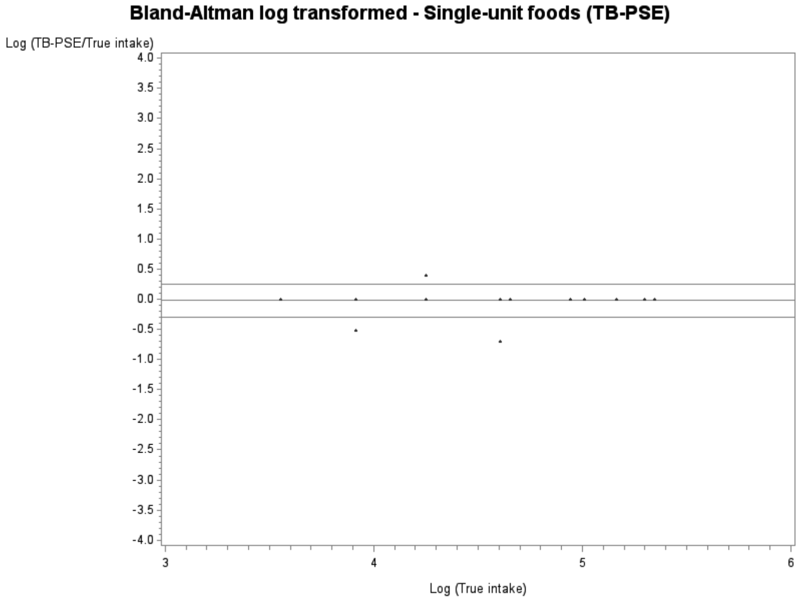


sFigure 7. Bland-Altman plot of log transformed proportion of TB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for liquids.

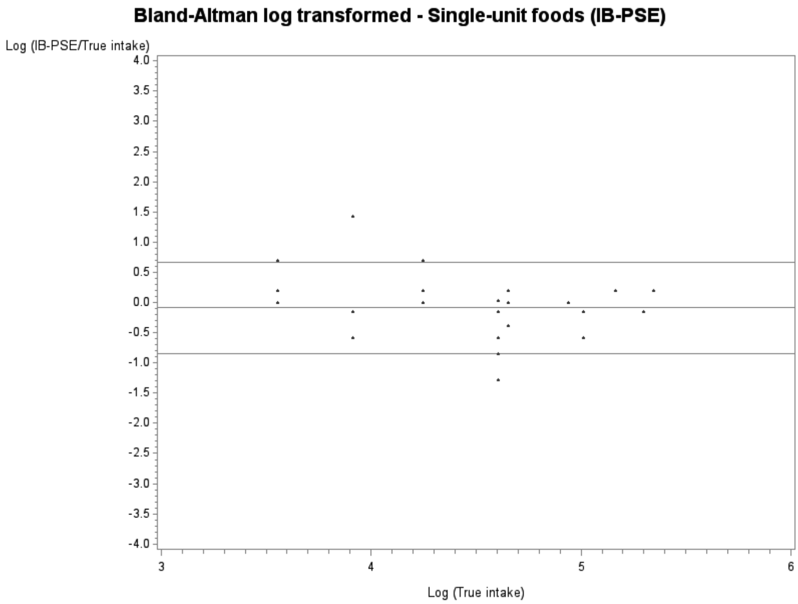


sFigure 8. Bland-Altman plots of log transformed proportion of IB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for liquids.

Single-unit foods

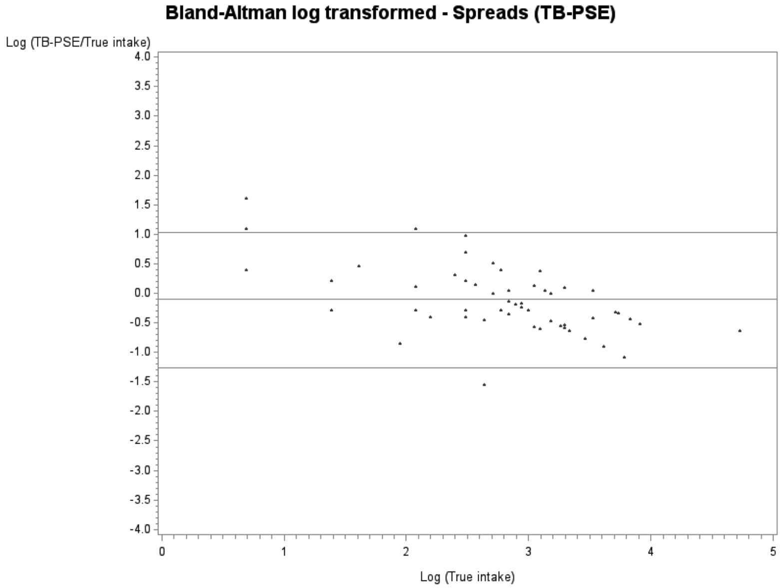


sFigure 9. Bland-Altman plot of log transformed proportion of TB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for single-unit foods.

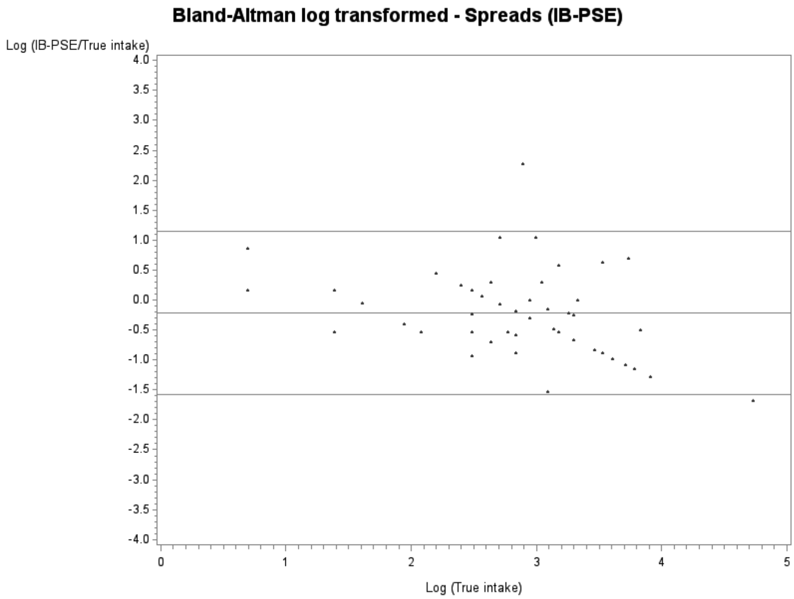


sFigure 10. Bland-Altman plots of log transformed proportion of IB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for for single-unit foods.

Spreads



sFigure 11. Bland-Altman plot of log transformed proportion of TB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for spreads.



sFigure 12. Bland-Altman plots of log transformed proportion of IB-PSE/true intake against log transformed true intake with mean proportion and 95% limits of agreement as reference lines for for spreads.

Chapter 3

Iterative development of an innovative smartphone-based dietary assessment tool: Traqq

Desiree A. Lucassen
Elske M. Brouwer-Brolsma
Anne m. van de Wiel
Els Siebelink
Edith J.M. Feskens

Journal of Visualized Experiments 2021; (169): e62032

ABSTRACT

To collect dietary intake data in a fast and reliable manner, a flexible and innovative smartphone application (app) called Traqq was developed (iOS/Android). This app can be used as a food record and 24-h recall (or shorter recall periods). Different sampling schemes can be created on either prespecified or random days/times within a predetermined period for both methods, with push notifications to urge the participants to register their food intake. In case of non-response, notifications are automatically rescheduled to ensure complete data collection. For use as a food record, respondents can access the app and log their food intake throughout the day. Food records close automatically at the end of the day; recalls close after submission of the consumed items. The recall as well as the food record module provide access to an extensive food list based on the Dutch food composition database (FCDB), which can be accustomed to fit different research purposes. When selecting a food item, respondents are simultaneously prompted to insert portion size, i.e., in household measures (e.g., cups, spoons, glasses), standard portion sizes (e.g., small, medium, large), or weight in grams, and eating occasion/time of consumption. Portion size options can be adjusted, e.g., only entry in grams in case of a weighed food record or time of consumption instead of eating occasion). The app also includes a My Dishes function, which allows the respondent to create their own recipes or product combinations (e.g., a daily breakfast) and only report the total quantity consumed. Subsequently, the app accounts for yield and retention factors. The data are stored on a secure server. If desired, additional questions, i.e., in general or those related to specific food items or eating occasions can be incorporated. This paper describes the development of the system (app and backend), including expert evaluations and usability testing.

INTRODUCTION

Accurate dietary assessment is crucial to ensure the quality of studies on the role of nutrition in health and disease prevention. Currently, such studies generally use established self-report dietary assessment methods, i.e., food frequency questionnaires, 24-h recalls (24hRs), and/or food records [1]. Despite the fact that these methods are of major importance for nutrition research, they also possess various drawbacks, e.g., memory-related bias, social desirability bias, and are burdensome for the respondent as well as the researcher [1, 2]. Recent technological inventions now offer the opportunity to overcome these drawbacks. During the past years, various research groups seized this opportunity and developed web-based and smartphone-based dietary assessment tools for nutrition research that address some of these known drawbacks (see Eldridge, Piernas [3] for an extensive overview of web- and smartphone-based tools), i.e., reduce causes of error, improve user-friendliness, and decrease the participant's and researcher's burden [1].

Nevertheless, the number of fully automated and validated smartphone applications (apps) that are appropriate for nutrition research is still limited. Most of the available dietary assessment apps (i.e., commercially or developed for research) are either not fully automated (i.e., require manual coding of food items) or are not (well) validated [3]. Moreover, most available validated apps have been developed for one specific research purpose and use in a specific country; due to rather fixed designs, re-using such apps for other research purposes or in other countries seems challenging [3-8]. Finally, despite the availability of food record-based apps, to date, there are no recall-based apps. Although food records are prone to reactivity bias, i.e., respondents may alter their food intake due to the awareness that they are being observed [2, 9], this is not the case for recalls, which emphasizes the need for the development of a validated recall-based app [10]. An innovative dietary assessment app called Traqq was developed for use in the Netherlands which can be used as a food record as well as a recall, depending on the research question [1].

Besides the possibility to alternate between the food record option and recall option, this app also differs from other dietary assessment tools because of its flexible nature. Specifically, regarding the food list, portion size estimates, sampling schemes, and the possibility to incorporate additional questions. The level of flexibility in the system enables tailoring to multiple research purposes that require accurate assessment of dietary behaviours. Currently, the app is in the process of being validated and will be ready to be used in various types of nutrition related research. The app can also be used, and perhaps further improved, for use in nutritional intervention programs to measure and influence dietary behaviours. As the development of reliable dietary assessment tools is challenging, and reports on these processes are scarce, especially with respect to user and expert involvement [3, 11, 12], this paper provides a detailed overview on how different information sources were integrated in the systematic and iterative development of this smartphone-based dietary assessment app. The process incorporates theory, expert consultation, and user engagement.

PROTOCOL

NOTE: All the procedures including human participants were conducted in a non-invasive manner by means of mostly qualitative research methods. Informed consent was obtained from all participants before the start of the evaluations. This protocol describes the iterative developmental process that can be roughly divided into four stages in which stages 1–3 are intertwined (Figure 1).

1. CONDUCT EXTENSIVE FORMATIVE RESEARCH IN PREPARATION OF THE DEVELOPMENTAL PROCESS.

1.1. Perform desktop research exploring existing web- and smartphone-based dietary assessment tools, with special attention to features known to be of key importance for accurate food intake data collection, i.e., method of food entry (including the food list and underlying FCDB) and portion size estimations.

1.2. Inspect existing web- and smartphone-based dietary assessment tools focusing on aspects such as dietary assessment methodology, information provision, reliability, search engine, and implemented features (e.g., images, barcode scanner, recipe functions).

1.3. Consult experts in the field of dietary assessment.
NOTE: Results of the desktop research and inspection of existing tools were discussed with experts in the field of dietary assessment, leading to a draft design plan for the development of the app. This draft design plan was evaluated by the experts and further improved as required.

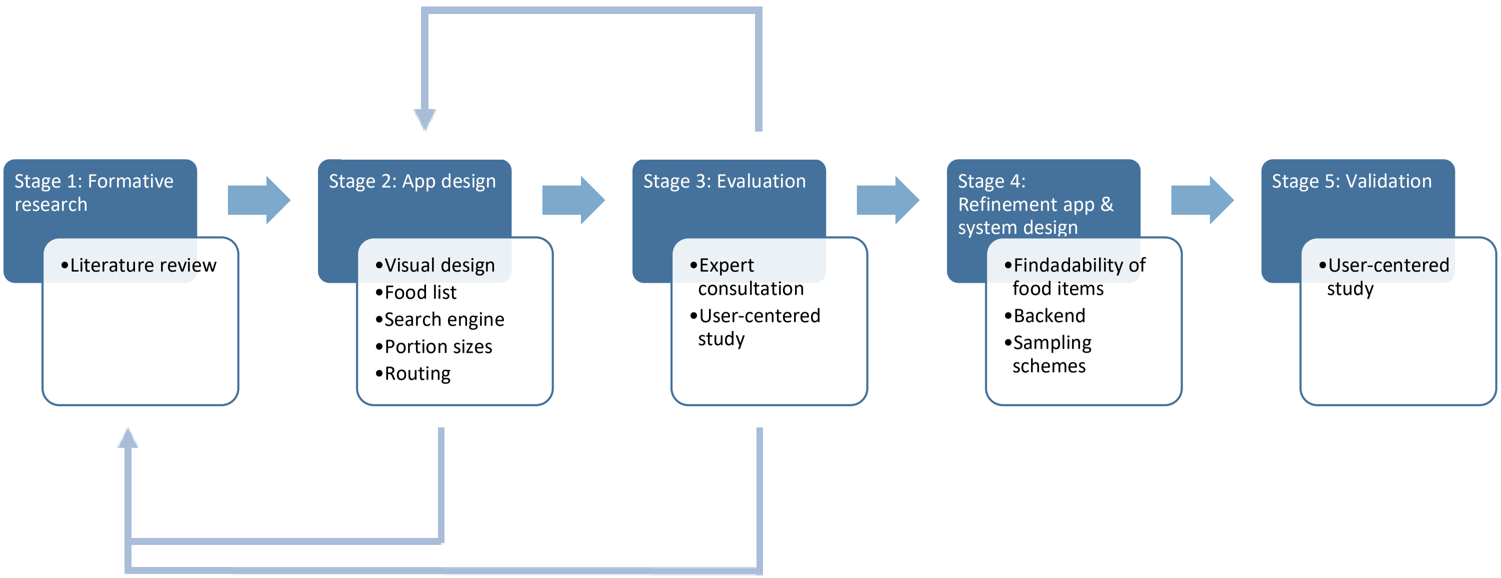


Figure 1. Overview of the stages of the iterative development process of the app. The development process consisted of five stages in total. However, the process was iterative which means that stages 1 through 3 were intertwined.

2. Design the dietary assessment app

2.1. Create the visual design of the app considering important aspects such as animation, branding, colour, layout, and typography [13].
NOTE: As space, colour, fonts, graphics, and interface elements highlight content and convey interactivity, it is essential to incorporate elements facilitating the functionality of the app.

2.2. Select a trustworthy FCDB (here, NEVO) to facilitate nutrient calculations of the collected food intake data.

2.3. Create a food list by critically evaluating the description of the food items mentioned in the FCDB.
NOTE: FCDBs are mostly developed for professional use; food descriptions are often complex and hinder searchability (e.g., "margarine low-fat 35% fat < 10 g of saturated fats unsalted" [14]).

2.4. Formulate search engine requirements; consider the use of punctuation marks, foreign names, misspellings, different search terms, and ranking of search results to facilitate searchability of food items.

2.5. Select portion size estimation (aid) by evaluating various existing dietary assessment tools and field testing of suitable options.

2.6. Design routing within the app to ensure that the user's navigation through the app is logical, predictable, and easy to follow.

2.7. Design backend features and requirements to control the app; include functions related to overall project management, project-specific management (e.g., participants, invitations, data collection), and user management (e.g., authorizations).

3. Evaluations by researchers

NOTE: Following each upgrade, the app was tested by nutrition scientists and research dieticians with expertise in dietary assessment (in-house testing) to verify whether functionalities improved as anticipated. The following instructions are to be executed by researchers.

3.1. Conduct expert evaluation by means of cognitive walkthroughs to simulate a first-time user experience so that the experts can explore the app individually and without guidance [15]. Ensure that the cognitive walkthroughs consist of the following steps.

3.1.1. Make sure that the expert completes a general questionnaire inquiring about the brand and type of smartphone.

3.1.2. Install the app on the expert's smartphone.

NOTE: To ensure proper installation and functioning and minimize the risk of interruptions during the evaluation, it is recommended that the researcher first verifies the app's functionality.

3.1.3. Instruct the expert on test procedures in which each expert is asked to take on the role of first-time user (i.e., research participant). Emphasize that the evaluation is performed from a user's perspective and not from the expert's own perspective.

NOTE: The user was assumed to be an experienced smartphone user and to have knowledge on the use of apps in general. However, this app was used for the first time.

3.1.4. Start the screen and audio recording.

3.1.5. Have the expert complete the cognitive walkthrough while using the app and carrying out a predetermined set of tasks [16]: 1) "I want to record my dinner. I started with a cup of tomato soup and a glass of milk.", 2) "Thereafter, I ate a pasta dish, which I consume regularly and want to enter it as a favourite (i.e., predecessor of My Dishes)." [recipe was provided], 3) "As I also consumed the pasta dish, I want to add this to today's food intake record.", and 4) "I entered everything I ate during dinner. I want to check my entry once more and then submit it.".

NOTE: While performing the tasks, the expert informs the researcher about his/her thought process, i.e., by explaining the steps needed to be completed to fulfil the described task.

3.1.6. Conduct a brief follow-up to clarify ambiguities [17], and provide the expert the opportunity for additional feedback.

3.1.7. Evaluate the results of each expert by checking the recordings to ensure that tasks were executed as intended and by reviewing the additional comments provided.

3.1.8. Share the results with the experts to assess whether assumptions made based on the recordings were correct.

NOTE: Results of the evaluation were discussed and prioritized in consultation with the experts. Based on the results of this evaluation, the app was further upgraded.

3.2. Conduct usability testing with intended users to evaluate the app's usability and likability among the intended users by means of think-aloud interviews and the system usability scale (SUS) [18] by following these steps:

3.2.1. Recruit participants who are representative of the target user population [19].

3.2.2. Instruct the participant regarding the study procedures, including the recording of screen and audio. Then, obtain informed consent from the participants.

NOTE: It is important that the researcher encourage the participant to "think-aloud" during the evaluation, i.e., explaining their thoughts on the required steps to complete each task whilst performing the task, as well as commenting on what functionalities did or did not work well.

3.2.3. Install the app on the participant's smartphone.

NOTE: To ensure proper installation and functioning and minimize the risk on interruptions during the evaluation, it is recommended that the researcher first verify the app's functionality.

3.2.4. Ask the participant to perform a practice task for the think-aloud interview: ask participants to visualize their bedroom and count the number of windows, while telling the researcher about what they saw and thought while counting the windows. Next, ask the participants to approach one of the windows in their bedroom and describe their experiences on their way to that window.

NOTE: A practice task was provided and repeated if needed to ensure that participants felt comfortable to think-aloud as desired [20].

3.2.5. Start the screen and audio recording.

3.2.6. Ask the participant to complete actual think-aloud interview with the predefined tasks: the participant must: 1) record everything they ate and drank during the previous day, and 2) record a regularly consumed dish through the My Dishes function.

3.2.7. During the session, observe, take notes, and stimulate the participants to keep thinking aloud, if needed, by simple prompts such as "Keep talking out loud", "Tell me what you think", or "Tell me what is on your mind". Minimize further interactions to prevent interference with the participant's thought process [15, 17].

3.2.8. Conduct a brief follow-up to clarify ambiguities [17].

3.2.9. Ask the participant to complete an evaluation questionnaire with general questions related to age, sex, educational level, type of smartphone, level of smartphone experience (i.e., experienced users are more likely to perform tasks quick and correctly [21]), as well as the SUS [18]—a 10-item questionnaire to assess the system's usability by means of Likert scale scoring ranging from 1 (strongly disagree) to 5 (strongly agree).

3.2.10. Analyse the data from each session by: 1) transcribing, coding, and creating (sub)themes, and 2) calculating the SUS score using a predefined formula resulting in a score between 0 to 100 [18], where a score of >68/100 indicates that the tool functions at above-average level of usability and a score >80/100 indicates excellent usability [22, 23].

NOTE: It is recommended that the researcher who guided the session analyse the data by using qualitative data analysis software. A second researcher can be consulted in case of ambiguities.

3.3. Conduct quantitative validation of dietary intake records against validated traditional methods and preferably independent measures [3].

NOTE: The app is being validated against web-based and telephone-based (i.e., interviews) 24hRs as well as independent urinary and blood biochemical markers. As the quantitative validation of the app is outside the scope of this paper, this will not be discussed further.

4. USING THE BACKEND SYSTEM FOR APP AND STUDY MANAGEMENT

NOTE: The system has three authorization levels: (1) administrator—this authorization level provides access to all sections of the backend (*i.e.*, creating new users, determining user authorization, and granting users access to one or more projects); (2) project managers—this authorization level allows access to specific projects and the possibility to create new projects; and (3) researchers—this authorization level only provides access to the specific projects that researchers are involved in.

4.1. Management of users and projects in the backend by administrators

4.1.1. Access the online backend via **traqq.idbit.net**, with login credentials (*i.e.*, username, password).

4.1.2. Create a new project by clicking on the **Projects** tab and then on **Create a new project**.

4.1.3. In the next screen, enter the requested project details (*i.e.*, project name, contact description, contact email, contact phone, contact website).

NOTE: Only the project name is mandatory to create a new project. The contact description, email, phone number, and website will become visible in the app under the **Contact & Info** button.

4.1.4. Select the desired features (*i.e.*, product list, ask eating occasion and/or time of consumption, record or recall).

NOTE: Each new project requires individual decision-making with respect to the most appropriate dietary assessment method (*i.e.*, record or recall), food list, portion size estimation, and eating occasion or meal times.

4.1.5. Save the new project by clicking on **Save**.

NOTE: When the screen closes, the administrator returns to the **Project overview** screen.

4.1.6. Next, create a new user by clicking on the **User** tab and then on **Add new user**.

4.1.7. In the following screen, enter a **Username**, a **Password**, and assign the user a **Role** (*i.e.*, administrator, manager, or user).

4.1.8. Save the new user by clicking on **Save**.

NOTE: When the screen closes, the administrator returns to the **User overview** screen.

4.1.9. Assign a user to a project by clicking on the notepad icon (*i.e.*, **Edit** column) for a specific user.

4.1.10. Assign a project by opening the dropdown menu under **Linked Projects**, selecting the desired project, and clicking on **Add**.

NOTE: This action needs be repeated for each project the user needs to be assigned to.

4.1.11. Communicate the log-in credentials to the new user along with the backend URL.

4.2. Management of projects in the backend by researchers (*i.e.*, Manager or User role)

4.2.1. Log in to the backend via **traqq.idbit.net** by using the credentials provided by the administrator.

4.2.2. Click on **Go to projects** to manage the projects.

4.2.3. Click on the arrow in the **View** column for the desired project.

NOTE: After doing this, the researcher is taken to a **Project Overview** page, and new tabs for this specific project appear.

4.2.4. Enter the participants in the backend by clicking on the **Participants** tab. Next, when a **Participant Overview** screen appears, click on **Add new participant.**

4.2.5. In the following screen, enter **Codename, Notes** (optional)**, Login ID, Login Key**, and ending with **Save**.

NOTE: It is recommended that the Participant's study ID be used as both codename and login ID. This minimizes confusion for the participant in case of multiple login credentials. Moreover, the codename is visible in the responses. Using the participant ID makes it easy to use the data. This option needs to be repeated for each participant. For larger groups, **Import participants from file**

(.csv) can be used. Here, the same details are required for each participant. The backend may not contain any personal information of participants.

4.2.6. Schedule invitations for each participant by clicking on the **Invitations** tab. Next, when an **Invitation Overview** screen appears, click on **Add new invitation.**

4.2.7. In the following screen, select a **Participant** from the dropdown menu, and enter **Period start time, Period end time, Opening time, Closing time, Survey URL** (*i.e.*, optional for implementation of additional questions)**, Notes** (optional)**, Enable** (always yes).

NOTE: The **Period start** and **end time** refer to reporting time frame (*i.e.*, what has been consumed between ..:.. and ..:..). In contrast, opening and closing time refer to the period in which the participant can actually report their intake. The correct implementation of an external survey requires some coding; for this, help from the administrator is recommended. For the majority of the invitations, the **Import invitations from** (.csv) option under **File** can be used. The file requires the same information as for the manual input. Invitations can also be created via **Sampling Schemes** (*i.e.*, where the system generates a random invitation scheme across different days and times based on a pre-set of rules such as sampling period, number of required invitations, response deadline). An advantage of the **Sampling Schemes** option is that the system automatically schedules a new invitation in case of non-response.

4.2.8. Track data collection via the **Calendar** tab by selecting a participant of interest from the dropdown menu.

NOTE: The calendar provides an overview of scheduled invitations within a project, either in general or for specific participants. Future invitations are portrayed in blue, completed past invitations are green, while past invitations without response are red. Responses to invitations can also be checked via the **Response** tab.

4.2.9. Track responses via the **Response** tab.

NOTE: In the **Response** section, the reported food intake data (*i.e.*, food item, consumed amount, eating occasion and/or time of consumption) is gathered.

4.2.10. Requests the administrator for data export.

NOTE: Data can be exported from the backend to a .csv file for further analysis (*e.g.*, responses/food intake data, compliance data) by the administrator. Responses include reported food items, selected portion sizes, consumed amounts in grams, and eating occasions/times.

4.2.11. Import the .csv file into nutrition calculation software for in-depth nutrient analyses.

NOTE: The data can be imported into nutrition calculation software that makes use of the Dutch FCDB.

5. USE OF THE APP BY THE PARTICIPANTS DURING THE STUDY

5.1. Download the freely available app from the App Store (iOS) or Google Play Store (Android), and access the app by logging in.
NOTE: Login credentials, as provided by the researcher, are required to access the app (step 4.2.5.). After logging in, the app sends invitations as scheduled in the backend based on the participant's credentials (step 4.2.7.).

5.2. After receiving an invitation via the app, report food intake.
NOTE: Participants can only register their food intake on predetermined days and times.

5.2.1. Open the app by clicking on the notification received or by opening the app via the app icon.
NOTE: After opening the app, an **Invitation Overview** screen appears where previous and current invitations are displayed.

5.2.2. Click on the open invitation.
NOTE: Participant is taken to an **Overview** screen where the invitation period is visible.

5.2.3. Enter the food item consumed first by clicking on **Product toevoegen** (Add food item).
NOTE: The participant is taken to the **Search** screen.

5.2.4. Start typing the name of the consumed item (*e.g.*, orange juice [jus d 'orange]). Click on the desired item as it appears whilst typing.

5.2.5. In the following screen, report the consumed amount (**Hoeveelheid**), corresponding portion size description (**Portie**), eating occasion (**Maaltijdmoment**) and/or time of consumption (**Tijdstip)**, and end by saving (**Opslaan**)**.**

5.2.6. Repeat the aforementioned steps until all food items are reported.

5.2.7. Submit the list (recall) by clicking on **Lijst versturen** (Send list), or the invitation automatically closes at the end of the day (record).
NOTE: The **Send List** option is also visible in the record version, so participants using the record can also send their input to the database. However, even if the data is already sent, the invite still closes at the end of the day, sending all data to the server.

REPRESENTATIVE RESULTS

The system (app and backend) was developed using the steps outlined in the above described protocol; the key results of this process are described below, concluding with the final design of the app.

FORMATIVE RESEARCH

In addition to extensive literature review, several web-based tools were inspected (e.g., Compl-eat [24], ASA24 [25], Foodbook24 [26], MyFood24 [27]) with respect to dietary assessment methodology and implemented features. In addition, the performance of several food tracking apps frequently used in the Netherlands were compared (e.g., MijnEetmeter [28], MyFitnessPal [29], Virtuagym Food [30]), focusing on aspects such as dietary assessment methodology, provision of information, reliability, search engine, and the use of additional features (e.g., images, barcode scanner, recipe functions). The results of this inspection led to the decision to develop the app in such a way that it can be used as a food record and a recall. Moreover, it led to the implementation of the My Dishes function, which can be used to create original recipes or frequently consumed product combinations (e.g., a daily breakfast). Within this function, yield and retention factors are automatically taken into account.

To accurately quantify food and nutrient intake, a complete, albeit practical, food list is crucial. Compiling such a food list requires a trade-off between the extensiveness of the food list and the searchability of the food items (i.e., food descriptions need to be clear, understandable, and easy to locate) [31, 32]. As food composition data form the fundamental basis for dietary assessment [33, 34], it is important to ensure that the developed food list can be linked to accurate food composition data. The food list included in the app is based on the Dutch FCDB (NEVO) [14], which was selected for its reliability and rich food composition data. Originally, the NEVO consists of 2,389 food items (version 2016/5.0), which was reduced to a food list of 1,449 items by eliminating "confusing items" (e.g., foods that cannot be consumed raw, foods that cannot be consumed without additions) or items that are not as essential to include (e.g., due to low consumption rates based on the Dutch Food Consumption Survey (DNFCS) [35]).

Additionally, the NEVO contains similar foods with different brand names; in such a case, only the generic option was included in the food list. To further facilitate usability, some food items were renamed to eliminate needless terminology such as 'prepared', 'frozen', 'average', and 'natural'. This "cleaning protocol" was developed by three well-trained research dietitians and executed by means of a syntax, which can be rerun once NEVO is updated. In addition, to optimize the searchability of food items, 1,019 well-known synonyms of the included foods were added to the food list. Thus, the food list included in the app eventually comprised 2,468 items. An overview of the food list development is displayed in Figure 2. To note, although this extensive food list has been developed for general use, the backend of the app does allow the import of alternative food lists if required.

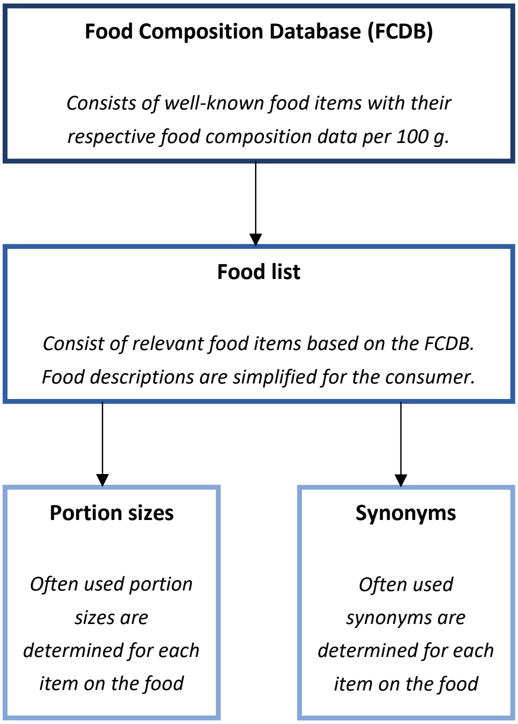


Figure 2. Structure of the food list developed for the app. The food list is based on the Dutch food composition database (FCDB) and corresponding portion size suggestions and synonyms were added for each item in the final food list.

Another crucial aspect of dietary assessment is the quantification of portion sizes. Although portion size estimation aids (PSEAs), e.g., images, referent objects, and standard portion sizes, support the reporting of the amounts of foods consumed [36-38], misreport of portion sizes is still a substantial source of bias [36, 39-41], and literature on the effectiveness of the different PSEAs is inconsistent [38]. Food images, portion size suggestions (i.e., standard sizes and household measures), and free entry of weight in grams are the most used PSEAs in web- and smartphone-based dietary assessment tools. For example, whereas portion size suggestions (e.g., cups, spoons, small, large) are used in tools such as Compl-eat [24] and Oxford WebQ [42], images aid portion size estimates in tools such as ASA24 [25] and

Myfood24 [27]. To investigate the most appropriate PSEA for the app, a pilot study was conducted to compare the accuracy of portion size suggestions (e.g., small, medium, large, or cup, spoon), free entry in grams, and portion size images. The results of this study led to the implementation of portion size suggestions as the PSEA in the app along with the option to enter amounts in grams [43].

EXPERT REVIEW

The aim of the expert evaluations was to qualitatively evaluate the app in terms of functionality and ease of learning. As many users prefer to learn software by exploration [44], a system's level of learnability is important. A total of 10 experts, i.e., 4 (research) dietitians and 6 nutrition and health behaviour experts (scientists) participated in the cognitive walkthroughs in which 60% used an Android smartphone. Most importantly, expert evaluations indicated that the first version of the app was not sufficiently intuitive, e.g., menu-structure was judged unclear due to vague buttons/icons, and the search engine generated an illogical order of results. Another critical point arising from the expert reviews related to the fact that selected items could not be modified. Based on these results, the design of the app was considerably upgraded from stage 2 onwards (Figure 1).

USABILITY EVALUATION

A total of 22 participants participated in the think-aloud interviews, which formed the basis of the usability evaluation. The initial sample size was set at 20 participants [45], after which data saturation was assessed. As data saturation was not reached after 20 interviews, inclusion continued while assessing data saturation after each successive interview. Participants had a mean ± standard deviation age of 48 ± 17 years (range 22–70 years); 36% were male, and the majority of the population was highly educated (55%). In addition, most participants used an Android device (n=14, 64%), and almost all participants had over 1 year of experience with smartphone use (n=21, 96%) (Table 1). All participants completed the tasks without or with minimal instruction.

Table 1. Characteristics of the study population and results of usability evaluation. Only the results of the system usability scale (SUS) are portrayed in this table along with the participant characteristics.

	Total (*n*=22)
Gender	
Male (%)	36.4
Female (%)	63.6
Mean age (mean, SD)	48.1 (17.2)
Educational level	
Low (%)	0
Medium (%)	45.5
High (%)	54.5
Smartphone type	
Android (%)	63.6
iOS (%)	36.4
Smartphone experience	
Shorter than 6m (%)	4.5
Between 6m and 1y (%)	0
Longer than 1y (%)	95.5
SUS (mean, SD)	79.4 (15.1)

Whereas some participants (n=13, 59%) indicated difficulties while using the My Dishes functionality; others (n=5, 23%) encountered minor functionality issues such as slow response of the menu button and difficulties using buttons related to insufficient screen size of smaller smartphones). Moreover, 15 (68%) participants indicated their preference for an option to enter consumed portion sizes in grams. Finally, evaluation of the SUS score indicated a rating of 79/100 (range 40–100), wherein only 3 out of the 22 participants rated the app below 68/100 and 13 rated >80/100, which suggests that the app can be considered user-friendly. Thus, overall, the suggested improvements were minor, and usability evaluations were promising. Subsequently, suggestions for improvement were discussed within the research team and, if deemed relevant, incorporated in the stage 4 upgrade to further optimize likability and usability of the app (Figure 1).

FINAL DESIGN

The steps described in the protocol and the results of the evaluation study eventually resulted in a final design for the app and the backend, which aimed for a simple visual design. This app can be used as a food record and a recall. As described previously, the food list is a modified version of the NEVO. Portion size estimation is supported by food-specific portion size suggestions; consumed portions can also be entered in grams. In case of the recall version of the app, the researcher has the possibility to select different timeslots (e.g., 2hR, 8hR, or 24hR). To collect food intake data on different days and times, various sampling schemes can be created within a predetermined period. Push notifications invite respondents to record their food intake. To ensure complete data collection, invitations are automatically rescheduled in case of non-response. Within the recall module, respondents can only report their food intake after receiving an invitation. In case of the food record, respondents can access the app and log their food intake throughout the day.

In contrast to most 24hR tools, the recall module of the app is not based on the Automated Multiple-Pass Method—a five-step method for collecting food intake data for the previous 24 h [46]—as this method is too elaborate and time-consuming for use in an app. More specifically, to increase usability and enhance the compliance of the food intake recordings [11, 21, 47], navigation was reduced to a minimum by limiting the number of screens that need to be accessed to 4 (Figure 3): 1) an Overview screen showing the reporting window; 2) consumed food items are reported through the Search screen, and once the desired item is selected 3) a dialog box appears probing eating occasion and consumed amount, after which 4) the user returns to the Overview screen now showing the recorded food items. In addition, the user could also use the My Dishes function to create recipes or product combinations, which can be entered via the Menu button.

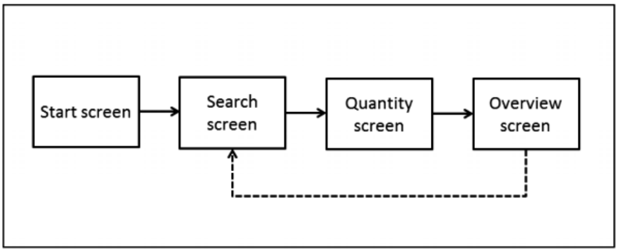
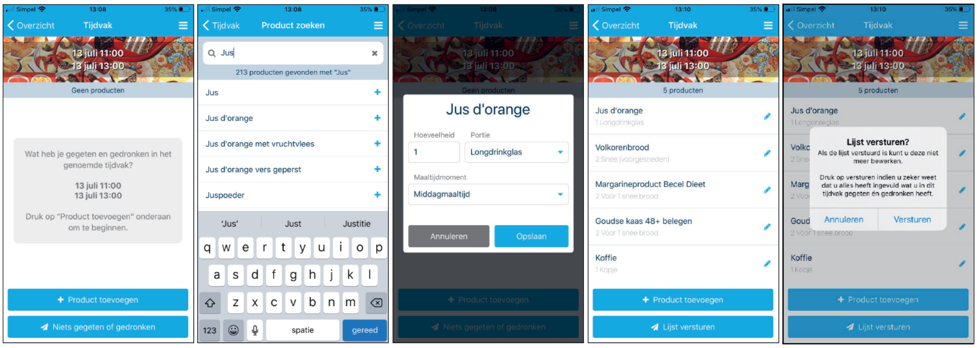


Figure 3. Schematic overview of the routing in the app.

The data are stored on a secure server. If desired, additional questions—general or related to specific eating occasions or food items—can be incorporated. The app can connect with online survey tools. Therefore, it is possible to conduct a survey unrelated to food intake via the app at prespecified times (e.g., context, behavioural, mood questions). It is also possible to ask specific questions related to reported food items or eating occasions (e.g., when apples are reported, when lunch is reported). The use of online survey tools provides an opportunity to ask many different questions via the app. The collected food intake data can be exported from the server and imported into nutrition calculation software for further analyses. In case of the use of additional questions, these data will be available in the survey tool as usual. The aim was to develop a well-structured and easy-to-use app. Some screenshots of the design can be seen in Figure 4A–E.



A B C D E

Figure 4. Screenshots of the final version of the app. (**A**) The **Start/Overview** screen, showing the invitation with the (in this case) 2 h-recall period. The user can press **Product toevoegen** (*i.e.*, Add item) to report a food item or **Niets gegeten of gedronken** (*i.e.*, I did not eat or drink anything) in case nothing was consumed during this time window. (**B**) The **Search** screen, showing results matching the search term “Jus” from the food list. The desired item can be selected from the search results. (**C**) A pop-up screen requires input of details on the selected item “Jus d ’orange”. In this case, the app asks for the amount consumed and eating occasion. The user can go back to the search result by pressing **Annuleren** (*i.e.*, cancel) or **Opslaan** (*i.e.*, save) to go further. (**D**) The **Overview** again, this time showing all the reported items. Another item can be added (**Product toevoegen**) or the input can be sent (**Lijst versturen**). (**E**) After selecting **Lijst versturen**, a pop-up appears asking the user if they are sure that they want to send, and reminds the user that it is not possible to make any more changes after the list has been sent. The user has the option to cancel (**Annuleren**) or send (**Versturen**).

DISCUSSION

This paper presents the iterative developmental process of the smartphone-based dietary assessment app Traqq. Balancing the required level of accuracy and user-friendliness posed the following main challenges in the development of the app related to decisions on 1) data entry (i.e., selecting the most accurate method for food identification and portion size quantification), 2) food composition data (i.e., selecting an accurate database and creating a full-fledged food list), 3) customization options (i.e., flexibility in food list, portion size quantification, and recipes), and 4) validation (i.e., against traditional methods and/or independent measures) [3, 48].

During the literature review, five validated and fully automated, smartphone-based, dietary assessment tools developed for research were identified [3], namely My Meal Mate [4], Electronic Dietary Intake Assessment (eDIA) [7], Easy Diet Diary [8], Electronic Carnet Alimentaire (e-CA) [5], and Eat and Track (EaT) [6]. Owing to the level of automatization of these five dietary assessment apps as well as this app, researcher burden and costs substantially decreases while data completeness increases compared to traditional dietary assessment methods. Additionally, this app, in turn, differs from the five existing dietary assessment tools in terms of flexibility. Specifically, whereas existing apps are all based on the food record method, this app can be used as a food record as well as a recall. Moreover, whereas the design of these apps is fixed, Traqq has the major advantage that it can be modified to fit different research purposes (e.g., dietary assessment method, food list, sampling schemes, additional questions) [3, 48]. Conversely, other existing dietary assessment apps contain valuable features, which are not implemented in the app (yet). To illustrate this point, some apps allow the user to take photographs of their food for food recognition and portion size estimation such as the semi-automated, technology-assisted dietary assessment (TADA) system [49, 50].

Participants in the usability study also indicated that the use of photographs could be a valuable addition to aid portion size estimation. However, there were still too many challenges to be addressed to implement such a feature at this stage, e.g., specifying and guiding with respect to the photographic angle (i.e., to assess depth), the need for a reference maker (i.e., to correct for sizes and colours), the essential before and after photo (i.e., to assess consumed amounts), and on how to process recipe dishes. Due to these technical challenges, the existing image-based dietary assessment apps are still semi-automated, which means manual image review must be done by the user, the researcher, or both [49, 50]. Technological advances, such as crowdsourcing and machine learning, have the potential to improve the use of food images for dietary assessment [51, 52]. In the future, these options will be explored to further improve the app.

The developmental process of the app was characterized by various critical steps. First, a formative research step was completed wherein the scientific concepts underpinning the rationale for app creation facilitated decision-making in setting up the general outline of the app. During this stage, special attention was paid to the selection of the FCDB and the selection of the PSEA—aspects that both directly influence data accuracy [33]. Regarding the FCDB, as the app has originally been developed for use in the Netherlands, its food list is based on the Dutch FCDB, NEVO [3]. In the future, the aim is to further develop the app for international use, which requires more extensive food composition data as many foods are country-specific. Currently, no international FCDB exists yet and if existent, its use might have been limited. More specifically, as the Dutch food list already contains 2,389 food items, the implementation of an international food composition table, e.g., for 5 countries would probably multiply this number of food items by about 5 and negatively affect the searchability of foods and consequently, the app's usability. Therefore, country-specific food lists will probably be most valuable and often also preferred by professionals [53]. This is facilitated by the app as it enables the import of alternative food lists and thus linkage to different (international) food composition tables.

Regarding the portion sizes, there are multiple options available to support the accuracy of the estimates, e.g., use of image booklets, referent objects, and/or textual portion size suggestions [38]. In view of user-friendliness, direct implementation of a PSEA in the app is preferred over using a PSEA alongside the app (e.g., image booklet, referent objects). During the development of the app, the decision was made to facilitate portion size quantification by offering the opportunity to enter portion sizes using portion size suggestions and entry in grams. Portion size suggestions are based on the only available Dutch portion size database [54]. Although Dutch dietary assessment tools such as Compl-eat and Eetmeter also rely on this database [24, 28], it needs to be noted that this portion size database dates from 2003, and tableware sizes have since increased [55]. Using this database may therefore underestimate food intake.

Currently, the portion size database is being updated by the Dutch National Institute for Public Health and the Environment (RIVM), the Dutch Nutrition Centre, and Wageningen University and Research [56], which will eventually be used to update the portion size suggestions in the app. Discrepancies between the old and new portions will be mapped and adjusted where needed. Although the use of portion size images (i.e., a series of images portraying different amounts of a selected food) may be a good alternative for text-based portion size suggestions

[57], research has shown that the accuracy of portion size estimation is highest when a series of portion size images is presented at once, instead of one image at a time [40, 58, 59]. Generally, currently available smartphones have relatively small screens, which limits the presentation of a series of images. Although new technologies facilitate the use of interactive portion size graphics wherein amounts of food on a virtual plate or cup can be increased or decreased by using a slider[60], these techniques are relatively new and still need to be thoroughly evaluated to assess their accuracy.

Another critical step in the development of the app included the involvement of experts and intended end-users. Although not often incorporated in the developmental process of tools (or not described) [11, 12], feedback from experts—as well as intended end-users—is crucial[61], allows maximization of usability, and maintains the required level of accuracy. The feedback of the intended end-users was particularly helpful in the final design of the My Dishes function. Overall, the users were satisfied with the possibility to create their own dishes. However, they did struggle with some of the procedures, for instance, although the function would automatically save data, this was not visible to the user. Therefore, many users kept searching for the Save button and got stuck, afraid to go back and lose their input. Based on these kinds of feedback, the function was improved to better fit the expectations of the user.

To conclude, Traqq is an innovative app with many advantages over existing apps and web-based tools. However, there are still various limitations. As the app still relies on self-report, self-report-related measurement errors still exist (e.g., memory bias (i.e., in case of recall), social desirability bias, and food intake modifications (i.e., in case of food records), inaccurate portion size estimations (i.e., in both)) [1]. In the coming years, recently launched novel technologies will be explored to further advance the app, e.g., by exploring the value of implementing features such as barcode scanners, voice recording, chatbots, and images, which could improve food identification and portion size estimation. Possibilities to connect with other apps (e.g., activity trackers, sleep trackers) and devices (e.g., accelerometers, heart rate monitors, chewing sensors) are being explored as well. Finally, the backend is also being subjected to further development e.g., through the expansion of sampling options.

REFERENCES

1. Brouwer-Brolsma, E.M., et al., Dietary Intake Assessment: From Traditional Paper-Pencil Questionnaires to Technology-Based Tools, in Environmental Software Systems. Data Science in Action. 2020. p. 7-23.
2. Thompson, F.E. and A.F. Subar, Dietary Assessment Methodology, in Nutrition in the Prevention and Treatment of Disease. 2017, Elsevier Inc. p. 5-48.
3. Eldridge, A.L., et al., Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients, 2018. 11(1).
4. Carter, M.C., et al., 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. British Journal of Nutrition, 2013. 109(3): p. 539-46.
5. Bucher Della Torre, S., et al., Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. Nutrients, 2017. 9(1).
6. Wellard-Cole, L., 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, 2019. 11(3).
7. Rangan, A.M., 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, 2015. 3(4): p. e98.
8. Ambrosini, G.L., et al., Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutrition Journal, 2018. 17(1): p. 5.
9. Subar, A.F., et al., Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. The Journal of Nutrition, 2015. 145(12): p. 2639-45.
10. Kipnis, V., et al., Bias in dietary-report instruments and its implications for nutritional epidemiology. Public Health Nutrition, 2002. 5(6A): p. 915-23.
11. Simpson, E., et al., Iterative Development of an Online Dietary Recall Tool: INTAKE24. Nutrients, 2017. 9(2): p. 118.
12. Chen, J., J.E. Cade, and M. Allman-Farinelli, The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment. JMIR Mhealth Uhealth, 2015. 3(4): p. e104.
13. Apple Inc. Human Interface Guidelines. 2017 [cited 2017 September 6th]; Available from: https://developer.apple.com/ios/human-interface-guidelines/.
14. RIVM. NEVO-online (2016/5.0). 2016 [cited 2018 August 15].
15. Jaspers, M.W., A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. International Journal of Medical Informatics, 2009. 78(5): p. 340-53.

16. Penha, A.D.S. and F.A.N.V. dos Santos, Evaluating the use of the cognitive walk-through usability. Human Factors in Design, 2012. 1(1).
17. Fonteyn, M.E., B. Kuipers, and S.J. Grobe, A Description of Think Aloud Method and Protocol Analysis. Qualitative Health Research, 2016. 3(4): p. 430-441.
18. Brooke, J., SUS - A quick and dirty usability scale. Usability evaluation in industry, 1996. 189(194): p. 4-7.
19. Kushniruk, A.W. and V.L. Patel, Cognitive and usability engineering methods for the evaluation of clinical information systems. Journal of Biomedical Informatics, 2004. 37(1): p. 56-76.
20. Davison, G.C., R.S. Vogel, and S.G. Coffman, Think-aloud approaches to cognitive assessment and the articulated thoughts in simulated situations paradigm. Journal of Consulting and Clinical Psychology, 1997. 65(6): p. 950-8.
21. Moumane, K., A. Idri, and A. Abran, Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. Springerplus, 2016. 5: p. 548.
22. Sauro, J., A practical guide to the system usability scale: Background, benchmarks & best practices. 2011: Measuring Usability LLC Denver, CO.
23. Bangor, A., P. Kortum, and J. Miller, Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of Usability Studies, 2009. 4(3): p. 114-123.
24. Meijboom, S., et al., Evaluation of dietary intake assessed by the Dutch self-administered web-based dietary 24-h recall tool (Compl-eat™) against interviewer-administered telephone-based 24-h recalls. Journal of Nutritional Science, 2017. 6.
25. Subar, A.F., et al., The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. Journal of the Academy of Nutrition and Dietetics, 2012. 112(8): p. 1134-7.
26. Timon, C.M., et al., The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. Journal of Medical Internet Research, 2017. 19(5): p. e158.
27. Carter, M.C., et al., Development of a UK Online 24-h Dietary Assessment Tool: myfood24. Nutrients, 2015. 7(6): p. 4016-32.
28. Voedingscentrum. Mijn Eetmeter. 2014 [cited 2017 September 5th]; Available from: https://itunes.apple.com/nl/app/mijn-eetmeter/id663382012?mt=8.
29. MyFitnessPal. MyFitnessPal. 2009-2015 [cited 2017 September 5th]; Available from: http://www.myfitnesspal.com/welcome/learn_more.
30. Virtuagym Food. Virtuagym Food. 2005-2017 [cited 2017 September 5th]; Available from: https://virtuagym.com/food.
31. Evans, K., et al., Development and evaluation of a concise food list for use in a web-based 24-h dietary recall tool. Journal of Nutritional Science, 2017. 6: p. e46.

32. Svensson, A., M. Magnusson, and C. Larsson, Overcoming Barriers: Adolescents' Experiences Using a Mobile Phone Dietary Assessment App. JMIR Mhealth Uhealth, 2016. 4(3): p. e92.
33. Finglas, P.M., R. Berry, and S. Astley, Assessing and improving the quality of food composition databases for nutrition and health applications in Europe: the contribution of EuroFIR. Advances in Nutrition, 2014. 5(5): p. 608S-614S.
34. Gibson-Moore, H., EuroFIR: Where we are now? Nutrition Bulletin, 2013. 38(3): p. 358-362.
35. Rossum van, C.T.M., et al., The diet of the Dutch. Results of the first two years of the Dutch National Food Consumption Survey 2012-2016. 2016, RIVM.
36. Byrd-Bredbenner, C. and J. Schwartz, The effect of practical portion size measurement aids on the accuracy of portion size estimates made by young adults. Journal of Human Nutrition and Dietetics, 2004. 17(4): p. 351-7.
37. Faggiano, F., et al., Validation of a method for the estimation of food portion size. Epidemiology, 1992. 3(4): p. 379-82.
38. Faulkner, G.P., et al., An evaluation of portion size estimation aids: precision, ease of use and likelihood of future use. Public Health Nutrition, 2016. 19(13): p. 2377-87.
39. Hernandez, T., et al., Portion size estimation and expectation of accuracy. Journal of Food Composition and Analysis, 2006. 19: p. S14-S21.
40. Nelson, M., M. Atkinson, and S. Darbyshire, Food photography. I: The perception of food portion size from photographs. British Journal of Nutrition, 1994. 72(5): p. 649-63.
41. Young, L.R. and M.S. Nestle, Portion sizes in dietary assessment: issues and policy implications. Nutrition Reviews, 1995. 53(6): p. 149-58.
42. Liu, B., et al., 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 Nutrition, 2011. 14(11): p. 1998-2005.
43. Lucassen, D.A., et al., The accuracy of portion size estimation using food images and textual descriptions of portion sizes: an evaluation study Journal of Human Nutrition and Dietetics, 2021. Accepted.
44. Wharton, C., et al., The Cognitive Walkthrough Method: A Practicionar's Guide, in Usability Inspection Methods, J. Nielsen and R.L. Mack, Editors. 1994, John Wiley & Sons: New York. p. 79-104.
45. Nielsen, J. How many test users in a usability test? 2012 June 4th, 2012 [cited 2017 December 5th]; Available from: https://www.nngroup.com/articles/how-many-test-users/.

46. USDA. AMPM - Features. 2016 09/08/2016 [cited 2018 February 27th]; Available from: https://www.ars.usda.gov/northeast-area/beltsville-md/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/ampm-features/.
47. Zhang, D.S. and B. Adipat, Challenges, methodologies, and issues in the usability testing of mobile applications. International Journal of Human-Computer Interaction, 2005. 18(3): p. 293-308.
48. Cade, J.E., Measuring diet in the 21st century: use of new technologies. Proceedings of the Nutrition Society, 2017. 76(3): p. 276-282.
49. Ahmad, Z., et al., A Mobile Food Record For Integrated Dietary Assessment. MADiMa16 (2016), 2016. 2016: p. 53-62.
50. Boushey, C.J., et al., New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proceedings of the Nutrition Society, 2017. 76(3): p. 283-294.
51. Fang, S., et al. cTADA: The design of a crowdsourcing tool for online food image identification and segmentation. in 2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI). 2018. IEEE.
52. Shao, Z., R. Mao, and F. Zhu. Semi-Automatic Crowdsourcing Tool for Online Food Image Collection and Annotation. in 2019 IEEE International Conference on Big Data (Big Data). 2019. IEEE.
53. Chen, J., et al., Designing Health Apps to Support Dietetic Professional Practice and Their Patients: Qualitative Results From an International Survey. JMIR Mhealth Uhealth, 2017. 5(3): p. e40.
54. Donders-Engelen, M.R., L.J.M. Van der Heijden, and K.F.A.M. Hulshof, Maten, Gewichten en Codenummers 2003. Food Portion Sizes and Coding Instructions. 2003, Zeist: Wageningen University: Division of Human Nutrition and TNO Nutrition,.
55. Van Ittersum, K. and B. Wansink, Plate Size and Color Suggestibility: The Delboeuf Illusion’s Bias on Serving and Eating Behaviour. Journal of Consumer Research, 2012. 39(2): p. 215-228.
56. RIVM. Portiegrootte voedingsmiddelen. 2019 [cited 2020 July 27]; Available from: https://www.rivm.nl/portiegrootte-voedingsmiddelen.
57. Timon, C.M., et al., A review of the design and validation of web- and computer-based 24-h dietary recall tools. Nutrition Research Reviews, 2016. 29(2): p. 268-280.
58. Kirkpatrick, S.I., et al., The Use of Digital Images in 24-Hour Recalls May Lead to Less Misestimation of Portion Size Compared with Traditional Interviewer-Administered Recalls. The Journal of Nutrition, 2016. 146(12): p. 2567-2573.
59. Subar, A.F., et al., Assessment of the accuracy of portion size reports using computer-based food photographs aids in the development of an automated self-administered 24-hour recall. Journal of the American Dietetic Association, 2010. 110(1): p. 55-64.

60. Figwee. Figwee - Learn More. 2021 NA [cited 2021 February 16]; Available from: https://figwee.com/learn-more/.

61. Preece, J., H. Sharp, and Y. Rogers, Interaction design: beyond human-computer interaction. 2015: John Wiley & Sons.

Chapter 4

DIetary ASSessment (DIASS) study: design of an evaluation study to assess validity, usability and perceived burden of an innovative dietary assessment methodology

Desiree A. Lucassen
Elske M. Brouwer-Brolsma
Anne Slotegraaf
Esther Kok
Edith J.M. Feskens

Nutrients 2022; 14(6): 1156

ABSTRACT

During recent years, the integration of technology has substantially improved self-reported dietary assessment methods, such as food frequency questionnaires (FFQ), food records, and 24-h recalls. To further reduce measurement error, additional innovations are urgently needed. Memory-related measurement error is one of the aspects that warrants attention, which is where new smartphone technologies and ecological momentary assessment (EMA) approaches provide a unique opportunity. In this article, we describe the DIASS study, which was designed to evaluate an innovative 2-h recall (2hR) smartphone-based methodology, against traditional 24-h recalls, FFQ, and biomarkers, to assess both actual and habitual dietary intake. It is hypothesized that a 2-h reporting window decreases reliance on memory and reporting burden, and increases data accuracy. We included 215 men (28%) and women (72%), with a mean ± SD age of 39 ± 19 years and a mean ± SD BMI of 23.8 ± 4.0. Most participants were highly educated (58%). Response rates for the various dietary assessment methods were >90%. Besides the evaluation of the accuracy, usability, and perceived burden of the 2hR methodology, the study set-up also allows for (further) evaluation of the other administrated dietary assessment tools.

INTRODUCTION

Accurate dietary assessment is one of the essential aspects of nutrition and health (behaviour) research, where 24-h recalls (24hRs), food frequency questionnaires (FFQs), and food records are currently the most commonly-used dietary assessment methods [1–4]. However, these methods have a range of drawbacks [2,5]. FFQs and 24hRs are retrospective and, thus, memory dependent, which makes recall bias [6,7] and misreporting nearly inevitable [2,8,9]. A food record is not memory dependent, but its prospective nature may introduce reactivity bias, due to, for instance, social desirability or to ease the recording task [2,7,9]. Finally, all these methods appear to heavily burden both the participant and the researcher [5,7,10].

Accordingly, there is a growing interest in more technology-based dietary assessment methods, which have the potential to improve accuracy and reduce the burden on both participant and researcher [3,11,12]. Numerous valuable computer- and web-based tools, mostly based on 24hRs and FFQs, have been developed during the past decade [3,10,12,13]. More recently, various smartphone applications (i.e., apps) have been developed to collect dietary intake data via digital food records [3,7,12]. Nevertheless, due to their self-report nature, prospective apps are still prone to reactivity bias and still highly intrusive [5,7,10], as illustrated by the time needed to register food intake [3].

To the best of our knowledge, no (validated) recall-based dietary assessment apps exist at present [12]. Although an innovative retrospective app still relies on the participants' memory [2,4,7], apps have the major advantage of enabling (near) real-time collection of dietary intake data [3,7,14,15]. In behavioural and social sciences, this is referred to as ecological momentary assessment (EMA); repeated real-time assessment of individual's behaviour in their natural 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 [15]. EMA opens the possibility of deviating from traditional dietary assessment methods and exploring new data collection efforts. More specifically, it offers the opportunity of deviating from the traditional 24hRs, to shorter recall periods (e.g., 2-h, 4-h, 8-h), which reduces the reliance on a participants' memory, takes less time to complete, and, thus, should have a lower burden for the respondent.

Therefore, we developed an innovative smartphone-based dietary assessment app [16] that can serve to collect dietary intake data in a faster, more flexible, and more reliable manner than the traditional methods. In order to facilitate tailored use, and depending on the purpose of dietary intake collection, the app can be used in the format of a food record or recall. The recall-module is also flexible, in terms of the reporting window; enabling 1-h recalls up to 24hRs. Within the current study, we explored the use of an innovative 2-h recall (2hR) methodology for (near) real-time data collection. A 2-h reporting period minimizes the reliance on memory compared to the traditional 24hRs. The 2-h reporting window was selected over a 1-h reporting window to avoid 'I did not consume anything' responses; e.g., overburdening of the participant. A longer reporting window of, e.g., 3-h or 4-h, was considered, but repudiated owing to a higher memory-related and reporting burden. Therefore, we felt that a 2-h reporting window would result in the lowest participant (perceived) burden relative to the report of a limited number items at once, while limiting memory-related bias. The 2hR methodology is flexible and can be used to assess actual food intake, by sending consecutive 2hRs on one or more full-days or dayparts, depending on the research question. The 2hRs can also be used to assess *habitual* intake, by sending random 2hRs over a longer period of time. However, it needs to be stressed that an adequate sampling scheme is crucial here, i.e., ensuring equal coverage of all eating occasions, allowing the assessment of *habitual* intake.

Although the 2hR methodology sounds promising, its validation against established methods is imperative, to judge its actual value. The DIASS study was designed to evaluate the accuracy of the smartphone-based 2hR methodology, to assess both *actual* and *habitual* intake of food groups, energy, and nutrients compared to established methods and independent biological markers. Secondary aims included the evaluation of the usability, perceived burden, and compliance of the 2hR method compared to established methods. Additionally, the DIASS study allows further evaluation of the other administrated dietary assessment tools. With this article, we aim to provide an overview of the (1) study design of the DIASS study, and (2) baseline characteristics of the study population, as a reference for future evaluation studies that will be performed using these data.

MATERIALS AND METHODS

DESIGN

The DIASS study had a cross-over design (12-weeks), with two experimental conditions; i.e., measuring *actual* intake and *habitual* intake. Dietary intake was assessed by means of the new 2hR app, as well as various established methods. In addition, information on demographics (e.g., educational level, occupation) was collected by means of an online questionnaire derived from the NQplus questionnaire [17]. Additional questions were included regarding the participant's weight stability and sleep pattern. Height and weight were measured on-site. During the study, activity trackers were used to assess physical activity levels. In the final study week participants were invited to complete an evaluation questionnaire regarding the various dietary assessment methods.

Originally, the 215 participants were randomly allocated into six groups (Figure 1). Groups differed in terms of the additional methods used to assess *actual* intake; groups did not differ in terms of the methods used to assess *habitual* intake. Therefore, the habitual groups in Figure 1 are combined. Unfortunately, due to the COVID19 pandemic, the urine and blood collections of 31 participants (groups 2, 3, 5, 6), scheduled from March 2020 onwards, were cancelled. Consequently, these participants were relocated to matching groups without urine and blood collections (groups 4, 6, and newly formed 7).

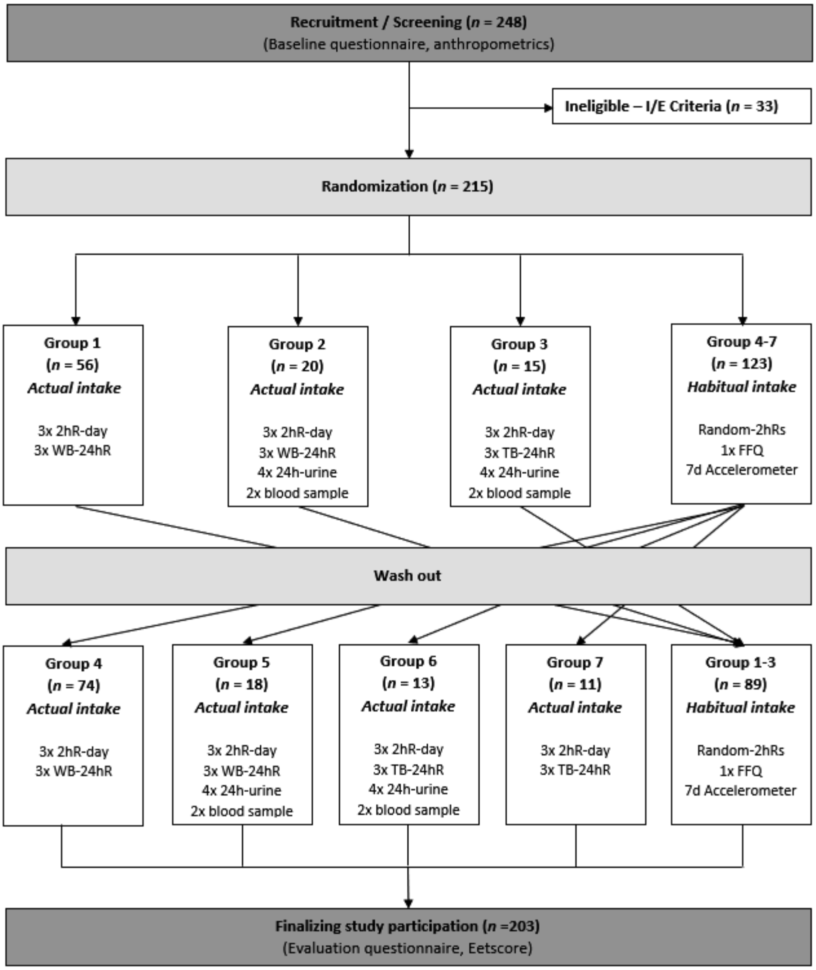


Figure 1. Flowchart study design. 2hR-day: Full day of consecutive 2-h recalls; WB-24hR: web-based 24-h recall; 24-h urine: 24-h urine collection; TB-24hR: telephone-based 24-h recalls; Random-2hRs: Randomly distributed 2-h recalls; FFQ: food frequency questionnaire; Eetscore: Web-based screener for diet quality.

After randomization in week 1, each participant completed two study periods of four weeks each (i.e., week 2–5 and week 8–11), during which, either *actual* or *habitual* intake was assessed, in random order. To minimize the participant burden, and as such optimize compliance, the two study periods were separated by a wash-out period of two weeks (i.e., week 6–7). Additionally, overall diet quality was assessed in week 12. The DIASS study was approved by the ethics committee of Wageningen University and Research (WUR) (ABR No.: NL69065.081.19) and conducted according to the guidelines laid down in the Declaration of Helsinki.

PARTICIPANTS

Recruitment took place between June 2019 and May 2020, and aimed to include 220 Dutch adults (men and women) aged 18–70 years. Eventually, 215 men and women were included in the study. Participants were recruited via the research volunteer database of the division of Human Nutrition and Health of WUR, social media accounts of the division (i.e., Twitter, Facebook), and through flyers and posters. Participants had to be Dutch speaking, not visually impaired, in possession of a smartphone with internet plan, metabolically stable (i.e., gained or lost ≤3 kg in the past 3 months), willing to maintain current dietary habits for the duration of the study, not participating in another dietary intervention study, not an employee of the division, and not having any formal training in the field of nutrition. Written informed consent was obtained from all participants prior to participation.

DIETARY INTAKE ASSESSMENT

Actual food intake was assessed on random non-consecutive days over one of the four-week study periods. Within this period, participants were invited to complete three 2hR-days and three 24hRs. In addition, a random subsample of 69 participants also provided four 24-h urine samples and two fasting blood samples. The urine collections were intentionally coupled to the recall days (i.e., 2× to 2hR-day and 2× to 24hR-day). Blood sampling occurred following two of the urine collections while the participants were at the study centre to hand in their urine containers. During the other four-week study period, habitual food intake was assessed by random 2hRs. The same number of 2hRs was used as for the 2hR-days. However, to assess *habitual i*ntake, the 2hRs were randomly distributed over the four-week period (i.e., 3× each time slot). In case of non-response, the 2hR was automatically rescheduled on the same time on another day. In addition, at the end of the study period, participants were invited to complete a FFQ. In the final study week, participants were invited to complete an additional short FFQ, to assess overall diet quality.

2HRS

The 2hRs were sent via Traqq®, a dietary assessment app developed by WUR [16]. By clicking on the notification/opening Traqq®, participants were able to report their food intake of the previous 2 h. The food intake report screen is supported by an extensive food list, based on the Dutch Food Composition Database [18]. Following the reporting of the food, participants are requested to enter the consumed amount and eating occasion (i.e., breakfast, lunch, dinner, snack). Amount was reported in household measures (e.g., cups, spoons), standard portion size (e.g., small, large) [19] or amount in gram. Participants could also report all ingredients of a recipe and the amount consumed (i.e., yield and retention factors are automatically taken into account) under the 'My Dishes' feature. In addition, this function could also be used to create frequently consumed food combinations (e.g., daily breakfast) to simplify the reporting of these items. When participants did not consume anything in the specified time window they could simply press the 'I did not eat or drink anything' button [16].

Actual intake was measured on thee random days by means of multiple consecutive 2hRs per day (Figure 2). A subsample also collected 24-h urine samples. Therefore, for this group, two of the 2hR-days were first randomly scheduled and then communicated to the participant to ensure 24-h urine collection on the recall day. The remaining 2hR-day was unannounced. On the recall days, participants received an invitation every 2 h to report their food intake during the previous 2 h (Figure 2). To ensure complete data collection, participants also received an additional prompt the following morning, i.e., to report any remaining food intake from the previous night. To stimulate a quick response [20], participants were informed that a 2hR closed after 60 min. However, in reality, the 2hRs remained open until the end of the day and then closed automatically. Participants were allowed to miss one invitation on a 2hR-day. In case of >1 missed 2hRs, the sampling was seen as incomplete and a new recall day was scheduled. On average, participants received eight consecutive 2hR invitations on a recall day. The 2hR-day sampling scheme was individualized according to the participant's sleep pattern (i.e., inquired though the baseline questionnaire).

Prompts	Recall period	Week 1							
Prompts	Recall period	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								1
2	8:00 – 10:00								1
3	10:00 – 12.00								1
4	12:00 – 14:00								1
5	14:00 – 16:00								1
6	16:00 – 18:00								1
7	18:00 – 20:00								1
8	20:00 – 22:00								1
Prevous night	22:00 – 06:00								1
								Total	9

Prompts	Recall period	Week 2							
Prompts	Recall period	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								
2	8:00 – 10:00								
3	10:00 – 12.00								
4	12:00 – 14:00								
5	14:00 – 16:00								
6	16:00 – 18:00								
7	18:00 – 20:00								
8	20:00 – 22:00								
Prevous night	22:00 – 06:00								
								Total	0

Prompts	Recall period	Week 3							
Prompts	Recall period	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								1
2	8:00 – 10:00								1
3	10:00 – 12.00								1
4	12:00 – 14:00								1
5	14:00 – 16:00								1
6	16:00 – 18:00								1
7	18:00 – 20:00								1
8	20:00 – 22:00								1
Prevous night	22:00 – 06:00								1
								Total	9

Prompts	Recall period	Week 4							
Prompts	Recall period	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								1
2	8:00 – 10:00								1
3	10:00 – 12.00								1
4	12:00 – 14:00								1
5	14:00 – 16:00								1
6	16:00 – 18:00								1
7	18:00 – 20:00								1
8	20:00 – 22:00								1
Prevous night	22:00 – 06:00								1
								Total	9

Figure 2. Example *actual* intake sampling scheme (i.e., three 2hR-days).

Habitual intake was measured by multiple random 2hRs. Participants received an invitation on random days and times to report their food intake during the previous 2 h (Figure 3). The same timeslots were used as for the actual intake measurement, but now, randomly divided over the four-week period instead of combined on full days. The 2hRs were restricted to a maximum of two per day to limit the number of recordings on one day. However, the additional question regarding the previous night remained linked to the final evening 2hR. Participants had a 60-min response deadline. In case of non-response, the 2hR closed, and a new invitation was automatically rescheduled for the same time on a different day.

Prompts	Recall period	Week 1 Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								1
2	8:00 – 10:00								1
3	10:00 – 12.00								1
4	12:00 – 14:00								
5	14:00 – 16:00								1
6	16:00 – 18:00								
7	18:00 – 20:00								2
8	20:00 – 22:00								
Prevous night	22:00 – 06:00								
								Total	6

Prompts	Recall period	Week 2 Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								1
2	8:00 – 10:00								1
3	10:00 – 12.00								
4	12:00 – 14:00								1
5	14:00 – 16:00								1
6	16:00 – 18:00								1
7	18:00 – 20:00								
8	20:00 – 22:00								1
Prevous night	22:00 – 06:00								1
								Total	7

Prompts	Recall period	Week 3 Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								
2	8:00 – 10:00								1
3	10:00 – 12.00								1
4	12:00 – 14:00								2
5	14:00 – 16:00								
6	16:00 – 18:00								1
7	18:00 – 20:00								
8	20:00 – 22:00								1
Prevous night	22:00 – 06:00								1
								Total	7

Prompts	Recall period	Week 4 Mon	Tue	Wed	Thu	Fri	Sat	Sun	Total
1	6:00 – 8:00								1
2	8:00 – 10:00								
3	10:00 – 12.00								1
4	12:00 – 14:00								
5	14:00 – 16:00								1
6	16:00 – 18:00								1
7	18:00 – 20:00								1
8	20:00 – 22:00								1
Prevous night	22:00 – 06:00								1
								Total	7

Figure 3. Example *habitual* intake sampling scheme (i.e., three times each 2hR-slot).

WEB-BASED 24HRS

A total of 168 participants were invited to complete thee random web-based 24hRs with Compl-eat™. A subsample of 38 of these participants also collected 24-h urine samples. Consequently, for this group, two of the 24hRs were randomly scheduled and then communicated to the participant to ensure 24-h urine collection on the recall day. The remaining 24hR was unannounced. Compl-eat™ is a validated self-administered web-based dietary 24hR-tool developed by WUR and is based on the automated five-step multiple-pass method [21]. In this method, participants first fill in a quick list of consumed foods and then in the next steps provide detailed information about eating occasion, type of foods, and consumed quantities [22]. The method of reporting intake in Compl-eat™ is similar to the reporting method in Traqq®. Items can be searched for in the food list, and consumed amounts can be reported in similar household measures, standard portion sizes, or in grams [19,22]. Compl-eat™ contains a recipe module similar to the 'My Dishes' function of Traqq®. Moreover, participants were able to make notes for clarifications. Invites for the web-based 24hRs were sent via email at 06:00 in the morning of the recall day. The questionnaire was accessible until midnight that same day. In case of non-response, a new 24hR was randomly scheduled.

TELEPHONE-BASED 24HRS

A subsample of 39 participants were asked to complete interviewer-administered telephone-based 24hRs instead of web-based 24hRs. A subsample of 28 of these participants also collected 24-h urine samples. Therefore, two of the 24hRs were randomly scheduled and then communicated to the participant to ensure 24-h urine collection on the recall day. The final 24hR remained unannounced. If needed, the dieticians made multiple attempts to reach the participant by phone on a recall day. In case of non-response, a new 24hR was randomly scheduled.

Since the validation of Compl-eat™, improvements have been made to the tool that still need to be validated. Therefore, the telephone-based 24hRs were used to ensure the accuracy of Compl-eat™. The telephone-based 24hRs were conducted by trained dieticians using a standardized protocol and the five-step multiple-pass approach [22]. However, to minimize reactivity bias, participants were not informed that the interviews were conducted by dieticians, in attempt to minimize socially desirable answers.

COMPUTATION OF DIETARY RECALL DATA

Data from both the 2hRs and the 24hRs were entered in the computation module of Compl-eat™ [21]. Total intakes of energy, macro-, and micronutrients, and food group intakes (g/d) were calculated using the Dutch Food Composition Database 2016 [18]. Dietary intake data were thoroughly checked by trained dieticians, according to a standardized protocol. The dieticians checked the data for completeness and unusual amounts. Errors were corrected according to a standardized approach, using standard portion sizes and recipes (e.g., 35 slices of bread was corrected to 1 slice of 35 g). Participants were not contacted in case of discrepancies.

FFQ

All participants were asked to complete a validated 183-item semi-quantitative FFQ, with a reference period of four weeks [23]. This extensive FFQ was administered online with the self-administered Dutch FFQ-tool™ [24]. Participants indicated the frequency of consumed food items by selecting answers ranging from 'not consumed' to '7 days per week'. In addition, portion sizes were estimated using natural portions and commonly used household measures. Energy and nutrient contents of foods were based on the Dutch Food Composition Database 2010 [25] and multiplied by the portion size and frequency of consumption to calculate mean daily intake of energy, macro-, and micronutrients. In addition, average daily intake (in grams) of food items were calculated by multiplying frequency of consumption by portion size. Trained dieticians conducted multiple quality checks to safeguard the quality of the data.

Diet Quality

All participants were asked to complete the Eetscore™ in the final study week. The Eetscore™ is a self-administered web-based screener for diet quality [26]. It consists of a 55-item FFQ and is scored with the Dutch Healthy Diet 2015-index to evaluate adherence to the Dutch food-based dietary guidelines [27]. This short FFQ was administered online with the Dutch Eetscore-tool™. Participants indicated frequency of consumed food items by selecting answers ranging from 'never' to 'every day' for regularly consumed foods and from 'not this month' to 'four times a month' for episodically consumed foods (e.g., legumes). Portion sizes were estimated using natural portions and commonly used household measures. Average daily intake of food items were calculated by multiplying frequency of consumption by portion size in grams. Sodium content of food items were based on the Dutch Food Composition Database 2010 [25] and multiplied by the portion size and frequency consumption.

Urine Collection

A total of 66 participants provided four 24-h urine samples during the 'actual intake' period. Two of these samples were linked to 2hR-days and the other two to 24hRs. The participants were instructed on 24-h urine sampling according to a standardized protocol and were provided with three-liter containers containing the preservative lithium dihydrogenphosphate (25 g). Participants also received three 100 mg para-aminobenzoic (PABA) tablets (KAL Vitamins, Salt Lake City, UT, USA), and were instructed to ingest one PABA tablet with each main meal. The 24-h urine collection started with the second voiding after waking up and was completed with the first voiding after waking up the next day. Participants were instructed to record the beginning and end times of the 24-h urine collection, the time of ingesting the PABA tablets, and any possible deviations from the protocol (e.g., missing urine). Urine samples were handed in at the study center where they were mixed, weighed, aliquoted, and stored at −80 °C until further analysis.

PABA was provided to check for completeness of the 24-h urine samples. Research has shown that providing PABA is recommended, but that it does not necessarily have to be analyzed [28]; often creating a feeling of being observed is enough. Moreover, when participants are willing to commit to four 24-h-urine collections, two blood samplings, and four extra visits to the study center, compliance often follows [29]. However, the 24-h urine collections were determined as valid if they met all of the following criteria: (1) collection time of 22–26 h, (2) sample volume ≥500 mL, (3) no more than 1 reported missed void, (4) estimated missed volume ≤5% of the total volume, and (5) creatinine levels of >10 mg/kg for women and >15 mg/kg for men [28]. Urinary creatinine was measured at 520 nm on the Synchon LX20 by the modified Jaffé procedure using a commercial kit.

The 24-h urine samples were assessed on nitrogen, potassium, and sodium content, which were used to estimate absolute intakes of protein, potassium, and sodium, respectively [30,31]. Urinary 24h-nitrogen (N) excretion was determined with the Kjeldahl technique (Foss KjeltecTM 2300 analyzer; Foss Analytical). Urinary protein content was calculated with the following formula: 6·25×(urinary N/0·81), accounting for an assumed 19% of fecal and skin losses [28,32]. Additionally, urinary potassium (K) concentration measurements were performed with an ion-selective electrode on a Roche 917 analyzer. The 24-h K-excretion was calculated by multiplying the total weight of the 24-h urine sample by the K-concentration. Following, this was divided by 0.77, assuming an urinary excretion of 77% [28]. Finally, urinary sodium (Na) was calculated the same way as urinary K. However, an urinary excretion of 86% was assumed [28]. The remaining 24-h urine samples were stored at −80°C for additional analyzes.

Blood Collection

The 66 participants that provided 24-h urine samples also provided two fasting blood samples. Following a 10-h overnight fast, these participants underwent a venipuncture at the study center. The venipunctures were conducted by experienced staff members and scheduled on days that participants were already at the study center to hand in their 24-h urine samples, sparing them extra visits. Biochemical analyses were performed either on a Dimension Vista 1500 automated analyzer (Siemens, Erlangen, Germany) or a Roche Modular P800 chemistry analyzer (Roche Diagnostics, Indianapolis, IN, USA). The blood samples were used to assess carotenoid, folate and n-3 fatty acid concentrations, to estimate habitual intake of fruit and vegetables, folate, and fish, respectively [31,33,34]. The remaining plasma and serum samples were stored at −80°C for additional analyzes.

Anthropometrics

Anthropometrics were conducted by trained staff, according to a standardized protocol in study week 1. Height was measured without shoes, using a stadiometer (SECA 213; SECO Corp., Hamburg, Germany) to the nearest 0.1 cm. Weight was measured without shoes, heavy clothing, and with empty pockets on a digital scale (SECA 877; SECA Corp., Hamburg, Germany) to the nearest 0.1 kg. Weight measurements were repeated in study weeks 7 and 12, to determine weight stability.

Total Energy Expenditure

Total energy expenditure (TEE) is estimated by calculating basal metabolic rate (BMR) and assessing physical activity level (PAL). BMR was calculated using the Harris and Benedict equation [35]. In addition, PAL was assessed over a 7-day period, in which the participant wore the ActiGraph wGT3X-BT accelerometer (ActiGraph LLC, Pensacola, FL, USA), except when showering, bathing, swimming, or involved in contact sports [36]. Accelerometers have been reported to be objective, practical, non-invasive, accurate, and reliable tools to assess physical (in)activity [36,37].

The raw accelerometer data was downloaded from the ActiGraph devices and thoroughly checked by an experienced data scientist. Participants were included if the devices collected accelerometer data of 7 consecutive days. Next, the data were imported into ActiLife version 6.13.4 (ActiGraph LLC, Pensacola, FL, USA), and participant's percentage of time spent in sedentary, light, moderate, vigorous, and very vigorous activity using the Troiano algorithm [38]. The daytime activity percentages were then extracted from ActiLife and multiplied with the corresponding PALs using Python version 3.7 (Python Software Foundation, Wilmington, DE, USA), 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) [39]. This process resulted in an individual PAL for each participant.

Demographics

The baseline questionnaire acquired general participant information (i.e., age, gender, educational level, daytime activities, sleeping pattern, intention to maintain current body weight). This questionnaire was derived from the NQplus study, a large cohort study in the Netherlands [17]. Additional questions were included to assess a participant's sleep pattern in order to personalize the 2hR sampling times.

Evaluation Questionnaire

In the final study week, participants completed an evaluation questionnaire on their experiences using the app and various commonly-used conventional dietary assessment methods. The evaluation questionnaire was based on previous studies and assessed aspects such as ease of use, convenience, perceived reporting burden, perceived accuracy, likelihood of future use, and overall experience [40,41]. Responses were based on a 5-point Likert scale (i.e., strongly agree, agree, neutral, disagree, strongly disagree), or participants could indicate what dietary assessment method matched best with a specific statement. In addition, the participants were asked to complete the system usability scale (SUS) for Traqq® [42]. This is an 10-item questionnaire with a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The SUS has been used in previous studies to assess the usability of dietary assessment apps [43,44].

Statistical Analyses

Baseline characteristics are presented as means with standard deviations (SD) and frequencies (n) with percentages (%). The SUS score was calculated with a predefined formula (range 0–100). A SUS score of >68/100 indicates an above-average usability and a score of >80/100 indicates an excellent usability [42]. Data analyses were performed using SPSS Statistics version 25.0 (SPSS Inc. Chicago, IL, USA).

Baseline Characteristics

In total, 215 men (28%) and women (62%) were included in the DIASS study (Table 1). The participants had a mean ± SD age of 39 ± 19 years and a mean ± SD BMI of 23.8 ± 4.0 kg/m2. According to the accelerometer data the participants had a sedentary lifestyle (mean ± SD PAL of 1.46 ± 0.02). The majority of the participant were either <25 years (42%) or ≥50 years (38%), most of the participants were highly educated (58%), the majority of the participants were either married/registered partners (32%) or single (42%), over half of the participants (52%) had a paid job, and the majority of the participants did not follow a diet regimen (71%). Overall, response rates for the dietary assessment methods were high (>90%). The completeness of the urine collections was 86%, and 100% of the fasting blood samples were collected. Finally, mean SUS score of the app was 72 ± 14.

Table 1. Baseline characteristics of the DIASS participants, including collected dietary intake data.

	N	Total	Men	Women
Men (*n*, %)	215	60 (28)	60 (100)	155 (0)
Mean age, years (SD)	215	39 (19)	45 (19)	37 (18)
Age category (*n*, %)	215			
<25 years		90 (42)	19 (32)	71 (46)
25–50 years		43 (20)	10 (16)	33 (21)
≥50 years		82 (38)	31 (52)	51 (33)
Mean BMI, kg/m^2 (SD)	215	23.8 (4.0)	25.0 (4.3)	23.4 (3.8)
BMI category (*n*, %)	215			
<18.5 kg/m^2		7 (3)	1 (1)	6 (4)
18.5–25 kg/m^2		148 (69)	34 (57)	114 (73)
≥25 kg/m^2		60 (28)	25 (42)	35 (23)
Mean BMR, kcal/day (SD)	215	1545 (211)	1799 (174)	1446 (122)
Mean PAL	203	1.46 (0.02)	1.46 (0.02)	1.46 (0.01)
Educational level (*n*, %)	215			
Low		5 (2)	0 (0)	5 (3)
Intermediate		85 (40)	26 (43)	59 (38)
High		125 (58)	34 (57)	91 (59)
Marital status (*n*, %)	215			
Married/registered partnership		69 (32)	25 (42)	44 (28)
Cohabiting		25 (12)	8 (13)	17 (11)
Serious relationship, not cohabiting		20 (9)	6 (10)	14 (9)
Single		90 (42)	17 (28)	73 (47)
Divorced		7 (3)	3 (5)	4 (3)
Widowed		3 (1)	0 (0)	3 (2)
Other		1 (1)	1 (2)	0 (0)
Paid job currently (*n*, %)	215			
Yes		112 (52)	33 (55)	79 (51)
No		103 (48)	27 (45)	76 (49)
Diet regimen (*n*, %)	204			
Yes, always		35 (17)	4 (7)	31 (21)
Yes, sometimes		24 (12)	6 (11)	18 (12)
Never		145 (71)	45 (82)	100 (67)
Number of complete dietary data collections (*n*, %)				
2hR-day [1]	214	591 (92)	158 (88)	433 (94)
WB-24hR	167	474 (90)	126 (88)	348 (91)
TB-24hR	39	117 (98)	33 (92)	84 (100)
Linked 24-h urine collections	66	238 (86)	73 (83)	165 (88)
Blood sample	66	138 (100)	44 (100)	94 (100)
Random 2hRs	212	4669 (96)	1322 (95)	3347 (96)
FFQ	212	204 (96)	55 (92)	149 (98)
Eetscore	203	192 (95)	54 (98)	138 (93)
Mean System Usability Score (SD)	190	72 (14)	73 (15)	72 (13)

[1] No more than one 2hR missed per day.

DISCUSSION

We have described the design of the DIASS study, which aimed to evaluate a newly developed smartphone-based dietary assessment methodology against established methods and objective markers. A total of 215 men and women were included (18–70 years); mainly women and highly educated. The overall response rates were high with >90% for the dietary assessment approaches and >86% for the collection of biological samples.

Various dietary assessment apps have been developed for research purposes, all based on the food record approach. Similarly to Traqq®, most of these apps rely on text entry for food identification and quantification; respondents select consumed foods from a fixed food list and quantify amounts by weights or household measures [12,41,45–48]. In contrast, some of the available apps rely on digital images for food identification and quantification, i.e., respondents take a before and after picture of each meal. Although this approach seems promising these apps are not fully automated, i.e., require some form of manual image review by user and/or researcher [12,43,44,49–51]. The text- and image-based dietary assessment apps all rely on national databases as the source of food composition data; thus, ensuring the quality of the nutrition calculations. Validation studies of these apps also show good agreement between the apps and the reference methods [12,43,48]. These comparisons are mostly made against established methods (e.g., 24hRs, weighed food records). A limited number of validation studies also included objective measures for total energy expenditure (TEE) from doubly-labelled water or accelerometers [49,50,52]. However, objective measures for nutrient intakes (i.e., biomarkers) are generally lacking [12,43,48]. Despite their limited availability [30,31], biomarkers are more sensitive for quantifying the magnitude and direction of potential measurement errors than traditional self-report dietary assessment methods [53]. Therefore, an important strength of this evaluation study was the collection of biological samples, which offered the opportunity to conduct established urine- and blood-based nutrient biomarker assessments, including nitrogen, potassium, sodium, folate, carotenoids and EPA/DHA, as well as more innovative food metabolomics [54]. Moreover, similarly to a few of the validated apps, we used accelerometers to obtain an objective measure for TEE.

The SUS score of 72 indicates that the app Traqq® has a good usability. This SUS rating was slightly lower compared to our prior usability test (mean SUS of 79) [16]. However, it should be noted that participants in our previous study only used the app for approximately one hour while performing specific tasks in a controlled environment, while during the DIASS study, participants used the app for two periods of four weeks in their day-to-day lives. Obviously, more issues occur with prolonged use (e.g., more missing food items or connectivity issues).

The DIASS study may be considered limited by the fact that about 70% of the participants were highly educated women, which may limit the generalizability of the results. However, this validation study was designed to assess the accuracy of the 2hR-based dietary intake estimates. Therefore, the lack of generalizability may not be as much of a limitation as it would have been in research into diet–disease relationships. However, acceptability and usability levels might be lower for individuals with a lower educational level [55,56]. Therefore, for usage in other study populations, it will be important to perform additional evaluations, and alterations to the method might be required.

Unfortunately, the COVID-19 pandemic forced us to cancel the final blood and urine collections, to safeguard the participants' health. This decreased the number of participants that collected independent biomarkers, from 100 to 69 participants. However, the data from the remaining participants in this subsample still provide valuable insights into the relation between true and assessed intake. Although this last sample of participants were in lockdown for the entire study duration, the self-reported dietary assessment and additional measures could proceed as planned.

In conclusion, the most important feature of the DIASS study is its elaborate study design. The results consist of actual dietary intake data obtained by multiple 2hR-days, multiple web-based or telephone-based 24hRs, and biochemical markers. This allows validation of actual intake assessment with the 2hR-days against both established methods and independent urinary biomarkers. Moreover, the same data can also be used to validate the most recent version of Compl-eat™. In contrast, the habitual dietary intake data obtained by multiple random 2hRs, FFQ, and, again, biochemical markers allows for validation of habitual intake assessment. This data also allow for further validation of the FFQ. Finally, the results of the Eetscore™ can be validated against the results of the FFQ and the 24hRs, as the diet quality score used in the Eetscore™ can also be calculated from these more extensive approaches. Additionally, the collected data can also be used to evaluate all administered methods, in terms of usability, response time, and perceived burden. As such, we believe that the DIASS study offers a unique opportunity for extensive evaluation of a variety of dietary assessment methods and contributes to the further improvements of these methods.

REFERENCES

1. Thompson, F.E.; Kirkpatrick, S.I.; Subar, A.F.; Reedy, J.; Schap, T.E.; Wilson, M.M.; Krebs-Smith, S.M. The National Cancer Institute's Dietary Assessment Primer: A Resource for Diet Research. J. Acad. Nutr. Diet. 2015, 115, 1986–1995. https://doi.org/10.1016/j.jand.2015.08.016.
2. Subar, A.F.; Freedman, L.S.; Tooze, J.A.; Kirkpatrick, S.I.; Boushey, C.; Neuhouser, M.L.; Thompson, F.E.; Potischman, N.; Guenther, P.M.; Tarasuk, V.; et al. Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. J. Nutr. 2015, 145, 2639–2645. https://doi.org/10.3945/jn.115.219634.
3. Shim, J.S.; Oh, K.; Kim, H.C. Dietary assessment methods in epidemiologic studies. Epidemiol. Health 2014, 36, e2014009. https://doi.org/10.4178/epih/e2014009.
4. Brouwer-Brolsma, E.M.; Lucassen, D.; de Rijk, M.G.; Slotegraaf, A.; Perenboom, C.; Borgonjen, K.; Siebelink, E.; Feskens, E.J.M.; de Vries, J.H.M. Dietary Intake Assessment: From Traditional Paper-Pencil Questionnaires to Technology-Based Tools. In Environmental Software Systems. Data Science in Action; Athanasiadis I., Frysinger S., Schimak G., Knibbe W. (eds).; Springer International Publishing: Cham, Switzerland 2020; pp. 7–23.
5. Forster, H.; Walsh, M.C.; Gibney, M.J.; Brennan, L.; Gibney, E.R. Personalised nutrition: The role of new dietary assessment methods. Proc. Nutr. Soc. 2016, 75, 96–105. https://doi.org/10.1017/S0029665115002086.
6. Rothman, K.J.; Greenland, S.; Lash, T.L. Modern Epidemiology; Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2008.
7. Thompson, F.E.; Subar, A.F. Dietary Assessment Methodology. In Nutrition in the Prevention and Treatment of Disease; Elsevier Inc.: Amsterdam, The Netherlands, 2017; pp. 5–48.
8. Young, L.R.; Nestle, M.S. Portion sizes in dietary assessment: Issues and policy implications. Nutr. Rev. 1995, 53, 149–158.
9. Vuckovic, N.; Ritenbaugh, C.; Taren, D.L.; Tobar, M. A qualitative study of participants' experiences with dietary assessment. J. Am. Diet. Assoc. 2000, 100, 1023–1028. https://doi.org/10.1016/S0002-8223(00)00301-1.
10. Timon, C.M.; van den Barg, R.; Blain, R.J.; Kehoe, L.; Evans, K.; Walton, J.; Flynn, A.; Gibney, E.R. A review of the design and validation of web- and computer-based 24-h dietary recall tools. Nutr. Res. Rev. 2016, 29, 268–280. https://doi.org/10.1017/S0954422416000172.
11. Timon, C.M.; Blain, R.J.; McNulty, B.; Kehoe, L.; Evans, K.; Walton, J.; Flynn, A.; Gibney, E.R. The Development, Validation, and User Evaluation of Foodbook24: A Web-Based

Dietary Assessment Tool Developed for the Irish Adult Population. J. Med. Internet Res. 2017, 19, e158. https://doi.org/10.2196/jmir.6407.
12. Eldridge, A.L.; Piernas, C.; Illner, A.K.; Gibney, M.J.; Gurinovic, M.A.; de Vries, J.H.M.; Cade, J.E. Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients 2018, 11, 55. https://doi.org/10.3390/nu11010055.
13. Ngo, J.; Engelen, A.; Molag, M.; Roesle, J.; Garcia-Segovia, P.; Serra-Majem, L. A review of the use of information and communication technologies for dietary assessment. Br. J. Nutr. 2009, 101 (Suppl. 2), S102–S112. https://doi.org/10.1017/S0007114509990638.
14. Stone, A.; Shiffman, S.; Atienza, A.; Nebeling, L. The Science of Real-Time Data Capture: Self-Reports in Health Research; Oxford University Press: New York, NY, USA 2007.
15. Shiffman, S.; Stone, A.A.; Hufford, M.R. Ecological momentary assessment. Annu. Rev. Clin. Psychol. 2008, 4, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415.
16. Lucassen, D.A.; Brouwer-Brolsma, E.M.; van de Wiel, A.M.; Siebelink, E.; Feskens, E.J.M. Iterative Development of an Innovative Smartphone-Based Dietary Assessment Tool: Traqq. JoVE 2021, 169, e62032, https://doi.org/10.3791/62032.
17. Brouwer-Brolsma, E.M.; van Lee, L.; Streppel, M.T.; Sluik, D.; van de Wiel, A.M.; de Vries, J.H.M.; Geelen, A.; Feskens, E.J.M. Nutrition Questionnaires plus (NQplus) study, a prospective study on dietary determinants and cardiometabolic health in Dutch adults. BMJ Open 2018, 8, e020228. https://doi.org/10.1136/bmjopen-2017-020228.
18. The Dutch National Institute for Public Health and the Environment (RIVM). NEVO. Dutch Food Composition Table version 2016/5.0. Dutch Nutrition Centre: Den Haag, The Netherlands, 2010
19. Donders-Engelen, M.R.; Van der Heijden, L.J.M.; Hulshof, K.F.A.M. Maten, Gewichten en Codenummers 2003. Food Portion Sizes and Coding Instructions; Wageningen University: Division of Human Nutrition and TNO Nutrition: Zeist, The Netherlands, 2003.
20. Freyne, J.; Yin, J.; Brindal, E.; Hendrie, G.A.; Berkovsky, S.; Noakes, M. Push Notifications in Diet Apps: Influencing Engagement Times and Tasks. Int. J. Hum.-Comput. Int. 2017, 33, 833–845. https://doi.org/10.1080/10447318.2017.1289725.
21. Meijboom, S.; van Houts-Streppel, M.T.; Perenboom, C.; Siebelink, E.; van de Wiel, A.M.; Geelen, A.; Feskens, E.J.M.; de Vries, J.H.M. Evaluation of dietary intake 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. 2017, 6, e49.

22. Conway, J.M.; Ingwersen, L.A.; Vinyard, B.T.; Moshfegh, A.J. Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. Am. J. Clin. Nutr. 2003, 77, 1171–1178. https://doi.org/10.1093/ajcn/77.5.1171.
23. Siebelink, E.; Geelen, A.; de Vries, J.H. Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. Br. J. Nutr. 2011, 106, 274–281. https://doi.org/10.1017/S0007114511000067.
24. Molag, M. Towards Transparent Development of Food Frequency Questionnaires: Scientific Basis of the Dutch FFQ-TOOL tm: A Computer System to Generate, Apply and Process FFQs; Wageningen University Research: Wageningen, The Netherlands, 2010.
25. The Dutch National Institute for Public Health and the Environment (RIVM). NEVO. Dutch Food Composition Table version 2010/2.0. Dutch Nutrition Centre: Den Haag, The Netherlands, 2010
26. de Rijk, M.G.; Slotegraaf, A.I.; Brouwer-Brolsma, E.M.; Perenboom, C.W.M.; Feskens, E.J.M.; de Vries, J.H.M. Development and evaluation of a diet quality screener to assess adherence to the Dutch food-based dietary guidelines. Br. J. Nutr. 2021, 1–11. https://doi.org/10.1017/S0007114521004499.
27. Looman, M.; Feskens, E.J.; de Rijk, M.; Meijboom, S.; Biesbroek, S.; Temme, E.H.; de Vries, J.; Geelen, A. Development and evaluation of the Dutch Healthy Diet index 2015. Public Health Nutr. 2017, 20, 2289–2299.
28. K
29. Lassale, C.; Castetbon, K.; Laporte, F.; Camilleri, G.M.; Deschamps, V.; Vernay, M.; Faure, P.; Hercberg, S.; Galan, P.; Kesse-Guyot, E. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. Br. J. Nutr. 2015, 113, 953–962. https://doi.org/10.1017/S0007114515000057.
30. Subar, A.F.; Midthune, D.; Tasevska, N.; Kipnis, V.; Freedman, L.S. 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. 2013, 67, 863–867. https://doi.org/10.1038/ejcn.2013.62.
31. Bingham, S.A. Biomarkers in nutritional epidemiology. Public Health Nutr. 2002, 5, 821–827.
32. Jenab, M.; Slimani, N.; Bictash, M.; Ferrari, P.; Bingham, S.A. Biomarkers in nutritional epidemiology: Applications, needs and new horizons. Hum. Genet. 2009, 125, 507–525.
33. Bingham, S.A. Urine nitrogen as a biomarker for the validation of dietary protein intake. J. Nutr. 2003, 133, 921S–924S.
34. Al-Delaimy, W.; Ferrari, P.; Slimani, N.; Pala, V.; Johansson, I.; Nilsson, S.; Mattisson, I.; Wirfalt, E.; Galasso, R.; Palli, D. Plasma carotenoids as biomarkers of intake of fruits

and vegetables: Individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). Eur. J. Clin. Nutr. 2005, 59, 1387–1396.

35. Saadatian-Elahi, M.; Slimani, N.; Chajès, V.; Jenab, M.; Goudable, J.; Biessy, C.; Ferrari, P.; Byrnes, G.; Autier, P.; Peeters, P.H. Plasma phospholipid fatty acid profiles and their association with food intakes: Results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. Am. J. Clin. Nutr. 2009, 89, 331–346.

36. Henry, C.J.K. Basal metabolic rate studies in humans: Measurement and development of new equations. Public Health Nutr. 2005, 8, 1133–1152. https://doi.org/10.1079/phn2005801.

37. Hills, A.P.; Mokhtar, N.; Byrne, N.M. Assessment of physical activity and energy expenditure: An overview of objective measures. Front. Nutr. 2014, 1, 5. https://doi.org/10.3389/fnut.2014.00005.

38. Ward, D.S.; Evenson, K.R.; Vaughn, A.; Rodgers, A.B.; Troiano, R.P. Accelerometer Use in Physical Activity: Best Practices and Research Recommendations. Med. Sci. Sports Exerc. 2005, 37, S582–S588. https://doi.org/10.1249/01.mss.0000185292.71933.91.

39. Troiano, R.P.; Berrigan, D.; Dodd, K.W.; Masse, L.C.; Tilert, T.; McDowell, M. Physical activity in the United States measured by accelerometer. Med. Sci. Sports Exerc. 2008, 40, 181.

40. Joint FAO/WHO/UNU Expert Consultation. Energy and Protein Requirements: Report of a Joint FAO/WHO/UNU Expert Consultation [Held in Rome from 5 to 17 October 1981]; 9241207248; World Health Organization: Geneva, Switzerland, 1985.

41. Spook, J.E.; Paulussen, T.; Kok, G.; Van Empelen, P. Monitoring dietary intake and physical activity electronically: Feasibility, usability, and ecological validity of a mobile-based Ecological Momentary Assessment tool. J. Med. Internet Res. 2013, 15, e214. https://doi.org/10.2196/jmir.2617.

42. Ambrosini, G.L.; Hurworth, M.; Giglia, R.; Trapp, G.; Strauss, P. Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr. J. 2018, 17, 5. https://doi.org/10.1186/s12937-018-0315-4.

43. Brooke, J. SUS—A quick and dirty usability scale. Usability Eval. Ind. 1996, 189, 4–7.

44. Ji, Y.; Plourde, H.; Bouzo, V.; Kilgour, R.D.; Cohen, T.R. Validity and Usability of a Smartphone Image-Based Dietary Assessment App Compared to 3-Day Food Diaries in Assessing Dietary Intake Among Canadian Adults: Randomized Controlled Trial. JMIR Mhealth Uhealth 2020, 8, e16953. https://doi.org/10.2196/16953.

45. Kerr, D.A.; Pollard, C.M.; Howat, P.; Delp, E.J.; Pickering, M.; Kerr, K.R.; Dhaliwal, S.S.; Pratt, I.S.; Wright, J.; Boushey, C.J. Connecting Health and Technology (CHAT): Protocol of a randomized controlled trial to improve nutrition behaviours using mobile devices

and tailored text messaging in young adults. BMC Public Health 2012, 12, 477. https://doi.org/10.1186/1471-2458-12-477.

46. Bucher Della Torre, S.; Carrard, I.; Farina, E.; Danuser, B.; Kruseman, M. Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. Nutrients 2017, 9, 76. https://doi.org/10.3390/nu9010076.
47. Rangan, A.M.; O'Connor, S.; Giannelli, V.; Yap, M.L.; Tang, L.M.; Roy, R.; Louie, J.C.; Hebden, L.; Kay, J.; Allman-Farinelli, M. 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 2015, 3, e98. https://doi.org/10.2196/mhealth.4613.
48. Wellard-Cole, L.; Chen, J.; Davies, A.; Wong, A.; Huynh, S.; Rangan, A.; Allman-Farinelli, M. Relative Validity of the Eat and Track (EaT) Smartphone App for Collection of Dietary Intake Data in 18-to-30-Year Olds. Nutrients 2019, 11, 621. https://doi.org/10.3390/nu11030621.
49. Blanchard, C.M.; Chin, M.K.; Gilhooly, C.H.; Barger, K.; Matuszek, G.; Miki, A.J.; Cote, R.G.; Eldridge, A.L.; Green, H.; Mainardi, F.; et al. Evaluation of PIQNIQ, a Novel Mobile Application for Capturing Dietary Intake. J. Nutr. 2021, 151, 1347–1356. https://doi.org/10.1093/jn/nxab012.
50. Rollo, M.E.; Ash, S.; Lyons-Wall, P.; Russell, A.W. Evaluation of a Mobile Phone Image-Based Dietary Assessment Method in Adults with Type 2 Diabetes. Nutrients 2015, 7, 4897–4910. https://doi.org/10.3390/nu7064897.
51. Pendergast, F.J.; Ridgers, N.D.; Worsley, A.; McNaughton, S.A. Evaluation of a smartphone food diary application using objectively measured energy expenditure. Int. J. Behav. Nutr. Phys. Act. 2017, 14, 30. https://doi.org/10.1186/s12966-017-0488-9.
52. Ahmad, Z.; Kerr, D.A.; Bosch, M.; Boushey, C.J.; Delp, E.J.; Khanna, N.; Zhu, F. A Mobile Food Record For Integrated Dietary Assessment. MADiMa16 2016, 2016, 53–62. https://doi.org/10.1145/2986035.2986038.
53. Boushey, C.J.; Spoden, M.; Delp, E.J.; Zhu, F.; Bosch, M.; Ahmad, Z.; Shvetsov, Y.B.; DeLany, J.P.; Kerr, D.A. Reported energy intake accuracy compared to doubly labeled water and usability of the mobile food record among community dwelling adults. Nutrients 2017, 9, 312.
54. Kipnis, V.; Subar, A.F.; Midthune, D.; Freedman, L.S.; Ballard-Barbash, R.; Troiano, R.P.; Bingham, S.; Schoeller, D.A.; Schatzkin, A.; Carroll, R.J. Structure of Dietary Measurement Error: Results of the OPEN Biomarker Study. Am. J. Epidemiol. 2003, 158, 14–21. https://doi.org/10.1093/aje/kwg091.
55. Brouwer-Brolsma, E.M.; Brennan, L.; Drevon, C.A.; van Kranen, H.; Manach, C.; Dragsted, L.O.; Roche, H.M.; Andres-Lacueva, C.; Bakker, S.J.L.; Bouwman, J.; et al.

Combining traditional dietary assessment methods with novel metabolomics techniques: Present efforts by the Food Biomarker Alliance. Proc. Nutr. Soc. 2017, 76, 619–627. https://doi.org/10.1017/s0029665117003949.
56. Illner, A.K.; Freisling, H.; Boeing, H.; Huybrechts, I.; Crispim, S.P.; Slimani, N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int. J. Epidemiol. 2012, 41, 1187–1203. https://doi.org/10.1093/ije/dys105.
57. Peng, W.; Kanthawala, S.; Yuan, S.; Hussain, S.A. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016, 16, 1158. https://doi.org/10.1186/s12889-016-3808-0.

Chapter 5

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

Desiree A. Lucassen
Elske M. Brouwer-Brolsma
Hendriek C. Boshuizen
Monica Mars
Johan de Vogel-Van den Bosch
Edith J.M. Feskens

Submitted to American Journal of Clinical Nutrition

ABSTRACT

BACKGROUND

Conventional dietary assessment methods are affected by measurement error. We developed a smartphone-based 2-hour recall (2hR) methodology to reduce participant burden and memory-related bias.

OBJECTIVE

Assessing the validity of the 2hR method against traditional 24-hour recalls (24hRs) and objective biomarkers.

METHODS

Dietary intake was assessed in 215 Dutch adults on six randomly selected non-consecutive days (i.e., three 2hR-days and three 24hRs) during a four-week period. Sixty-three participants provided four 24-hour urine samples, to assess urinary nitrogen and potassium concentrations.

RESULTS

Intake estimates of energy (2,052±503 kcal vs. 1,976±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 24hRs. Comparing self-reported protein and potassium intakes 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). Intakes of and energy, nutrients and food groups showed similar reproducibility (ICC) for 2hR-days and 24hRs.

CONCLUSIONS

Comparing 2hR-days with 24hRs showed 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 to 24hRs, suggesting that 2hR-days are a valid approach to assess intake of energy, nutrients and food groups.

INTRODUCTION

Research on the role of nutrition in health and disease prevention mostly relies on self-reported dietary intake data, i.e., 24-hour 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. On the other hand, 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, both researcher and participant burden is high [3].

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 and colleagues [4] for an overview). Compared to 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 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 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].

Yet, all available research and commercial dietary assessment apps are based on the food record approach, and 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 only validated against traditional self-report methods and not against objective measures such as doubly-labelled 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 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-hours, 4-hours) according to the EMA principle. This enables 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 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, food groups, to traditional 24hRs and urinary recovery biomarkers.

METHODS

PARTICIPANTS

The DIASS study was conducted between June 2019 and May 2020 and included 215 participants aged 18-70 years. 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 ≤3 kg in the past 3 months), and willing to maintain their dietary habits for the duration of the study. The DIASS study had a cross-over design with two study periods; one study period focused on *actual* intake (i.e., 2hR-days vs. 24hRs) and one 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 three 2hR-days and three 24hRs (n=162; Supplemental Figure 1), and four 24-hour 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 participants.

STUDY DESIGN

Food intake was assessed on randomly selected non-consecutive days over a four-week study period. Participants completed three 2hR-days and three 24hRs (i.e., either web-based or interviewer-administered). Recall days were randomly selected and scheduled over the four-week study period using the statistical analyses system (SAS) version 9.4 (SAS Institute Inc, Cary, NC, USA), except when in combination with urine collections. Urine collections were matched to the recall days (i.e., 2x to 2hR-day and 2x to 24hR-day) where recall days were randomly scheduled and then preannounced to facilitate the 24-hour urine collection on the recall days. In case of non-response, a new day was randomly selected and scheduled.

5

METHODS OF DIETARY ASSESSMENT

2-hour 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 two hours. On average, participants received eight consecutive 2hR invitations on a recall day, see Figure 1 for an example scheme of a 2hR-day. Notifications were sent at the end of each 2-hour interval with a reporting window of 1 hour (e.g., interval 6:00-8:00; notification at 8:00; reporting deadline at 9:00). The morning after the recording day, another invitation was sent to report on potential nighttime food intake (e.g., nighttime interval 22:00-6:00; notification at 8:00; reporting deadline at 9:00). The 2hR-day sampling scheme was individualized based on the participant's sleeping pattern, as inquired via the baseline questionnaire, to minimize the risk that participants where disturbed while sleeping. To illustrate, if a participant indicated to wake up at 9:00, the first notification was sent at 10:00 instead of 8:00. 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 occasion, i.e., breakfast, lunch, dinner, snack. Quantity can be reported in household measures (e.g., spoon, cup), standard portion sizes (e.g., small, large) or amount in gram. 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.

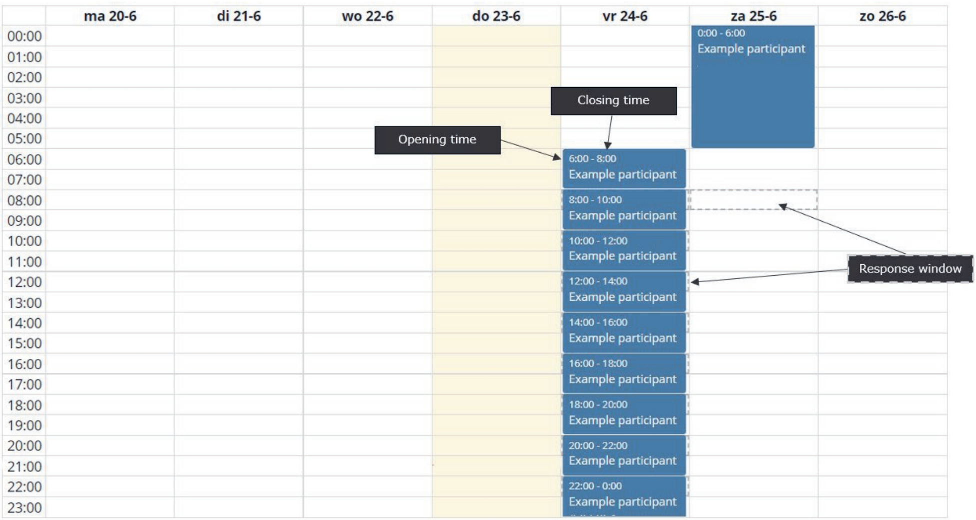


Figure 1. Example of a 2hR-day sampling scheme.

24-hour recalls

Participants also completed three random non-consecutive web-based 24hRs (n=128, 79%) or three random non-consecutive 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 five-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 detailed information about type of foods, consumed quantities, and eating occasion [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 gram [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 email at 6:00 in the morning on 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 gram. The interviewer-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 as the most accurate version of the 24hR method and included to ensure the accuracy of Compl-eat™, no major differences were found between results of 24hRs administered via Compl-eat™ and by telephone (results not included here). 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 was 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

Urine collection (24h) was performed according to a standardized protocol. Participants received three-liter containers containing the preservative lithium dihydrogenphosphate (25 g), three 100 mg para-aminobenzoic acid (PABA) tablets (KAL Vitamins, Salt Lake City, UT, USA), 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 completeness of the 24h-urine samples, participants were instructed to ingest one PABA tablet with each main meal (i.e., breakfast, lunch, diner), and 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 completeness of the urine sample. Urinary creatinine concentrations were measured at 520 nm on the Synchron 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 22-26h, 2) sample volume ≥500 ml, 3) no more than 1 reported missed void, 4) estimated missed volume ≤5% of the total volume, and 5) creatinine levels of >10 mg/kg for women and >15 mg/kg for men [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:

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

Finally, 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-hour potassium intake was calculated with the following formula:

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

OTHER VARIABLES

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

Physical activity levels were assessed over a 7-day period by means of the ActiGraph wGT3X-BT (ActiGraph LLC, Pensacola, FL, USA). The ActiGraph was not worn during showering, bathing, swimming, or contact sports [23]. The accelerometer data was 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 physical activity level (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, intake-related bias, person-specific bias and a constant bias were assumed. Biomarker is assumed to be an unbiased measurement. To evaluate the comparability of the 2hR-days and the 24hR-days with the biomarkers as reference method, the following measurement error modes were used:

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

2hR-days or 24hRs (R) $$R_{ij} = \alpha_R + \beta_R(\bar{T}_i + \Delta T_{ij}) + w_{Ri} + \epsilon_{Rij} \qquad (2)$$

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 $\bar{T}_i$, while the true intake on day j is given by $\bar{T}_i + \Delta T_{ij}$. The person-specific bias of the method is given by w_{Ri} and the random error by ϵ_{Rij}. $\bar{T}_i$, ΔT_{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 $\mathrm{var}\epsilon_{Xij}$ 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 absence of proportional scaling bias in the reference method ($\beta_X = 1$) were made to enable estimation of the model parameters. To evaluate the 2hR-days with the 24hRs as 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 cut-offs for both methods (n=16). Participants were identified as dietary under- or over-reporters if their ratio of average daily total energy intake to basal metabolic rate (EI:BMR) fell outside an individualized cut-off. BMR was calculated using the Harris and Benedict equation, taking into account gender, age, weight, and height [27]. Individual cut-offs were estimated using the method recommended by Black [28]. 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 [29]. 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) [29]. Second, absolute differences between the 2hR-days and the 24hRs were evaluated using paired t-tests. Third, 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 [29]. 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 in addition Bland-Alman plots were created to examine the level of agreement.

Validity coefficients and attenuation factors were calculated using the estimates of the measurement error models. 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 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, and can be used to correct for measurement error in future studies on diet-disease 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:

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

Attenuation factor: $\lambda_X = \frac{\rho^2{}_{XT}}{\beta_X}$ (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 24hR comparison; k=1 for biomarker comparison).

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

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

All analyses were performed using IBM SPSS Statistics version 25.0 (SPSS Inc. Chicago, IL, USA) and SAS Software version 9.4 (SAS Institute Inc, Cary, NC, USA).

5

RESULTS

PARTICIPANT CHARACTERISTICS

Participants had a mean±SD age of 40.4±18.8 years, 73% were women 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).

Table 1. General characteristics of the participants included in this validation study

	Total (146)	Men (39)	Women (107)
Mean age, years (SD)	40.4 (18.8)	46.8 (18.8)	38.1 (18.3)
Age category (n, (%))			
<25 years	55 (38)	10 (26)	45 (42)
25-50 years	34 (23)	7 (18)	27 (25)
>50 years	57 (39)	22 (56)	35 (33)
Mean BMI, kg/m^2 (SD)	23.6 (3.8)	24.8 (4.2)	23.2 (3.5)
BMI category (n, (%))			
<25 kg/m^2	107 (73)	22 (57)	85 (79)
25.0-29.9 kg/m^2	30 (21)	13 (33)	17 (16)
≥30 kg/m^2	9 (6)	4 (10)	5 (5)
Mean estimated BMR, kcal/d (SD)[1]	1,530 (200)	1,780 (144)	1,439 (127)
Mean PAL (SD)	1.46 (0.01)	1.46 (0.02)	1.46 (0.01)
Educational level (n, (%))			
Low[2]	3 (2)	0 (0)	3 (3)
Intermediate[3]	53 (36)	16 (41)	37 (34)
High[4]	90 (62)	23 (59)	67 (63)
Diet regimen (n, (%))			
Yes, always	25 (17)	3 (8)	22 (21)
Yes, sometimes	18 (12)	4 (10)	14 (13)
Never	103 (71)	32 (82)	71 (66)
Preferred method (n, (%))			
2hR-days	126 (87)	35 (90)	91 (85)
24hRs	15 (10)	4 (10)	11 (10)
Unknown[5]	5 (3)	0 (0)	5 (5)

[1] Based on the Harris and Benedict equation, for men: BMR=(66.4730+(13.7516*weight))+(5.0033*height)-(6.7750*age), and for women: BMR=(655.0955+(9.5634*weight))+(1.8496*height)-(4.6756*age) [27].

[2] Primary or lower education.

[3] Secondary or higher vocational education.

[4] University or college.

[5] Answering was not compulsory.

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

Estimated 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). 2hR-days and 24hRs showed relatively similar intakes of energy and nutrients (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, vitamin B2, 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.

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

Statistically significant paired-t-tests were only found for ‘alcoholic beverages’, ‘grains and cereals’, ‘non-alcoholic 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 confectionary’, and ‘vegetables’. Group-level bias was large (>10%) for ‘composite dishes’, ‘non-alcoholic 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 also 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, 0.09, respectively). In agreement, 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 ‘non-alcoholic beverages’.

Table 2. Energy and nutrient intakes assessed by 2hR-days and 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 (95% CI)[3]	Validity coefficient (95% CI)	Attenuation factor (95% CI)
	Mean ± SD	Mean ± SD					
Energy (kcal)	2,052 ± 503	1,976 ± 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 protein (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)
Carbohydrates (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)	3,114 ± 820	2,994 ± 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)	2,922 ± 3,582	3,299 ± 4,470	-11.4	0.23	0.41 (0.26, 0.54)	0.73 (-1.03, 2.48)	0.42 (-0.27, 1.12)
Vitamin B1 (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)
Vitamin B2 (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) x 100 - 100.

[2] Paired-t-test between mean intake assessed with 2hRs and 24hRs.

[3] Spearman correlation.

5

Table 3. Intake of food groups (g/d) assessed by 2hR-days and 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	P [2]	Correlation	Validity coefficient	Attenuation
	Mean ± SD	Mean ± SD	bias (%)[1]		coefficient (95% CI)[3]	(95% CI)	factor (95% CI)
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, 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)
Meats 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)
Non-alcoholic 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, 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, 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) x 100 - 100.

[2] Paired-t-test between mean intake assessed with 2hRs and 24hRs.

[3] Spearman correlation.

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 coefficients; 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) x 100 - 100.

[2] Paired-t-test between mean intake assessed with 2hRs and 24hRs.

[3] Spearman correlation.

5

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 correlation between the self-report measures and urinary nitrogen 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 vs. urinary potassium (0.62), which was lower and acceptable for 24hRs (0.45). Validity coefficients for protein intake (0.41 for 2hR-days vs. biomarker; 0.44 for 24hRs vs. biomarker) where acceptable. For potassium, a good validity coefficient was found for 2hR-days vs. 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 intake 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 ($\beta = -0.24$, $p = 0.12$), while the regression line of differences for the comparison to 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 while the intake increased.

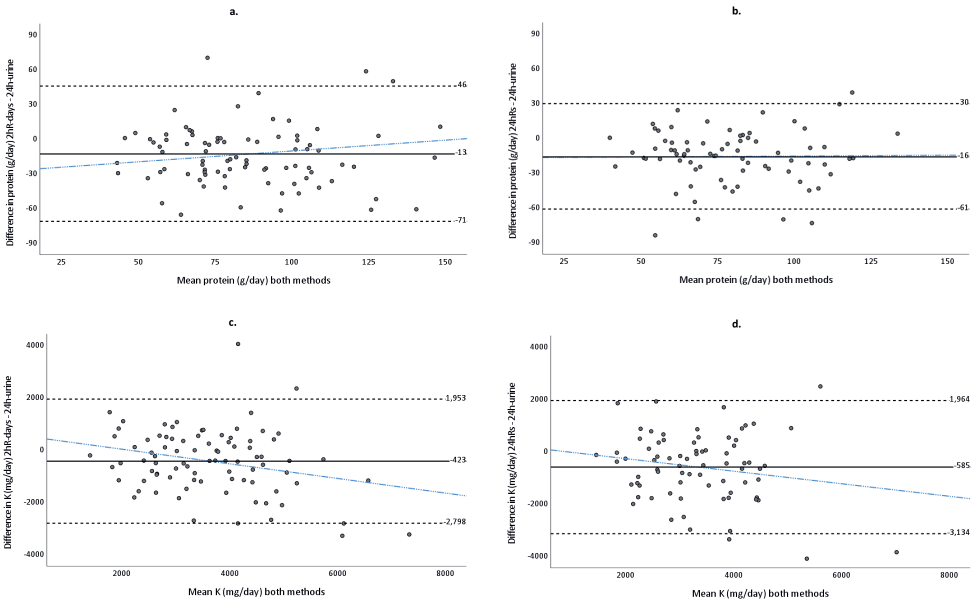


Figure 2. Bland-Altman plots of the differences in intake estimated with the self-report measure and the biomarker, plotted against the mean of both methods (g/d). Mean difference (solid line), 95% limits of agreement (1.96xSD of mean difference; dashed line), and linear regression line (blue dashed 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.

REPRODUCIBILITY OF THE 2HR-DAYS COMPARED TO THE 24HRS

The ICC for repeated 2hR-days showed acceptable reproducibility for energy and macronutrients (range 0.27-0.49), which is similar as the reproducibility observed for the repeated 24hRs (range 0.21-0.51) (Table 5). 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 (Table 6), except for the group 'fats, oils and savory sauces' (0.31 and 0.14, respectively), 'grains and cereals' (0.33 and 0.18, respectively), and 'vegetarian products' (0.53 and 0.48, respectively) where the ICC was higher for the 2hR-days than for the 24hRs.

Table 5. Intra-class correlations (ICC) and 95%CI for energy and nutrient intake between the three 2hR-days and between the three 24hRs

	ICC 2hRs	95% CI	ICC 24hRs	95% CI
Energy (kcal)	0.42	0.32, 0.52	0.48	0.39, 0.58
Protein (en%)	0.30	0.19, 0.40	0.21	0.10, 0.31
Protein (g)	0.36	0.25, 0.46	0.45	0.35, 0.55
Plant-based protein (g)	0.44	0.35, 0.53	0.51	0.42, 0.60
Animal protein (g)	0.32	0.21, 0.42	0.41	0.31, 0.51
Fat (en%)	0.45	0.35, 0.55	0.24	0.14, 0.35
Fat (g)	0.43	0.33, 0.53	0.40	0.30, 0.50
SFA (g)	0.37	0.27, 0.47	0.37	0.27, 0.48
MUFA (g)	0.44	0.34, 0.54	0.32	0.21, 0.42
PUFA (g)	0.36	0.26, 0.47	0.33	0.23, 0.44
Cholesterol (mg)	0.27	0.18, 0.37	0.25	0.16, 0.35
Carbohydrates (en%)	0.48	0.38, 0.57	0.37	0.28, 0.47
Carbohydrates (g)	0.45	0.36, 0.54	0.49	0.40, 0.58
Mono and disaccharides	0.48	0.38, 0.57	0.45	0.35, 0.55
Polysaccharides	0.42	0.33, 0.52	0.43	0.33, 0.52
Fiber (g)	0.49	0.39, 0.58	0.49	0.39, 0.58
Alcohol (en%)	0.32	0.22, 0.41	0.37	0.27, 0.46
Alcohol (g)	0.31	0.21, 0.41	0.32	0.22, 0.42
Ca (mg)	0.32	0.21, 0.52	0.37	0.27, 0.48
Fe (mg)	0.49	0.39, 0.58	0.37	0.27, 0.48
K (mg)	0.49	0.39, 0.58	0.46	0.36, 0.56
β-Carotene (μg)	0.48	-1.33, 2.28	0.80	0.78, 0.82
Vitamin B1 (mg)	0.20	0.11, 0.30	0.21	0.12, 0.31
Vitamin B2 (mg)	0.41	0.31, 0.51	0.42	0.32, 0.52
Vitamin B6 (mg)	0.30	0.19, 0.41	0.26	0.15, 0.37
Vitamin B12 (μg)	0.04	-0.01, 0.09	0.26	0.16, 0.37
Vitamin C (mg)	0.25	0.15, 0.36	0.21	0.11, 0.32
Vitamin D (μg)	0.19	0.08, 0.29	0.25	0.15, 0.36
Vitamin E (mg)	0.34	0.23, 0.44	0.24	0.13, 0.34
Folate (μg)	0.32	0.21, 0.42	0.35	0.24, 0.45

Table 6. Intra-class correlations and 95%CI for intake of food groups between the three 2hR-days and between the three 24hRs

	ICC 2hRs	95% CI	ICC 24hRs	95% CI
Alcoholic beverages	0.27	0.16, 0.38	0.39	0.29, 0.50
Bread	0.43	0.34, 0.52	0.38	0.29, 0.47
Cheese	0.20	0.11, 0.30	0.19	0.09, 0.28
Composite dishes	0.01	-0.03, 0.05	0.06	-0.03, 0.15
Dairy	0.51	0.42, 0.60	0.50	0.41, 0.59
Eggs	0.15	0.06, 0.24	0.13	0.04, 0.22
Fats, oils, savory sauces	0.31	0.20, 0.41	0.14	0.03, 0.24
Fish	0.00	-0.02, 0.03	0.13	0.03, 0.23
Fruit	0.41	0.31, 0.51	0.35	0.24, 0.45
Grains and cereals	0.33	0.23, 0.44	0.18	0.08, 0.29
Legumes	0.11	0.01, 0.22	0.09	-0.01, 0.19
Meats and poultry	0.31	0.21, 0.42	0.25	0.15, 0.36
Non-alcoholic beverages	0.71	0.65, 0.78	0.72	0.66, 0.79
Nuts, seeds, snacks	0.09	0.00, 0.19	0.06	-0.01, 0.14
Pastry, cake, biscuits	0.23	0.13, 0.33	0.20	0.11, 0.29
Potatoes	0.09	0.01, 0.17	0.07	-0.00, 0.15
Savory sandwich fillings	0.23	0.14, 0.32	0.27	0.18, 0.37
Soups	0.14	0.04, 0.24	0.17	0.07, 0.28
Sugar and confectionery	0.34	0.23, 0.44	0.32	0.22, 0.43
Vegetables	0.44	0.34, 0.53	0.35	0.25, 0.44
Vegetarian products	0.53	0.44, 0.62	0.48	0.39, 0.57

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 to validated 24hRs. Validation against objective urinary biomarker 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.

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, Dutch National Food Consumption (DNFCS) data show a median 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 [30]. 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); thus we can conclude that 2hR-days provides a more precise and accurate estimate of protein intake than the 24hRs. In terms of alcohol, DNFCS data show a 11 g/d (95% CI: 10-12 g/d) median intake estimate, which is close to our 24hR estimate (9±12g/d) but higher than our 2hR-day intake estimate (7±11 g/d) [30]. Accordingly, a similar difference between 2hR-day and 24hR is observed for the food group 'alcoholic beverages' (109 g/d vs. 141 g/d). As it is well known that alcohol consumption varies highly accross days [31], it is difficult to detemine the exact origin and direction of this difference between 2hR-day and 24hR [32]. A possible explanation could be the short reporting deadline of the nighttime recall (i.e., 1 hour), 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 episodical consumed foods such as alchoholic beverages.

Differences in absolute micronutrient intakes were relatively small, only β-carotene, vitamin B2, 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 reported intake between the 2hR-days (0.04 and 0.19, respectively) as compared to 24hRs (0.26 and 0.25, respectively), Therefore, these differences could be caused by day-to-day variation in, for instance, reported 'fish'. 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) [31].

Group-level bias was low for the majority of regurlarly consumed food groups (≥5 days/week according to the DNFCS), i.e., 'bread', 'dairy', 'fruit', 'meats and poultry', 'vegetables' [30] suggesting an 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 'non-alcoholic 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 suggests better recollection and thus reporting with shorter recall periods. In contrast, intake estimates for non-alcoholic beverages were higher for 2hR-days than 24hRs, which may be explained by the fact that non-alcoholic beverages are often consumed throughout the day, not always linked to specific eating ocassions, 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 'non-alcoholic beverages' being the highest. Relatively similar attenuation factors were observed for urinary protein (2hR-days: 0.27 vs. 24hR: 0.38) and potassium (2hR-days: 0.52 vs. 24hRs 0.55). Freedman and colleagues observed a similar range of attenuation factors for protein in their biomarker analyses (0.14-0.54) [33].

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 [34], evaluation studies mostly show lower intake estimates of energy and macronutrients by food record apps compared to 24hRs [9-12, 35, 36]. 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 two hours of the day and immediately send it to an external server after which data are not visible anymore for the participant. With regular food records, food intake reports remain visible throughout the day, which increases the likelihood of introducing reactivity bias. All in all, 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 to food record based-apps.

Although the design of this validation study is well thought-out, 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 the 24hR and 2hR 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 women, which may have affected the generalizability of the validation results. Therefore, additional validation with a more diverse population might be needed. Nevertheless, considering that 2hRs have a lower reliance on memory makes it a promising approach for use in populations with decreasing cognitive abilities. A strength of the current study is that we used multiple tests to assess validity of 2hR-days, which has been suggested as the most optimal approach to assess validity of a dietary assessment method [29, 32].

In conclusion, the use of 2hR-days are a reliable approach to assess actual intake of energy, nutrients and food groups. The group-levels bias was relatively low for the majority of the macro- and micronutrients and food groups. Comparisons with biomarkers showed a smaller underestimation of protein and potassium intake by 2hR-days as compared to 24hRs. Finally, the majority of the participants indicated to prefer the use of 2hR-days of the 24hRs.

REFERENCES

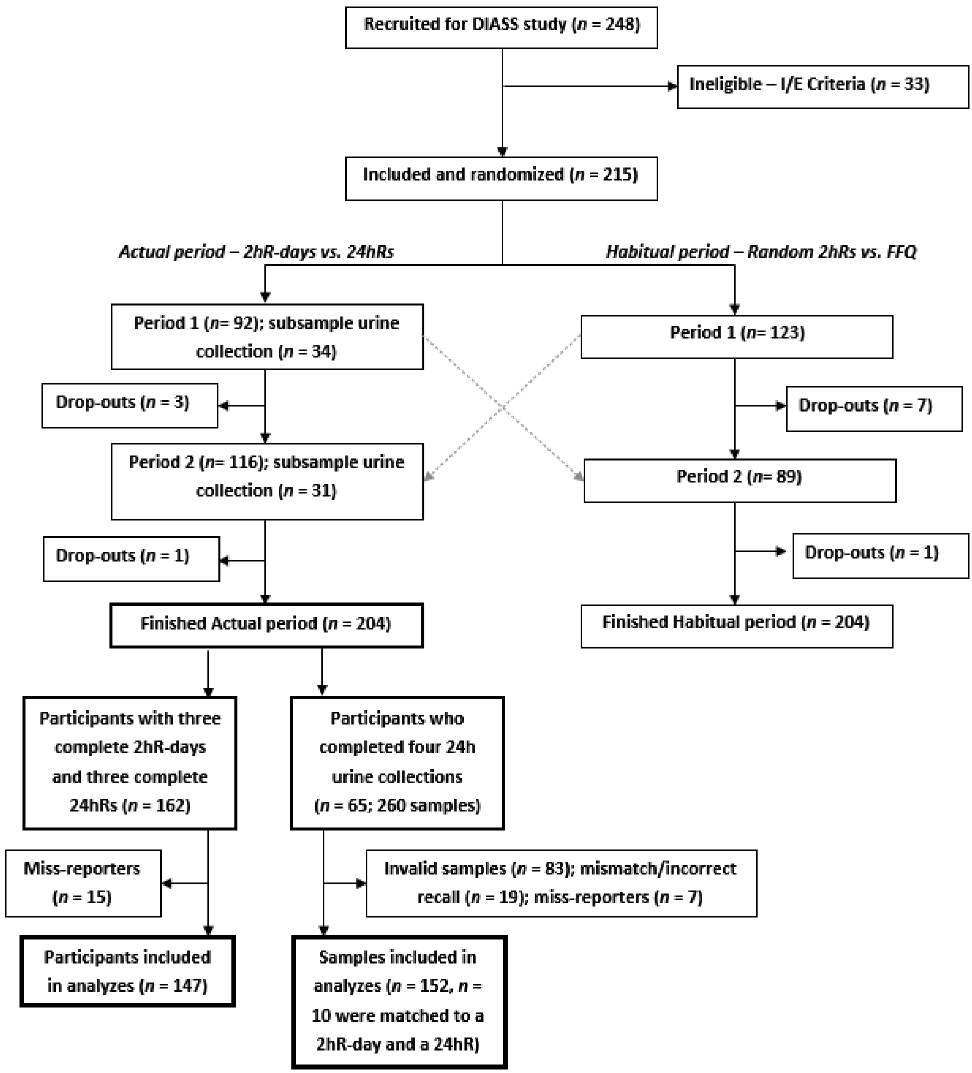
1. Shim, J.S., K. Oh, and H.C. Kim, Dietary assessment methods in epidemiologic studies. Epidemiol Health, 2014. 36: p. e2014009.
2. Subar, A.F., et al., Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. J Nutr, 2015. 145(12): p. 2639-45.
3. Thompson, F.E. and A.F. Subar, Dietary Assessment Methodology, in Nutrition in the Prevention and Treatment of Disease, A.M. Coulston, et al., Editors. 2017, Elsevier Inc. p. 5-48.
4. Eldridge, A.L., et al., Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients, 2018. 11(1).
5. Brouwer-Brolsma, E.M., et al. Dietary intake assessment: From traditional paper-pencil questionnaires to technology-based tools. in International Symposium on Environmental Software Systems. 2020. Springer.
6. Sharp, D.B. and M. Allman-Farinelli, Feasibility and validity of mobile phones to assess dietary intake. Nutrition, 2014. 30(11-12): p. 1257-66.
7. Stone, A., et al., The science of real-time data capture: Self-reports in health research. 2007: Oxford University Press.
8. Shiffman, S., A.A. Stone, and M.R. Hufford, Ecological momentary assessment. Annu Rev Clin Psychol, 2008. 4(1): p. 1-32.
9. Bucher Della Torre, S., et al., Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. Nutrients, 2017. 9(1).
10. Carter, M.C., et al., 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr, 2013. 109(3): p. 539-46.
11. Wellard-Cole, L., 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, 2019. 11(3).
12. Rangan, A.M., 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, 2015. 3(4): p. e98.
13. Ambrosini, G.L., et al., Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr J, 2018. 17(1): p. 5.
14. Lucassen, D.A., et al., Iterative Development of an Innovative Smartphone-Based Dietary Assessment Tool: Traqq. JoVE, 2021(169): p. e62032.

15. Lucassen, D.A., et al., DIetary ASSessment (DIASS) Study: Design of an Evaluation Study to Assess Validity, Usability and Perceived Burden of an Innovative Dietary Assessment Methodology. Nutrients, 2022. 14(6): p. 1156.
16. RIVM. NEVO-online (version 2016/5.0). NEVO-online (2016/5.0). 2016 [cited 2018 15 August]; Available from: https://nevo-online.rivm.nl/.
17. Meijboom, S., et al., Evaluation of dietary intake assessed by the Dutch self-administered web-based dietary 24-h recall tool (Compl-eat™) against interviewer-administered telephone-based 24-h recalls. Journal of Nutritional Science, 2017. 6.
18. Cade, J.E., Measuring diet in the 21st century: use of new technologies. Proc Nutr Soc, 2017. 76(3): p. 276-282.
19. Conway, J.M., et al., Effectiveness of the US Department of Agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. Am J Clin Nutr, 2003. 77(5): p. 1171-8.
20. Subar, A.F., et al., 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, 2013. 67(8): p. 863-7.
21. Lassale, C., et al., Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. Br J Nutr, 2015. 113(6): p. 953-62.
22. Bingham, S.A., Urine nitrogen as a biomarker for the validation of dietary protein intake. The Journal of nutrition, 2003. 133(3): p. 921S-924S.
23. Hills, A.P., N. Mokhtar, and N.M. Byrne, Assessment of physical activity and energy expenditure: an overview of objective measures. Front Nutr, 2014. 1: p. 5.
24. Troiano, R.P., et al., Physical activity in the United States measured by accelerometer. Medicine and science in sports and exercise, 2008. 40(1): p. 181.
25. Joint FAO/WHO/UNU Expert Consultation, Energy and protein requirements: report of a Joint FAO/WHO/UNU Expert Consultation [held in Rome from 5 to 17 October 1981], in Technical Report Series, World Health Organization, Editor. 1985, World Health Organization: Geneva,.
26. Jenab, M., et al., Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet, 2009. 125(5-6): p. 507-25.
27. Henry, C.J.K., Basal metabolic rate studies in humans: measurement and development of new equations. Public Health Nutrition, 2005. 8(7a).
28. Black, A.E., 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. International Journal of Obesity, 2000. 24(9): p. 1119-1130.
29. Lombard, M.J., et al., Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J, 2015. 14: p. 40.

30. Van Rossum, C., et al., The diet of the Dutch: Results of the Dutch National Food Consumption Survey 2012-2016. 2020.
31. Willett, W., Nutritional Epidemiology. 3rd ed. 2013, New York,: Oxford University Press.
32. Kirkpatrick, S.I., et al., Best Practices for Conducting and Interpreting Studies to Validate Self-Report Dietary Assessment Methods. J Acad Nutr Diet, 2019. 119(11): p. 1801-1816.
33. Freedman, L.S., et al., Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. Am J Epidemiol, 2014. 180(2): p. 172-88.
34. Zhang, L., et al., A Systematic Review and Meta-Analysis of Validation Studies Performed on Dietary Record Apps. Adv Nutr, 2021. 12(6): p. 2321-2332.
35. Ocke, M., et al., Relative Validity of MijnEetmeter: A Food Diary App for Self-Monitoring of Dietary Intake. Nutrients, 2021. 13(4).
36. Lemacks, J.L., K. Adams, and A. Lovetere, Dietary Intake Reporting Accuracy of the Bridge2U Mobile Application Food Log Compared to Control Meal and Dietary Recall Methods. Nutrients, 2019. 11(1).

SUPPLEMENTAL MATERIAL

Supplemental Figure 1. Flowchart participant selection from the DIASS population.



Chapter 6

Evaluation of the smartphone-based dietary assessment tool 'Traqq' for assessing habitual dietary intake by random 2-hour recalls in adults: comparison with FFQ and blood concentration biomarkers

Desiree A. Lucassen
Elske M. Brouwer-Brolsma
Hendriek C. Boshuizen
Michiel Balvers
Edith J.M. Feskens

In preparation

ABSTRACT

INTRODUCTION

Accurate dietary assessment is essential in nutritional epidemiology for studying diet-disease relations. These studies mainly rely on food frequency questionnaires (FFQ) and/or repeated 24-hour recalls (24hRs) for assessing habitual dietary intake. Yet, these methods are affected by memory-related bias and heavily burden respondents. To reduce respondent burden and associated measurement error, we recently developed an innovative smartphone-based 2-hour recall (2hR) method. In this study, we evaluated the use of random 2hRs for assessing habitual intake against a validated FFQ and blood concentration markers.

METHODS

Dietary intake was assessed in 215 Dutch adults by a number of 2hRs on randomly selected days and times (i.e., equivalent to three full days of 2hRs) over a four-week period. At the end of the study period, participants completed an FFQ. Sixty-five participants also provided two fasting blood samples, to assess plasma carotenoid and plasma n-3 fatty acid concentrations.

RESULTS

Intake estimates of energy (2,132±665 kcal vs. 2,017±572 kcal) and most nutrients (e.g., protein: 79±29 g vs. 74±22 g; fat: 85±34 g vs. 82±29 g; carbohydrates: 233±75 g vs. 221±65 g) were slightly higher with 2hRs than FFQ. Spearman correlations between 2hRs and FFQ ranged from 0.33 to 0.69 for energy and macronutrients and from 0.32 to 0.58 for micronutrients. For all nutrients, ≥72% of the participants classified in the same/adjacent quartile. Spearman correlations between 2hRs and plasma carotenoids and n-3 fatty acids ranged between 0.34 and 0.57 and cross-classification ranged between 61% to 83% in the same/adjacent quartile.

CONCLUSION

Comparing 2hRs with FFQ and blood concentration markers showed good ranking ability for energy, most nutrients, and most frequently consumed foods. More variation was seen for episodically consumed foods and nutrients. Sampling scheme optimalization will probably result in the capturing of episodically consumed foods and, thus, provide a complete estimate of habitual dietary intake.

INTRODUCTION

Adequate assessment of dietary intake is an essential part of nutritional epidemiology, not only to study diet-disease relations, but also for nutritional surveillance to evaluate a (sub)population's nutritional status [1]. In many of such studies, food frequency questionnaires (FFQ) or repeated 24-hour recalls (24hR) are the method of choice to assess habitual dietary intake [1, 2].

FFQ is a closed method, specifically developed to assess long-term habitual dietary intake. A FFQ consists of a list of foods and drinks and asks the respondent to report frequency of consumption for each of these items [3]. In large studies, FFQ is often the preferred method as they are relatively easy and inexpensive to process [1, 2]. However, respondents often perceive completing the FFQ as extremely burdensome [4]. To illustrate, an extensive FFQ assessing intake of energy, macro-, and most micronutrients takes 45-60 minutes to complete [3]. This often results in incomplete questionnaires. In addition, questions at the end of an FFQ are more likely to be affected by measurement error compared to questions in the beginning of an FFQ [5].

Repeated 24hRs are often used as an alternative method to assess habitual dietary intake. In contrast to FFQ, 24hR is an open method in which a participant is asked to recall all foods and drinks consumed during the previous 24-hours [1-3]. Data of three 24hRs can be used to gain insight in the habitual intake of frequently consumed foods, whereas more than three days are needed to capture the day-to-day variation of a variety of nutrients and foods that are episodically consumed (e.g., vitamin A, vitamin C, fish) [6]. Although repeated 24hRs result in a more detailed account of habitual dietary intake, they are also more difficult to process and, therefore, more expensive. Also for the respondents, the burden remains high as completing one 24hR takes approximately 30-45 minutes [2, 3]. Moreover, both FFQ and 24hR, heavily depend on the respondents ability to correctly recall food intake of the past 24-hours (24hR) or month (FFQ) [1, 3].

Novel technologies are explored and integrated to improve dietary assessment methods, striving to reduce respondent's burden, recall bias, and to improve accuracy. For FFQ and 24hRs, this mostly resulted in web-based applications. In these applications the reporting process is usually standardized, for instance by facilitating self-reporting by the respondent, integration of a fixed food-list linked to a reliable food composition database, which in turn allows automatic coding and calculation of dietary intake [1, 7]. Although these applications significantly improved the accuracy of dietary assessment and lowered researcher burden, issues related to reliance on self-reported data remain (i.e., (un)intentional misreporting of food intake) [2, 7].

To reduce respondent burden and associated measurement error, we recently developed a smartphone-based 2-hour recall (2hR) method [8, 9]. The 2hR method is based on the ecological momentary assessment (EMA) principle; repeated real-time assessment of individual's behaviour 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 [10]. In a recent study, we already showed that the use of repeated 2hRs on one day results in a more accurate account of actual dietary intake data as compared to traditional 24hRs (Lucassen, Brouwer-Brolsma, Boshuizen, et al., *submitted for publication*). In the current study, we explored the use of randomly assessed 2hRs over a longer period of time for assessing habitual dietary intake (i.e., by randomly distributing the equivalent of three full 2hR-days over a period of four weeks). If proven effective, this approach has major potential for improving the accuracy of collected dietary intake data in nutritional epidemiology as 2hRs minimizes the reliance on memory and take little time to complete.

Therefore, in this paper we describe the validation of the collected dietary intake data using the EMA principle. We compared the use of repeated random 2hRs for (near) real-time monitoring of habitual food intake, i.e., energy, macro/micronutrients, food groups, to a traditional validated FFQ and blood concentration markers.

METHODS

PARTICIPANT SELECTION AND STUDY DESIGN

The DIASS study included 215 participants aged 18-70 years and was conducted between June 2019 and May 2020. Participants were eligible for participation if they were able to speak and read Dutch, metabolically stable (i.e., gained or lost ≤3 kg in the past 3 months), willing to maintain their dietary habits for the duration of the study, and in possession of a smartphone with internet plan. The DIASS study had a cross-over design with two study periods; one study period focused on *actual* intake (i.e., 2hR-days vs. 24hRs) and one on *habitual* intake (i.e., random 2hRs vs. FFQ). More details on the DIASS study can be found elsewhere [9].

The present study describes the data of the habitual study period. Food intake was assessed on randomly selected days and times over a four-week study period. Participants completed a number of random 2hRs, equivalent to three full days of 2hRs, and an FFQ. A subsample also provided fasting blood samples. Venepunctures were scheduled in agreement with these participants and conducted at the study centre. Included participants completed a full scheme of random 2hRs (i.e., each timeslot three times) and an FFQ (*n*=145; Supplemental Figure 1), and provided two fasting blood samples (*n*=41; subsample). Under- and over-reporters were identified and excluded based on the Willet cut-offs for both methods [11]. Participants with unreliable or incomplete 2hR and/or FFQ data (i.e., men with energy intakes <800 kcal or >4200 kcal; women <500 kcal or >3500 kcal) were excluded (*n*=4). The DIASS study was approved by the ethics committee of Wageningen University and Research (ABR No.: NL69065.081.19) and conducted according the guidelines laid down in the Declaration of Helsinki. Written informed consent was obtained from all participants.

METHODS OF DIETARY ASSESSMENT

2-hour recalls

The smartphone app Traqq® was used for the 2hRs [8]. Habitual intake was measured by repeated random 2hRs over a four-week study period. Within this period, participants received invitations on random days and times to report their food intake of the previous 2 hours (Figure 1). A personalized, automated 2hR sampling scheme was created in Traqq for each participant. First, a scheme of consecutive 2hRs was created that covered a full day (i.e., on average eight 2hRs). The full day was completed by another invitation, linked to the final evening 2hR, to report on night-time food intake. Next, the timeslots in this full-day scheme were randomly distributed over a four week period; where each time slot was assessed three times to ensue most optimal coverage of the habitual intake (Figure 1). This approach was based on the traditional approach of using three 24hRs to assess habitual intake [6].

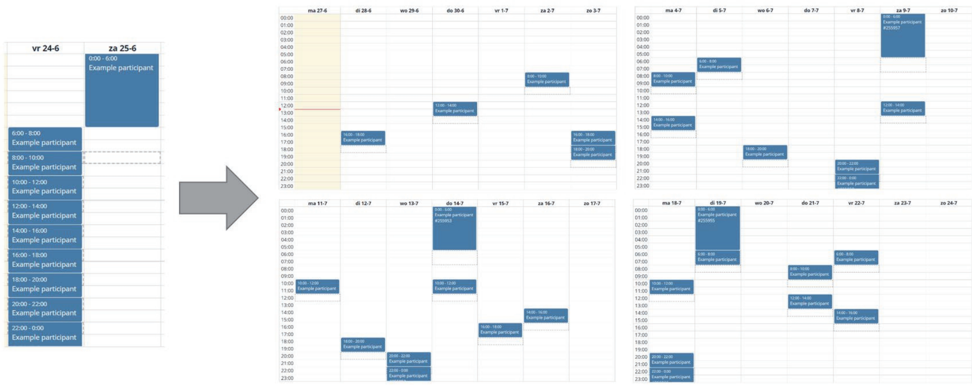


Figure 1. Example random 2hR sampling scheme (right), based on full days of 2hRs (left).

2hRs were restricted to a maximum of two per day to limit the number of recordings on one day where the night-time recall was linked to the preceding final evening 2hR. For all recalls participants had a 1-hour response deadline. Participant's sleeping pattern was taken into account (i.e., inquired in the baseline questionnaire) to avoid disturbing participants while they were sleeping. In general, no invitations were send after 22:00. In case of non-response, the 2hR closed, and a new invitation was automatically rescheduled for the same time on a different day.

Notifications were send at the end of each 2-hour interval with a reporting window of one hour (e.g., interval 6:00-8:00; notification at 8:00; reporting deadline at 9:00). The morning after a final evening 2hR, another invitation was send to report on potential night-time food intake (e.g., night-time interval 22:00-6:00; notification at 8:00; reporting deadline at 9:00). Participant reported their food intake by clicking on the notification or opening the app. Thereafter, the search screen opened and food items could be selected from an extensive food list based on the Dutch Food Composition Database. Subsequently, participants were prompted to report quantity and eating occasion, i.e., breakfast, lunch, dinner, snack. Quantity could 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 could 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 could also be used to create frequently consumed product combinations (e.g., daily breakfast products), which simplified reporting these items and decreased (mis)calculation errors.

Data from the 2hRs was entered in the computation module of Compl-eat™ [12]. Total intakes of energy and nutrients were calculated using the Dutch Food Composition Database 2016 [13]. Data was thoroughly checked by well-trained dieticians according to a standardized protocol, particularly focusing on unusual amounts. Unusual amounts were corrected using standard portion sizes and recipes (e.g., 150 cups of coffee was corrected to 1 cup of 150 g). Habitual intake was calculated per participant by adding up reported intake of all the 2hRs and dividing them by 3 (i.e., as each time point was assessed three times).

Food frequency questionnaire

At the end of the four-week study period, participants were asked to complete a 183-item semi-quantitative FFQ, with a reference period of four weeks. This extensive FFQ was validated for energy, macronutrients, and a number of vitamins [14-16], and administered online with the self-administered Dutch FFQ-tool™ [17]. Participants indicated the frequency of consumed food items by selecting answers ranging from 'not consumed' to '7 days per week'. In addition, portion sizes were estimated using natural portions and commonly used household measures. Energy and nutrient contents of foods were based on the Dutch Food Composition Database 2010 [18] and multiplied by the portion size and frequency of consumption to calculate mean daily intake of energy, macro-, and micronutrients. In addition, average daily intake (in gram) of food items were calculated by multiplying frequency of consumption by portion size. Trained dieticians conducted multiple quality checks to safeguard the quality of the data.

BLOOD COLLECTION

Following a 10-hour overnight fast, blood was drawn from an antecubital vein using venepuncture. Venepunctures were conducted at the study centre, by experienced staff members. Blood was immediately centrifuged, and plasma was stored at −80°C until further analyses. Plasma carotenoids and plasma fatty acids were available for 65 participants, of which fatty acids were determined in replicate blood samples of all participants, and carotenoids were determined in replicate blood samples of 13 participants. Blood samples were used to assess carotenoid and n-3 fatty acid concentrations, to estimate habitual intake of fruit and vegetables and fish, respectively [19-21].

Plasma carotenoids, including α- and β-carotene, β-cryptoxanthin, lutein, and zeaxanthin, were determined using ultra-pressure liquid chromatography coupled to diode array detection (UPLC-DAD). In short, 375 µL plasma was denatured using ethanol in the presence of the internal standard retinyl acetate. The carotenoids were subsequently extracted using 0.01% w/v butylated hydroxytoluene in hexane. The extracts were dried under nitrogen in a TurboVap evaporator (Biotage, Uppsala, Sweden) at 35 °C, reconstituted in 150 µL acetonitrile, and transferred to a LC vial. LC analysis was performed on a Acquity H-class UPLC coupled to an Acquity PDA eLambda detector (Waters, Etten-Leur, the Netherlands) using an Acquity UPLC HSS T3 column (2.1 x 150 mm, 1.8 µm; Waters). Gradient elution was performed using a mixture of acetonitrile-dichloromethane-methanol (ratio 85:5:10) containing 0.1% ammonium acetate (eluens A) and ULC-MS grade water (eluens B) with a constant flow of 400 µL/min and a runtime of 30 min. The detector was set at 292, 325 and 450 nm. Concentrations were calculated using a 6-point calibration curve. A pooled plasma sample was analysed in duplicate in each analytical batch to monitor the quality of the analyses. The interbatch CV values ranged between 4 and 11% for all compounds.

Plasma fatty acid profiles in cholesterol esters, including n-3 fatty acids, were determined using gas chromatography with flame ionization detection (GC-FID). In short, the lipid fraction was extracted from 650 µL plasma using 2-propanol and n-octane. The octane layer was evaporated to dryness and reconstituted in a 1:39 mixture of diethylether-hexane. Lipid fractions were separated using a 500 mg SPE column (Sopachem, Ede, the Netherlands) which allowed to isolate the cholesterol-bound lipid fraction. Lipids were de-esterified in sulfuric acid in methanol for one hour at 90°C, after which MQ water and hexane were added. After shaking the samples, the hexane layer was used for GC analysis, which was performed on an 8690 GC system (Agilent, Amstelveen, the Netherlands). Separation of fatty acid methyl esters was performed on a CP WAX 58 CB column (25 m, i.d. 0.25 mm; Chrompack). The GC oven temperature was set to increase from 60°C to 245°C over a runtime of 45 minutes.

OTHER VARIABLES

General participant characteristics (i.e., age, sex, educational level, daytime activities, sleeping pattern, intention to maintain current body weight) was acquired with a standardized questionnaire. Height was measured without shoes using a stadiometer (SECA 213; SECO Corp., Hamburg, Germany) and weight was assessed without shoes, heavy clothing and empty pockets on a digital scale (SECA 877; SECA Corp., Hamburg, Germany). BMI was calculated as weight/height^2. At the end of the study period, participants were asked to indicate which dietary assessment method they preferred (i.e., random 2hRs or FFQ).

STATISTICAL ANALYSES

Results are presented as means with standard deviations (mean±SD) and frequencies with percentages (n (%)). Macronutrients and alcohol were additionally expressed in energy densities to adjust for energy. To evaluate the random 2hRs against the FFQ for habitual intake of energy, nutrients and food groups multiple analyses were performed [22]. First, absolute intake differences between methods were calculated and expressed as group-level bias ((mean intake 2hRs) / (mean intake FFQ) * 100 – 100). A group-level bias of ≤10% was classified as acceptable (i.e., indication of a relatively similar mean intake) [22]. Second, absolute differences between the 2hRs and the FFQ were evaluated using paired t-tests. Third, Spearman correlation coefficients were calculated to assess the strength and direction of the association between the methods [23]. Correlation coefficients of <0.20 were regarded as poor, 0.20-0.49 as acceptable, and ≥0.50 as good [22]. The ranking ability of the random 2hRs was assessed by dividing the intake of nutrients and foods as assessed by the 2hRs and FFQ over quartiles after which we examined whether persons were ranked into the same, adjacent or extreme quartile. Classification of ≥50% of the participants in the same quartile [22], ≥75% in the same or adjacent quartile [24], and <10% in the extreme quartile [22], was considered a good outcome. Presence, direction and extent of bias at group level for total energy intake was visualized by plotting the difference between 2hRs and FFQ against the mean of the two methods [25]. All analyses were performed using IBM SPSS Statistics version 28.0 (SPSS Inc. Chicago, IL, USA).

RESULTS

Participants had a mean±SD age of 38.5±18.4 years, were mainly women (74%), highly educated (56%), normal weight (74%; BMI <25 kg/m^2), and did not follow a diet regimen (73%). Overall, 77% of the participants preferred random 2hRs over an extensive FFQ (Table 1).

Table 1. General characteristics of the participants included in this validation study

	Total (141)	Men (37)	Women (104)
Mean age, years (SD)	38.5 (18.4)	42.0 (19.1)	37.3 (18.1)
Age category (n, (%))			
<25 years	58 (41)	14 (38)	44 (42)
25-50 years	32 (23)	6 (16)	26 (25)
>50 years	51 (36)	17 (46)	34 (33)
Mean BMI, kg/m^2 (SD)	23.8 (4.0)	24.7 (4.2)	23.5 (3.9)
BMI category (n, (%))			
<25 kg/m^2	104 (74)	24 (65)	80 (77)
25.0-29.9 kg/m^2	25 (18)	10 (27)	15 (14)
≥30 kg/m^2	12 (8)	3 (8)	9 (9)
Educational level (n, (%))			
Low[1]	4 (3)	0 (0.0)	4 (4)
Intermediate[2]	58 (41)	14 (38)	44 (42)
High[3]	79 (56)	23 (62)	56 (54)
Diet regimen (n, (%))			
Yes, always	21 (15)	1 (3)	20 (19)
Yes, sometimes	17 (12)	4 (11)	13 (13)
Never	103 (73)	32 (86)	71 (68)
Preferred method (n, (%))			
Random 2hRs	108 (77)	27 (73)	81 (78)
FFQ	26 (18)	8 (22)	18 (17)
Unknown[4]	7 (5)	2 (5)	5 (5)

[1] Primary or lower education.

[2] Secondary or higher vocational education.

[3] University or college.

[4] Answering was not compulsory.

Estimated intakes of energy and most macronutrients were higher with the 2hRs than with the FFQ, with the exception of fat (en%), PUFA, ALA, EPA, DPA, and alcohol (en%) (Table 2). Yet, differences in estimated energy and most macronutrient intakes were small (group-level bias ≤10%). Group-level bias was >10% for animal protein (13%), EPA (-16%), DHA (-23%) and alcohol (en%, -12%). A larger variation was found for the micronutrients with percentual differences ranging between -27% (β-carotene) and 12% (vitamin B1).

Spearman correlation coefficients between the 2hRs and FFQ ranged from acceptable to good for both energy and macronutrients (range 0.33 – 0.69) and for micronutrients (range 0.32 –

0.58). Moreover, the 2hRs classified ≥75% of the participants in the same or adjacent quartile as the FFQ except for DHA (72%), β-carotene (74%), and lycopene (73%). Misclassification in the extreme quartile was below 10% for energy and all nutrients (range 0 – 8%). The Bland-Altman plot shows an increasing difference between 2hRs and FFQ with increasing energy intake (β = 0.26, p = <0.01) (Figure 2).

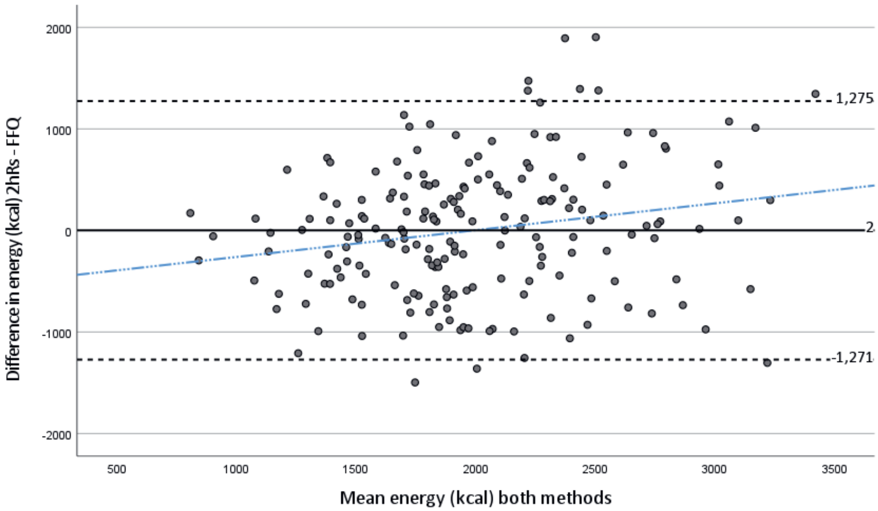


Figure 2. Bland-Altman plot of the differences in energy intake (kcal) estimated with random 2hRs and FFQ, plotted against the mean of the both methods (g/d). Mean difference (solid line), 95% limits of agreement (1.96xSD of mean difference; dashed line), and linear regression line (blue dashed line) are included.

For food groups, a larger variation in intake estimates was found ranging from -65% (legumes) to +62% (pastry, cake and biscuits) (Table 3). Group-level bias was relatively small (≤10%) for ‘alcoholic beverages’, ‘dairy’, ‘eggs’, ‘fruit’, ‘potatoes’, ‘soups’, and ‘vegetables’. For the remaining food groups, group-level bias exceeded 10%, which was also supported by mostly statistically significant paired-t-tests for these groups. Higher intakes were estimated by 2hRs than FFQ for the food groups ‘bread’, ‘cheese’, ‘dairy’, ‘eggs’, ‘fruit’, ‘non-alcoholic beverages’, ‘pastry, cake and biscuits’, ‘savoury sandwich fillings’, ‘soups’, ‘sugar and confectionery’, and ‘vegetables’. For the remaining food groups, intake estimates were higher by FFQ than 2hRs.

Spearman correlation coefficients for most food groups were acceptable to good (range 0.27 – 0.67). Except for the correlation for ‘soups’, which was poor (0.14). For the majority of the food groups, ≥75% of the participants were classified in the same or adjacent quartile except for ‘fats, oils and savoury sauces’ (72%) and ‘potatoes’ (70%). Similar as to the nutrients, misclassification in the extreme quartile was below 10% (range 0 – 9%).

6

Table 2. Energy and nutrient intakes assessed by random 2hRs and FFQ with corresponding group-level bias, paired-t-tests, Spearman correlation coefficients between the 2hRs and FFQ, and corresponding cross-classification (n = 141).

	2hRs	FFQ	Group-level bias (%)[1]	P[2]	Correlation coefficient[3] (95% CI)	Cross-classification by quartiles		
	Mean ± SD	Mean ± SD				Same (%)	Same + adjacent (%)	Extreme (%)
Energy (kcal)	2,132 ± 665	2,017 ± 572	5.7	0.04	0.41 (0.26, 0.54)	39	78	6
Protein (en%)	15 ± 3	15 ± 2	0.5	0.74	0.48 (0.33, 0.60)	37	79	3
Protein (g)	79 ± 29	74 ± 22	7.2	0.02	0.48 (0.33, 0.60)	39	78	4
Plant-based protein (g)	37 ± 14	37 ± 13	1.5	0.60	0.59 (0.46, 0.70)	45	84	1
Animal protein (g)	42 ± 23	37 ± 17	13.2	<0.01	0.63 (0.51, 0.73)	48	84	0
Fat (en%)	35 ± 6	36 ± 5	-2.6	0.07	0.34 (0.18, 0.48)	36	77	6
Fat (g)	85 ± 34	82 ± 29	3.0	0.43	0.36 (0.20, 0.50)	37	76	8
SFA (g)	30 ± 12	28 ± 11	8.6	0.03	0.45 (0.30, 0.58)	40	78	5
MUFA (g)	30 ± 15	30 ± 11	0.2	0.97	0.37 (0.21, 0.51)	35	77	6
PUFA (g)	17 ± 8	17 ± 7	-4.7	0.27	0.40 (0.25, 0.53)	39	75	6
EPA (g)	0.08 ± 0.15	0.10 ± 0.09	-15.9	0.24	0.43 (0.28, 0.56)	36	77	3
DHA (g)	0.10 ± 0.25	0.13 ± 0.14	-23.0	0.18	0.39 (0.23, 0.53)	37	72	6
Cholesterol (mg)	222 ± 141	220 ± 110	1.3	0.78	0.41 (0.26, 0.54)	33	76	6
Carbohydrates (en%)	44 ± 7	44 ± 6	0.5	0.67	0.50 (0.36, 0.62)	40	79	4
Carbohydrates (g)	233 ± 75	221 ± 65	5.3	0.04	0.49 (0.34, 0.61)	43	77	5
Mono/disaccharides (g)	101 ± 34	94 ± 29	7.6	<0.01	0.54 (0.40, 0.65)	41	83	4
Polysaccharides (g)	131 ± 51	127 ± 46	3.7	0.26	0.51 (0.37, 0.63)	43	80	4
Fibre (g)	25 ± 10	24 ± 8	0.5	0.86	0.58 (0.45, 0.69)	45	85	2
Alcohol (en%)	3.0 ± 4.6	3.4 ± 3.7	-12.0	0.12	0.69 (0.57, 0.77)	52	84	3
Alcohol (g)	10 ± 15	10 ± 10	0.3	0.97	0.69 (0.58, 0.78)	45	87	3
Ca (mg)	988 ± 388	935 ± 345	5.7	0.08	0.58 (0.45, 0.69)	45	83	1
Fe (mg)	12 ± 4	11 ± 3	3.0	0.28	0.43 (0.28, 0.56)	38	75	3
β-Carotene (µg)	2,998 ± 2,891	4,102 ± 2,606	-26.9	<0.001	0.33 (0.17, 0.47)	34	74	6
Lycopene (µg)	2,934 ± 4,178	3,020 ± 3,084	-2.8	0.75	0.32 (0.16, 0.46)	32	73	7
Vitamin B1 (mg)	1.01 ± 0.42	0.90 ± 0.27	11.6	<0.001	0.48 (0.33, 0.60)	44	78	4

Vitamin B2 (mg)	1.43 ± 0.55	1.35 ± 0.44	6.0	0.06	0.52 (0.38, 0.64)	41	79	1
Vitamin B6 (mg)	1.36 ± 0.51	1.50 ± 0.44	-9.0	<0.01	0.42 (0.27, 0.55)	36	78	4
Vitamin B12 (µg)	3.92 ± 2.87	4.11 ± 2.29	-4.5	0.45	0.58 (0.45, 0.69)	37	84	1
Vitamin C (mg)	99 ± 59	90 ± 37	9.6	0.06	0.51 (0.37, 0.63)	43	82	5
Vitamin D (µg)	2.34 ± 1.74	2.97 ± 1.58	-21.2	<0.001	0.48 (0.33, 0.60)	33	82	4
Vitamin E (mg)	13 ± 6	14 ± 5	-8.3	0.02	0.40 (0.25, 0.53)	43	76	6
Folate (µg)	270 ± 99	253 ± 74	6.4	0.03	0.48 (0.33, 0.60)	43	78	6
Folate equivalents (µg)	273 ± 98	272 ± 83	0.4	0.88	0.48 (0.33, 0.60)	38	81	4

[1] Group-level bias = (mean 2hRs) / (mean FFQ) x 100 - 100.

[2] Paired-t-test between mean intake assessed with 2hRs and 24hRs.

[3] Spearman correlation.

6

Table 3. Intake of food groups (g/d) assessed by random 2hRs and FFQ with corresponding group-level bias, paired-t-tests, Spearman correlation coefficients between the 2hRs and FFQ, and corresponding cross-classification (n = 141).

	2hRs	FFQ	Group-level bias (%)[1]	P [2]	Correlation coefficient [3] (95% CI)	Cross-classification by quartiles		
	Mean ± SD	Mean ± SD				Same (%)	Same + adjacent (%)	Extreme (%)
Alcoholic beverages	137 ± 241	140 ± 169	-2.2	0.83	0.65 (0.54, 0.75)	51	81	4
Bread	130 ± 75	117 ± 70	10.7	0.04	0.47 (0.33, 0.60)	41	83	6
Cheese	31 ± 23	25 ± 22	22.0	<0.01	0.51 (0.37, 0.63)	36	80	3
Dairy	245 ± 193	223 ± 179	9.7	0.12	0.65 (0.53, 0.74)	48	90	2
Eggs	21 ± 31	21 ± 22	0.5	0.96	0.52 (0.38, 0.64)	43	79	6
Fats, oils, savoury sauces	27 ± 24	44 ± 26	-39.3	<0.001	0.32 (0.16, 0.46)	33	72	5
Fish	15 ± 28	18 ± 18	-16.7	0.30	0.35 (0.19, 0.49)	28	78	0
Fruit	195 ± 158	181 ± 127	7.8	0.20	0.59 (0.46, 0.70)	48	86	1
Grains and cereals	55 ± 58	86 ± 59	-36.1	<0.001	0.46 (0.31, 0.58)	43	77	4
Legumes	8 ± 18	23 ± 27	-65.3	<0.001	0.29 (0.13, 0.44)	34	77	2
Meats and poultry	55 ± 62	62 ± 51	-11.0	0.16	0.66 (0.54, 0.75)	55	86	0
Non-alcoholic beverages	1795 ± 913	1498 ± 522	20.6	<0.001	0.67 (0.55, 0.76)	50	90	2
Nuts, seeds, snacks	32 ± 34	37 ± 32	-14.5	0.14	0.27 (0.10, 0.41)	35	76	9
Pastry, cake, biscuits	48 ± 36	30 ± 22	62.2	<0.001	0.48 (0.33, 0.60)	34	80	4
Potatoes	41 ± 44	45 ± 31	-8.6	0.34	0.29 (0.13, 0.44)	33	70	8
Savoury sandwich fillings	15 ± 16	13 ± 17	13.8	0.28	0.38 (0.22, 0.52)	38	74	8
Soups	48 ± 68	45 ± 66	6.5	0.70	0.14 (-0.02, 0.30)	31	77	4
Sugar and confectionery	30 ± 26	24 ± 21	25.0	<0.01	0.54 (0.40. 0.65)	38	82	1
Vegetables	178 ± 134	164 ± 107	8.6	0.17	0.36 (0.20, 0.50)	39	76	9
Vegetarian products	27 ± 55	37 ± 74	-27.4	0.03	0.67 (0.55, 0.76)	31	91	0

[1] Group-level bias = (mean 2hRs) / (mean FFQ) x 100 - 100.

[2] Paired-t-test between mean intake assessed with 2hRs and FFQ.

[3] Spearman correlation.

Spearman correlation coefficients between plasma carotenoids and fruit and vegetables intake were acceptable to good for intake reported by 2hRs (range 0.42 – 0.57) (Table 4). For all plasma carotenoid comparisons, correlations were lower, yet still acceptable, for intake reported by FFQ (range 0.20 – 0.39). In contrast, correlation between plasma n-3 fatty acids (EPA and DHA) and fish intake was acceptable for 2hRs but better for FFQ (0.34 vs 0.52, respectively).

For most of the concentration markers and self-reported intakes, ≥75% of the participants were classified in the same or adjacent quartile except for sum of carotenoids and vegetable intake by 2hRs (61%) and FFQ (71%), α-carotene and fruit/vegetable intake by FFQ (71%), β-carotene and fruit/vegetable intake by 2hRs (71%), and β-cryptoxanthin and fruit intake by 2hRs (68%). Misclassification in the extreme quartile was ≤10% for both methods (range 0 – 7%), except for sum of carotenoids and fruit intake by FFQ (10%) and sum of carotenoids and vegetable intake by FFQ (10%).

Table 4. Spearman correlation coefficients and cross-classification of reported intakes of fish, fruit and vegetables in 2hRs and FFQ and related blood concentration biomarkers

	Correlation coefficient[1] (95% CI)	Cross-classification by quartiles		
		Same (%)	Same + adjacent (%)	Extreme (%)
Fruit and vegetable intake and sum of carotenoids				
2hRs	0.57 (0.30, 0.76)	39	83	0
FFQ	0.36 (0.05, 0.61)	34	81	7
Fruit intake and sum of carotenoids				
2hRs	0.49 (0.20, 0.70)	46	78	7
FFQ	0.37 (0.06, 0.62)	44	83	10
Vegetable intake and sum of carotenoids				
2hRs	0.54 (0.26, 0.74)	44	61	7
FFQ	0.20 (-0.12, 0.48)	32	71	10
Fruit and vegetable intake and α-carotene				
2hRs	0.48 (0.19, 0.70)	44	76	2
FFQ	0.34 (0.03, 0.59)	42	71	5
Fruit and vegetable intake and β-carotene				
2hRs	0.44 (0.14, 0.67)	32	71	0
FFQ	0.35 (0.04, 0.60)	39	78	5
Fruit intake and β-cryptoxanthin				
2hRs	0.42 (0.12, 0.65)	29	68	5
FFQ	0.39 (0.08, 0.63)	37	78	7
Vegetable intake and lutein + zeaxanthin				
2hRs	0.42 (0.12, 0.65)	29	76	7
FFQ	0.35 (0.04, 0.60)	29	81	7
Fish intake and EPA + DHA				
2hRs	0.34 (0.03, 0.59)	24	83	0
FFQ	0.52 (0.23, 0.72)	49	83	0

[1] Spearman correlation

DISCUSSION

In this study we explored the accuracy of the EMA-based 2hR methodology for assessing habitual intake by repeated random 2hRs. The 2hRs yielded slightly higher energy and macronutrient intake estimates compared to the validated FFQ. Most importantly, as shown by correlations ≥0.40 and a ranking agreement ≥ 75% (i.e., ranking in the same or adjacent quartile as the FFQ), the ranking ability of the 2hRs were promising for energy, most nutrients, and most frequently consumed foods. Validation against blood concentration markers for carotenoid and n-3 fatty acids showed acceptable to good correlations for both methods. Finally, most participants preferred repeated random 2hRs over one extensive FFQ.

Our results showed higher intake estimates for energy and most macronutrients with 2hRs, except for fat (en%), EPA, DHA and alcohol (en%). The FFQ used in this study is developed to cover at least 96% of absolute intake [14-16, 26], therefore, some underestimation by FFQ was expected. Still, difference between methods were mostly small (group-level bias ≤10%), except for animal protein (13%), EPA (-16%), DHA (-23%), and alcohol (en%, -12%). These results are in line with habitual intake estimates by repeated 24hRs, of which is known that three days is enough to gain insight in the habitual intake of frequently consumed foods [6, 16, 27]. Yet, difference in alcohol estimates was only (slightly) larger for en%, whereas differences in alcohol (g) and alcoholic beverages were very low (<1% and 2%, respectively). The larger differences for EPA and DHA, and also for fish intake, can be attributed to the fact that these are episodically consumed, and are therefore more difficult to capture with the 2hRs. This is also the case for repeated 24hRs [6]. Episodically consumed foods are more difficult to capture with open methods such recalls or records, due to day-to-day variation in intake, whereas an FFQ specifically asks about consumption of these foods. Yet, correlations between 2hRs and FFQ intake estimates of EPA, DHA and fish were similar or higher as compared to results from previous studies investigating correlations between repeated 24hRs and similar Dutch FFQs [16, 27-30]. This is also in line with results from our previous validation study, showing more accurate intake estimates by repeated 2hRs on one day compared to 24hRs (Lucassen, Brouwer-Brolsma, Boshuizen, et al., unpublished results).

As expected, larger differences were found in intake estimates of micronutrients between random 2hRs and FFQ, especially for β-carotene (-27%), vitamin B1 (12%), and vitamin D (21%). Yet, correlations for micronutrients were all acceptable to good, and ranking agreement was ≥ 75% for most micronutrients, except for β-carotene and lycopene who showed only a slightly lower ranking agreement (73% and 74%, respectively). Interesting, a similar study comparing an FFQ with repeated 24hRs, showed much lower correlations for vitamin B1 (0.21) and vitamin D (0.27) than in our study (vitamin B1: 0.48; vitamin D: 0.48) [27]. Although β-carotene was not included, this study did include comparisons of FFQ estimates with plasma carotenoids. As mentioned, plasma carotenoids are considered biological markers for the past intake of fruits and vegetables [20]. When comparing our results to other studies investigating correlations between self-reported fruit and vegetable intake (by FFQ) and plasma carotenoids, we found similar results between our FFQ and plasma carotenoids (range 0.20 – 0.39), and mostly higher correlations between the 2hRs and plasma carotenoids (range 0.42 - 0.57). To illustrate, Sluik and colleagues found correlations between FFQ reported fruit and vegetable intake and plasma levels between 0.24 and 0.43 [27]. A similar trend can be seen when comparing to international studies [20, 31]. In contrast, the correlation between 2hR reported fish intake and plasma EPA/DHA) was lower than the correlation between FFQ reported fish intake and plasma EPA/DHA (i.e., 0.34 vs. 0.52, respectively). Interestingly, correlations between FFQ reported fish intake and plasma EPA/DHA varies over studies (e.g., 0.43 [27], 0.29 [21], 0.17 [32]). These differences can probably be attributed to the large day-to-day variation in fish intake, which not only affects recalls but also seems to affect FFQs.

Regarding food groups, we see a similar trend as for the nutrients, with correlations of ≥0.40 for most frequently consumed foods (i.e., ≥5 days/week according to the Dutch National Food Consumption Survey) and some episodically consumed foods (<5 days/week) [33]. Only for 'fats, oils and savoury sauces', 'fish', 'legumes', 'nuts, seeds and snacks', 'potatoes', 'savoury sandwich fillings', 'soups', and 'vegetables' correlations were lower. With the exception of vegetables, these are mostly episodically consumed foods of which we know that they are difficult to capture with an open method [6, 33]. However, recall days do not automatically have to result in better coverage of episodically consumed foods, due to the infrequent consumption of for instance fish and legumes. Perhaps a better option would be to combine the recalls with a short food propensity questionnaire, which is specifically focused on assessment of episodically consumed food groups [34, 35].

Overall, the correlations between food group intake estimates by 2hR and FFQ were similar, as compared to other studies comparing repeated 24hRs and Dutch FFQ [16, 27, 28]. Although absolute differences in intake are less relevant for nutritional epidemiology, some differences do stand out. Significantly higher intakes were estimated for 'non-alcoholic beverages' by 2hRs than FFQ (1,795±913 g vs. 1,489±522 g; 21% difference). Interestingly, similar results were found for 'non-alcoholic beverages' in our previous study, comparing intake estimated by 2hR-days with 24hRs (i.e., 1,670±788 vs. 1,405±644; 19% difference) (Lucassen, Brouwer-Brolsma, Boshuizen, et al., *submitted for publication*). These results strengthen our previous explanation that reporting of non-alcoholic beverages is easily forgotten as they are often consumed throughout the day, without being linked to a specific eating occasion, which makes it more difficult to recall. Also in line with our previous study, are the significant lower intake estimated for ‘fats, oils and savoury sauces’ with 2hRs than FFQ (27±24 g vs. 44±26 g; 39% difference). It seems that reporting of fat, oils and savoury sauces is more easily forgotten with 2hRs. This can be explained by the fact that FFQ specifically asks about the consumption of these foods, whereas with 24hRs, respondents are prompted by either the tool or the interviewer to report consumption for these items. Adding a prompt to specific eating occasions to report added fats could increase the recall of these items (e.g., butter/margarine on bread for lunch; added fats/oils while cooking diner).

Although we used data from an extensive validation study [9], there are still some methodological issues that warrant attention. We developed our 2hR sampling schemes according to the assumption that three 24hRs are enough to estimate habitual intake. Yet, it is also known that three 24hRs is often not enough to capture the daily variation in episodically consumed foods (and related nutrients) [6]. The effect on intake estimates of these foods can also be seen in our results (on fish and legumes). To enable capturing habitual intake of all foods, a more extensive sampling scheme might be more efficient (e.g., equivalent of 7 days) or to combine the 2hRs with a short food propensity questionnaire. Next, the majority of our sample consisted of highly educated women, potentially limiting generalizability of our results, which is extremely important for use in nutritional epidemiology. Therefore, additional evaluation in a more diverse population is needed before the 2hR method can be adopted in nutritional epidemiology to study diet-disease relations [36]. Still, the current evaluation clearly indicate the potential of using 2hRs to assess habitual dietary intake. Especially as our results are based on multiple statistical tests, which has been suggested as the most optimal approach to evaluate a dietary assessment method [22, 37].

In conclusion, the use of random 2hRs shows great potential as an alternative method for assessing dietary intake. A sampling scheme, based on the equivalent of three full recall days, shows higher intake estimates and good ranking ability for energy, most nutrients, and most frequently consumed foods. More variation was seen for episodically consumed foods and nutrients. A more extensive sampling scheme will probably be able to provide a complete estimate of habitual dietary intake. Yet, more research is warranted to prove these assumptions.

REFERENCES

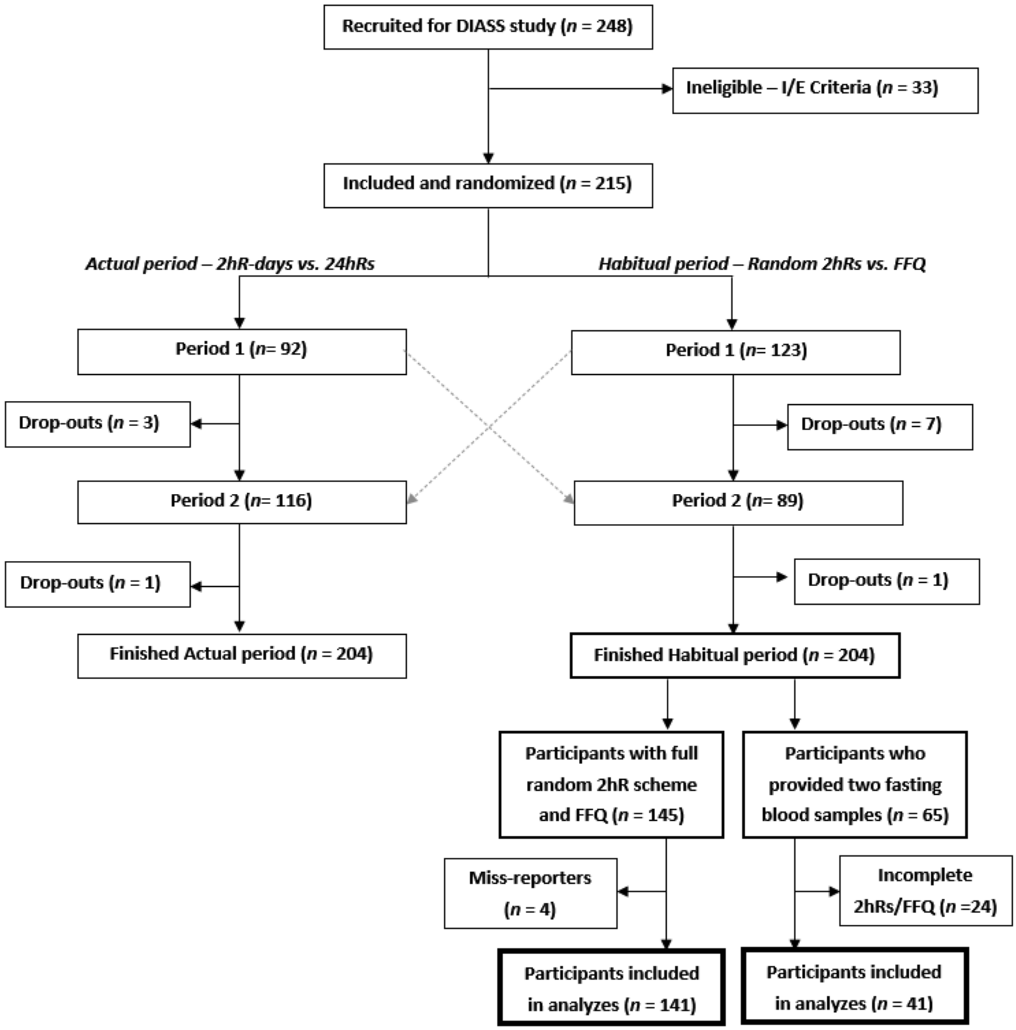
1. Naska, A., A. Lagiou, and P. Lagiou, Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Research, 2017. 6.
2. Shim, J.-S., K. Oh, and H.C. Kim, Dietary assessment methods in epidemiologic studies. Epidemiology and health, 2014. 36.
3. Thompson, F.E. and A.F. Subar, Dietary assessment methodology. Nutrition in the Prevention and Treatment of Disease, 2017: p. 5-48.
4. Cade, J., et al., Development, validation and utilisation of food-frequency questionnaires–a review. Public health nutrition, 2002. 5(4): p. 567-587.
5. Peytchev, A. and E. Peytcheva. Reduction of measurement error due to survey length: Evaluation of the split questionnaire design approach. in Survey Research Methods. 2017.
6. Willett, W., Nutritional Epidemiology. 3rd ed. 2013, New York,: Oxford University Press.
7. Brouwer-Brolsma, E.M., et al. Dietary intake assessment: From traditional paper-pencil questionnaires to technology-based tools. in International Symposium on Environmental Software Systems. 2020. Springer.
8. Lucassen, D.A., et al., Iterative Development of an Innovative Smartphone-Based Dietary Assessment Tool: Traqq. JoVE, 2021(169): p. e62032.
9. Lucassen, D.A., et al., DIetary ASSessment (DIASS) Study: Design of an Evaluation Study to Assess Validity, Usability and Perceived Burden of an Innovative Dietary Assessment Methodology. Nutrients, 2022. 14(6): p. 1156.
10. Shiffman, S., A.A. Stone, and M.R. Hufford, Ecological momentary assessment. Annu. Rev. Clin. Psychol., 2008. 4: p. 1-32.
11. Rhee, J.J., et al., Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies. American journal of epidemiology, 2015. 181(4): p. 225-233.
12. Meijboom, S., et al., Evaluation of dietary intake assessed by the Dutch self-administered web-based dietary 24-h recall tool (Compl-eat™) against interviewer-administered telephone-based 24-h recalls. Journal of Nutritional Science, 2017. 6.
13. RIVM. NEVO-online (version 2016/5.0). NEVO-online (2016/5.0). 2016 [cited 2018 15 August]; Available from: https://nevo-online.rivm.nl/.
14. Feunekes, G.I., et al., Relative and biomarker-based validity of a food-frequency questionnaire estimating intake of fats and cholesterol. The American journal of clinical nutrition, 1993. 58(4): p. 489-496.

15. Siebelink, E., A. Geelen, and J.H. de Vries, Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. Br J Nutr, 2011. 106(2): p. 274-81.
16. Streppel, M.T., et al., Relative validity of the food frequency questionnaire used to assess dietary intake in the Leiden Longevity Study. Nutrition journal, 2013. 12(1): p. 1-8.
17. Molag, M., Towards transparent development of food frequency questionnaires: scientific basis of the Dutch FFQ-TOOL tm: a computer system to generate, apply and process FFQs. 2010.
18. RIVM. NEVO-online (2010/2.0). 2010.
19. Jenab, M., et al., Biomarkers in nutritional epidemiology: applications, needs and new horizons. Human genetics, 2009. 125(5-6): p. 507-525.
20. Al-Delaimy, W., et al., Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). European journal of clinical nutrition, 2005. 59(12): p. 1387-1396.
21. Saadatian-Elahi, M., et al., Plasma phospholipid fatty acid profiles and their association with food intakes: results from a cross-sectional study within the European Prospective Investigation into Cancer and Nutrition. The American journal of clinical nutrition, 2009. 89(1): p. 331-346.
22. Lombard, M.J., et al., Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J, 2015. 14: p. 40.
23. Masson, L.F., et al., Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. Public health nutrition, 2003. 6(3): p. 313-321.
24. Forster, H., et al., Online dietary intake estimation: the Food4Me food frequency questionnaire. Journal of medical Internet research, 2014. 16(6): p. e3105.
25. Bland, J.M. and D. Altman, Statistical methods for assessing agreement between two methods of clinical measurement. The lancet, 1986. 327(8476): p. 307-310.
26. Brouwer-Brolsma, E.M., et al., Development and external validation of the 'Flower-FFQ': a FFQ designed for the Lifelines Cohort Study. Public Health Nutr, 2021: p. 1-12.
27. Sluik, D., et al., A national FFQ for the Netherlands (the FFQ-NL 1.0): validation of a comprehensive FFQ for adults. Br J Nutr, 2016. 116(5): p. 913-23.
28. Ocke, M.C., et al., The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. International journal of epidemiology, 1997. 26(suppl_1): p. S37.

29. Ocke, M.C., et al., The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. International journal of epidemiology, 1997. 26(suppl_1): p. S49.
30. van Dongen, M.C., et al., The Maastricht FFQ: Development and validation of a comprehensive food frequency questionnaire for the Maastricht study. Nutrition, 2019. 62: p. 39-46.
31. Burrows, T.L., et al., Fruit and vegetable intake assessed by food frequency questionnaire and plasma carotenoids: a validation study in adults. Nutrients, 2015. 7(5): p. 3240-51.
32. Welch, A.A., et al., Dietary fish intake and plasma phospholipid n–3 polyunsaturated fatty acid concentrations in men and women in the European Prospective Investigation into Cancer–Norfolk United Kingdom cohort. The American journal of clinical nutrition, 2006. 84(6): p. 1330-1339.
33. Van Rossum, C., et al., The diet of the Dutch: Results of the Dutch National Food Consumption Survey 2012-2016. 2020.
34. Subar, A.F., et al., The food propensity questionnaire: concept, development, and validation for use as a covariate in a model to estimate usual food intake. J Am Diet Assoc, 2006. 106(10): p. 1556-63.
35. Smiliotopoulos, T., E. Magriplis, and A. Zampelas, Validation of a Food Propensity Questionnaire for the Hellenic National Nutrition and Health Survey (HNNHS) and Results on This Population’s Adherence to Key Food-Group Nutritional Guidelines. Nutrients, 2020. 12(6): p. 1808.
36. Illner, A.K., et al., Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol, 2012. 41(4): p. 1187-203.
37. Kirkpatrick, S.I., et al., Best Practices for Conducting and Interpreting Studies to Validate Self-Report Dietary Assessment Methods. J Acad Nutr Diet, 2019. 119(11): p. 1801-1816.

SUPPLEMENTAL MATERIAL

Supplemental Figure 1. Flowchart participant selection from the DIASS population.



6

Chapter 7

Short and long-term innovations on dietary behaviour assessment and coaching: present efforts and vision of the Pride and Prejudice consortium

Desiree A. Lucassen
Marlou P. Lasschuijt
Guido Camps
Ellen J. Van Loo
Arnout R.H. Fischer
Roelof A.J. de Vries
Juliet A.M. Haarman
Monique Simons
Emely de Vet
Marina Bos-de Vos
Sibo Pan
Xipei Ren
Kees de Graaf
Yuan Lu
Edith J.M. Feskens
Elske M. Brouwer-Brolsma

International Journal of Environmental Research and Public Health 2021; 18(15): 7877

ABSTRACT

Overweight, obesity and cardiometabolic diseases are major global health concerns. Lifestyle factors, including diet, have been acknowledged to play a key role in the solution of these health risks. However, as shown by numerous studies, and in clinical practice, it is extremely challenging to quantify dietary behaviours as well as influencing them via dietary interventions. As shown by the limited success of 'one-size-fits-all' nutritional campaigns catered to an entire population or subpopulation, the need for more personalized coaching approaches is evident. New technology-based innovations provide opportunities to further improve the accuracy of dietary assessment and develop approaches to coach individuals towards healthier dietary behaviours. Pride & Prejudice (P&P) is a unique multi-disciplinary consortium consisting of researchers in life, nutrition, ICT, design, behavioural and social sciences from all four Dutch Universities of Technology. P&P focuses on the development and integration of innovative technological techniques such as artificial intelligence (AI), machine learning, conversational agents, behaviour change theory and personalized coaching to improve current practices and establish lasting dietary behaviour change.

INTRODUCTION

Globally, poor diet quality is acknowledged to significantly impact health, and associated health care costs. The Global Burden of Disease study even indicated that a suboptimal diet is the second-leading risk factor for disability-adjusted life years and deaths worldwide after tobacco [1].

Despite the importance of a healthy diet, nutrition guidelines set by the Dutch Health Council are often not met by Dutch consumers. To illustrate, the Dutch Health Council (2015) recommends daily intakes of ≥250 g of vegetables and ≥200 g of fruits, one portion of fatty fish per week (17 g/day) and to limit intake of red meat and processed meat (<43 g/day) [2]. According to the Dutch Food Consumption Survey 2012–2016, a substantial proportion of Dutch women and men (19–50 year) do not adhere to these recommendations, i.e., shown by self-reported mean daily intakes of 128 g and 131 g of vegetables, 115 g and 97 g of fruits, 14 g and 14 g of fish, along with 88 g and 131 g of meat (products), respectively [3]. Meeting healthy diet requirements is implicated in benefiting long-term health, i.e., a lower risk of overweight, obesity, coronary heart disease, stroke, diabetes, and colon and lung cancer [2]. Moreover, in order to prevent overweight and obesity, which often precede the above listed adverse health outcomes, food overconsumption (i.e., energy intake exceeding energy expenditure) should also be prevented [4]. In 2019, more than 50% of Dutch adults were overweight (volksgezondheidenzorg.info, accessed on 26 September 2020). This high overweight prevalence may to some extent relate to an obesogenic environment offering many high-caloric fast foods such as sugar-sweetened beverages and cakes/cookies [4-13]; often liquid foods or foods with a soft texture can be consumed quickly and in larger quantities compared to foods that require chewing [8, 14, 15]. It has for instance been shown that participants consuming a liquid product consumed 30% more of the product compared to participants consuming a semi-solid product despite similar palatability, energy density and macronutrient composition [16]. Therefore, targeting eating rate may reduce overconsumption and consequently excessive weight gain [17], for instance by stimulating the intake of hard textured foods such as a fruit salad instead of smoothies or by creating awareness of individual eating speed to stimulate people to eat more slowly.

To date, already many interventions have been designed to steer individuals to a healthier diet. However due to their ‘one-size-fits-all’ approach with merely generic advice catered to a (sub)population, these nutritional interventions have shown only limited success. To effectively influence dietary behaviours, interventions need to be more tailored to the individual’s needs and preferences [18, 19], which is where mobile applications or technologies offer opportunities to coach individuals towards a healthier diet and prevent overconsumption [20]. However, at least three central questions need to be addressed. First, how to capture problems related to diet and eating behaviour on an individual level? It is an enormous challenge to accurately quantify an individual's food intake (i.e., food identification and portion size estimation) and as such to target personal diet optimization. There are many apps on the market that assess food intake but only very few are validated [21, 22]. Validated dietary assessment apps are key to ensure the collection of reliable food intake data. Nevertheless, due to their self-report nature, even validated apps still contain various sources of error, such as inaccurate portion size estimates, inaccurate food identification, and incomplete recordings [23]. Second; how to tailor the dietary intervention to the individual? The adoption of behaviour change theories is important to identify and influence key constructs related to behaviour change. However, multiple reviews indicate that the integration of behaviour change theories in diet-related apps is still limited [24-26]. Third, how to engage individuals in long-term use of diet-related apps? Perceived usability, perceived benefit, and trust in an app are important aspects for long-term user acceptance [27]. Moreover, intrinsic motivation is essential to commit to long-term use [28, 29]. As existing diet-related apps are quite burdensome when related to food recording, decreasing this burden is assumed to beneficially impact user engagement, preferably while simultaneously improving the accuracy of the records. The above aspects can already be addressed in the developmental process by involving potential end-users as well as other stakeholders (intermediate users) (e.g., general practitioners, dieticians) [30], which can be especially valuable when developing for specific target groups (e.g., children, low SES, older adults). This approach is not only assumed to benefit app quality, but also trustworthiness of the app, the likelihood of professionals' implementing the app in daily practice, and the prospect of adoption of the app by the target group [31].

New innovative technologies offer the opportunity for more personalized dietary interventions through apps, even more so by utilizing technologies such as wearables, different types of cameras, artificial intelligence (AI), and machine learning techniques [32-34]. The use of such smartphone applications allows for real-time personalized nutritional coaching, which requires less effort and has the potential to be more accurate compared to traditional methods [35]. However, the development of effective diet-related apps and technologies requires a multidisciplinary approach. Nutrition-, behaviour change-, (AI) technology-, health- and human-technology interaction expertise is crucial to develop and implement dietary interventions that have an actual impact on health.

THE PRIDE & PREJUDICE CONSORTIUM

Pride & Prejudice (P&P) is a consortium that features a unique combination of disciplines, involving life, nutritional, ICT, design, behavioural and social sciences, and is composed of researchers from all four Dutch Universities of Technology (Delft University of Technology, Eindhoven University of Technology, University of Twente, and Wageningen University & Research). Within this consortium in-depth nutrition and nutrition behaviour knowledge is combined with the technological knowledge of machine learning, lifestyle sensors, and knowledge of the design of effective health behaviour change interventions. As such, P&P offers an exceptional platform to jointly develop new technology-based opportunities to build upon existing tools (Figure 1), for example by studying features that simplify the reporting of dietary intake as well as increase its accuracy by combining photo, video and/or speech recording with machine-learning or AI. This will not only facilitate a more accurate dietary behaviour assessment among healthy adults, but may also aid assessment among specific target groups such as children and older adults. Moreover, the accuracy of dietary behaviour assessment directly influences the efficacy of nutritional behaviour coaching as dietary problems can be targeted more specifically [35]. Therefore, once a more accurate assessment can be achieved, individual, personalized dietary advice can be generated rather than general nutritional guidelines. P&P started its funded activities in January 2019 and will continue until January 2023. In this paper we outline our vision with respect to innovating our current dietary behaviour assessment and intervention tools.

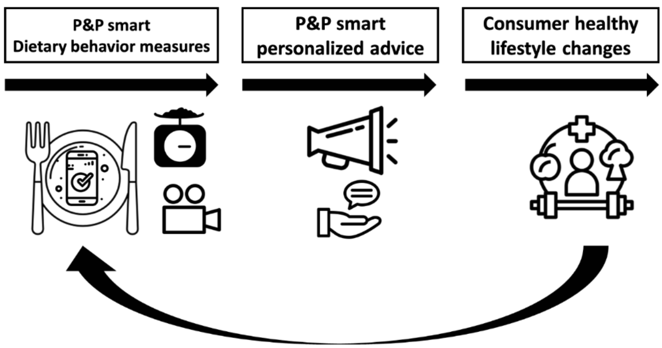


Figure 1. Overview of P&P efforts on technology driven dietary behaviour assessment, and personalized interventions and coaching.

CURRENT TOOLS TO ASSESS DIETARY INTAKE

Academia has long been investing considerable effort into creating validated tools to accurately assess dietary intake [21, 23]. The mainstay of dietary assessment is based on self-report methods, i.e., food frequency questionnaires (FFQ), 24-h recalls (24hR), and food records. Originally, these methods were paper-pencil based but these evolved into computer- and web-based tools. The level of automatization within these tools has multiple benefits compared to traditional methods, e.g., decreased level of error, improved level of accuracy, increased user-friendliness, lower burden, and reduced costs. More recently, these dietary assessments tools were further innovated through the implementation of smartphone technology, enabled by almost universal adoption of smartphones in the population. Smartphone-based tools (i.e., apps) have the major advantage of enabling real-time data collection at any location at which the owner of the phone is present. Moreover, smartphones have multiple build-in sensors that may provide valuable information without any effort to provide input by the consumer (e.g., GPS, pedometer), and smartphones also offer the opportunity to easily connect to other apps/sensors (e.g., heart rate monitors, wearables). Eldridge et al. [21] created a clear overview of the available dietary assessment tools. Strikingly, whereas web-based tools are generally based on the FFQ or 24hR, all smartphone-based tools are based on the food record method.

The division of Human Nutrition and Health of Wageningen University & Research (WUR) created a solid foundation to assess diet in the general population by means of FFQs, 24hRs and food records [23]. Since the early 2000s, the original paper-pencil methods evolved into web-based tools (i.e., FFQ-tool™ and Compl-eat™) [23]. In addition, the most recent innovative tool developed by WUR researchers is the app “Traqq”, which can be used as a recall and food record. In the food record module, users can enter consumed foods throughout the day. Conversely, the recall module invites the user to record food intake in a specific time period by means of notifications. Once the user opens the app, access is given to an extensive food list based on the Dutch food composition database. The food list is flexible and can be modified by the researcher to fit different research purposes or different target groups (i.e., include sports nutrition or infant foods). Once a food item is selected, a consumed amount needs to be inserted, which can be reported in household measures (e.g., cups, spoons, glasses), standard portion sizes (e.g., small, medium, large), and weight in grams. The app also allows the user to enter all ingredients of an original recipe in combination with the quantity of the meal consumed. Yield and retention factors (i.e., retained weight and nutrients after cooking) are automatically taken into account. Traqq also enables linkage to online survey tools (e.g., Qualtrics). This way, additional questions can be incorporated linked to specific foods, eating occasions or times (e.g., context-related questions, mood questions) [36].

User-friendliness and trustworthiness of Traqq was addressed by consulting intended end-users during the development process, which eventually resulted in a clear and simple tool that can be used with minimal instruction. The effectiveness of this process was confirmed by the results of the evaluation study showing system usability scores (SUS) of 79/100, representing above average (SUS > 68) to excellent usability (SUS > 80) [36]. The logo of WUR, a respected research university, was incorporated in the app to underline solid scientific underpinning and enhance system credibility [37]. Currently, the validation of the recall-app is ongoing. Participants record their dietary intake by means of the app as well as validated traditional web-based dietary-assessment tools, with further validation using blood and urine samples. To evaluate upgrades since the previous evaluation, usability is reassessed as well. The preliminary results in terms of user-friendliness, validity and reproducibility seem promising; the final results are expected in mid-2021.

Traqq distinguishes itself from other dietary assessment apps in terms of its flexible nature. The app can be tailored to fit different research purposes in terms of the food list, portion size options, eating occasion or time questions, including additional questions and sampling scheme options. Moreover, it is possible to alternate between the recall and the food record module. To the best of our knowledge, Traqq is the only dietary assessment app that can also function as a recall, while other dietary assessment apps are all based on the food record method [21, 36] that are prone to reactivity bias and social desirability bias, thus affecting dietary assessment accuracy [38, 39]. In turn this will influence reliability of provided personalized advice and negatively influence the intervention’s efficacy.

In addition to self-report methods, there are also sensor-based technologies available that may facilitate dietary intake assessment (e.g., body-worn cameras, tooth sensors, smart dishes with weighing scales). Vu et al, [34] created an extensive overview of wearable dietary assessment technologies that aid objective food intake measurement, and as such eliminate various self-report related errors (e.g., estimation errors). These technologies are often invasive, not fully automated, or only provide partial food intake assessment; for example, only consumed amounts without food identification (smart dishes) or only a limited number of nutrients (tooth sensors). Consequently, these technologies are mostly exclusively used in a laboratory-setting and not validated for use in a real-life setting [33, 34].

CURRENT TOOLS TO ASSESS EATING BEHAVIOUR

INDIVIDUAL EATING BEHAVIOUR

Sensory-based technology such as weighing scales and video recordings can be used to assess dietary behaviour (i.e., what and how much is consumed) and also to assess eating behaviour (i.e., how is it consumed, e.g., chewing rate, bite size) [40, 41]. Current gold standards to measure individual eating behaviour are labour intensive and not suited for use beyond a controlled environment. To detect automated eating behaviours or oral processing such as chews and bites, participants are filmed while eating. Subsequently, recordings are annotated by two independent trained observers. Programs often used to annotate videos are Observer® XT (Noldus Information Technology) and ELAN (Max Planck Institute) [42, 43]. Chews and bites are annotated over time with readouts such as meal duration (i.e., time between first and last bite), eating rate (i.e., gram per minute), bite size (i.e., gram per bite) and oral processing duration (i.e., time between first chew and main swallow) [15, 44]. Eating behaviour and food texture (liquid, solid) can also be assessed with ear sensors that measure sound, PPG and accelerometery [41, 45]. Oral processing in turn can also be measured with tracking dots (i.e., stickers on nose and chin) using Kinovea [46, 47], which provides detailed and objective information on chew cycle duration, number of chews, and oral processing duration. Moreover, oral processing can be assessed by electromyography (EMG) or electromagnetic articulography (EMA), which tracks electrical activity of the jaw muscles [48, 49]. EMG provides information on muscle effort or eating effort, number of chews, and chewing duration. EMA provides information on tongue movements, displacement, speed, and acceleration, which requires adhesive sensors on the tongue, jaw and cheeks. Other, more invasive methods of oral processing are video fluorograph and magnetic resonance imaging (MRI) that can be used to visualize the inside of the mouth and throat to track food and to observe movement of the jaw, tongue and oesophagus [50, 51]. The later methods are particularly suited to identify anatomical abnormalities and tracking of the swallowing process (e.g., evaluation of choking hazards in patients). To measure food properties in relation to eating behaviour, bolus properties can be measured such as dry matter, particle size and mechanical properties [52]. Additionally, actual food intake (i.e., consumed amounts) is mostly measured by weighing before and after the meal.

As stated, most methods used to study eating behaviour are not suited for field studies, but do allow for measures of food intake and eating behaviour outside the lab [40, 41]. Recently, WUR combined these technologies in a weighing-tray, hereafter referred to as the mEETr (derived from the Dutch words for 'measurement device' and 'to eat'). The mEETr consists of a regular dining tray with three built-in weighing stations. These three weighing stations continuously measure the weight of a bowl, plate, and drinking cup or glass. Each weighing station consists of three triangle positioned measurement points (sensors) to balance weight. Besides these weighing stations, the mEETr tray includes a video-camera holder. Using a camera holder on the plate ensures that the camera is well positioned for the dietary and eating behaviour measurements. Based on these video images, eating behaviour characteristics can be determined as an extension of emotion detection software [53]. Eating behaviour characteristics that can be determined are number of bites, sips, chews and swallows. Combining eating behaviour information with weighing data of the meal facilitates calculation of bite size, eating duration per bite and eating rate, and the order in which different meal components are consumed. Tray weight and video data are transported to a PC using a wireless receiver. Here the data is cleaned after which outcome measures are calculated. The mEETr gives insight into how much people eat and the way food is consumed, which can be used for research purposes and to provide personalized eating behaviour advice. More specifically, this could result in recommendations on food type (whole fruit instead of juices and smoothies) and eating rate, which adds to current dietary-assessment methods such as the 24hR, FFQ or food records that do not provide information on eating rate. Therefore, the mEETr is especially of use in specific population groups in which current dietary questionnaires cannot be used, such as in children [54, 55].

SOCIAL EATING BEHAVIOUR

Eating is not only about what, how much, and how food is consumed, it also about the social environment: 'the social space of eating'. Sharing a meal can involve overt social aspects such as serving oneself or someone else food, passing on plates or serving trays, and adjusting overall dining time to table partners. There are also more subtle or covert social aspects involved such as going for seconds or thirds, synchronizing bites or eating speed, and synchronizing eating or serving quantity [56]. To understand the implicit social dynamics of eating together and how this impacts food choices, studies of social interaction are needed to also address social environmental factors in personalized dietary coaching. Social eating behaviour is predominantly assessed through video recordings, which provides valuable insights into effects of conversations, eye contact between table members, or gesture mimicking [57]. However, video recordings miss out on dimensions like quantities of foods, which can relate to individual food intake but also food sharing. Although such dimensions can be monitored through weighing scales or instrumented cutlery, such attributes create awareness and may intervene with natural behaviours of the table members, and such introduce reactivity bias.

To address this, the University of Twente developed the Sensory Interactive Table (SIT) to measure both individual and social eating behaviours [58]. The SIT is an instrumented,

interactive round dining table (⌀1.45 m). The table surface is composed of 199 individually controllable, hexagon-shaped modules, each embedded with a load cell (199 load cells total) and 42 LEDs each (8358 LEDs total) just below the tabletop surface. Each module is covered with hexagon-shaped 15 mm thick white plexiglass to diffuse the LED light. Modules are replaceable, providing the option to use other sensors and feedback modalities. A plastic foil is placed on top of the plexiglass to create a waterproof surface. A tablecloth makes up the last layer of the table to create a visually appealing unobtrusive measurement instrument [58].
The load cells measure the weight of the items on the table, over the course of a meal. As a result, many overt and covert aspects of eating behaviour related to mass become measurable, such as bite size, total amount of food on a plate, or synchronicity of eating speed between individual table members. Additionally, the SIT provides the option for coaching through use of the LEDs, which allow for communication with table members by use of light interactions, potentially providing feedback and advice about their behaviour, habits and eating choices [58].
Currently, the table is controlled through Unity (Unity Technologies), a cross-platform game engine that collects and processes the data from the loadcells (input) and control the interactions that is sent back towards the LEDs (output). The software allows for individual processing of input and output alone, or can interconnect the two, creating a feedback loop to the user. It creates a flexible set-up, suitable to study the eating behaviours of people in a social setting, the social interactions between people in a dining setting, and the continuous cycle of feedback and the response to this feedback in real time [58].

Technology to Improve Measures of Dietary Behaviour

IMAGE AND SPECTROSCOPY TO IMPROVE DIETARY INTAKE MEASURES

In order to improve the quality of dietary intake assessment, a feature that may contribute to a more accurate dietary assessment is the use of images [59, 60]. Images can be particularly interesting when tools are devoted to specific populations, such as children or individuals with intellectual impairment, as related to limited skills in terms of literacy, writing, food recognition and dependency of the care-giver [61-63]. Smartphone-based systems that estimate food intake by means of images already exist [e.g., 64], but their usability still seems limited and results are often insufficiently validated.

Current efforts are particularly focused on two approaches, including image-assisted and/or food recognition image-based approaches. Image-assisted approaches are especially useful as part of retrospective methods (i.e., 24hR). Interviewees capture all food and drinks consumed through pictures, which subsequently assists the reporting of the foods and drinks consumed and associated portion sizes. However, for this approach image-review is required by the interviewee and/or researcher, which makes the method quite tedious. To circumvent this problem, automated food recognition (and volume) image-based approaches can be used as part of prospective methods (i.e., food records); interviewees record their intake by taking before and after pictures of all foods and drinks consumed. As image-based recording may facilitate automatic food identification and portion size estimation, this approach is especially promising to reduce respondent and researcher burden and increases the accuracy of current food and nutrient intake estimates by reducing reactivity bias, portion size errors and errors due to incorrect food identification. However, these methods still have many limitations as pictures must be taken from a specific angle and often in combination with a reference marker (i.e., to assess size and depth). Additionally, food recognition is still a challenge, especially in case of mixed-dishes or when differentiating between a diet soda and a sugar-containing version of that same soda [60, 64]. Portion size estimation from pictures is also still a challenge as the weight estimate is based on food volume. However, volume is food specific, i.e., whereas a salad is voluminous and light, a candy bar is often dense and heavy [65, 66]. Differences in food volume can be better assessed by using 3D-images over regular images. A 3D-camera, which is already part of the newest smartphone models, is able to detect the shape and volume of a food item. This allows the assessment of portion sizes.

Besides regular images, spectroscopic images of food products may serve future energy and macronutrient estimates of foods, so called chemical fingerprinting [67, 68]. A spectroscopic camera can detect many different frequencies of light outside of the visible spectrum, which may facilitate product identification. Besides, spectroscopy uses near-infrared and infrared, which serve the assessment of food composition. The translation from chemical fingerprint to geometric and dynamic dietary information requires advanced 'chemometric' data-analyses techniques [59]. To determine a set of wavelengths needed to quantify macronutrient content, a 400–800 nm wavelength hyperspectral camera is needed. Such cameras with a limited wavelength range are expected to be incorporated in smartphones in the near future. Thus, integrating a hyperspectral camera and a 3D-camera could provide more detailed food data, and lead to more objective measures of food intake.

CONVERSATIONAL AGENTS TO IMPROVE DIETARY INTAKE MEASURES

Exploring the potential use of a conversational agent or chatbot could be another valuable supplementary input source to assess dietary intake, particularly among (older) adults with functional impairments (e.g., visual and/or motor impairment) and individuals with limited (E-)health literacy [69, 70]. Implementation of a conversational agent may further simplify the recording process and increase accuracy [71, 72].

A distinction can be made between rule-based chatbots and AI-based chatbots. Rule-based bots are the most common chatbots and are programmed according to a decision tree architecture; users have to answer specific (often closed) questions via text- and/or button-entry, after which the bot will respond based on the fixed decision tree. Therefore, rule-based bots are only useful for 'simple' tasks (e.g., tracking and stimulating fruit intake). In case of more complex tasks, AI-based chatbots are more suitable due to use of AI and Natural Language Processing (NLP), which enable more advanced 'conversations'. By converting and interpreting text/speech and even images, AI-bots have the ability to make 'intelligent decisions', 'learn' (i.e., machine learning, deep learning) and provide better and more accurate answers in case of more long-term use [71]. Due to this learning process, AI-based bots are able to identify frequently consumed foods related to eating occasions and/or identify habitual consumption patterns, which in its turn enables the AI-bot to send personalized reminders at opportune moments to remember users to report their food intake.

Currently, multiple diet-related chatbots are available to assess and (often) influence dietary behaviour [73-76], but these are not yet validated. Choi and Kim [79] evaluated the feasibility and acceptability of ICT based mobile chatbot technology to reduce dietary sugar intake: >60% of the participants reported difficulties related to the use of the chatbot and forgot to record their food intake, resulting in incomplete food intake registration. AI and machine learning techniques are assumed to further stimulate the use of conversational agents to assist dietary intake assessment, especially among specific target groups. Nevertheless, although chatbots may reduce participant burden, reporting an entire diet via chatbot may still be tedious due to the question-answer structure. Consequently, to further reduce participant burden, integrating the use of chatbots with other tools is needed. Utilization of available smartphone features and advanced AI and machine learning techniques are eminent to facilitate multiple data entry methods (text, speech, images), integrate personalized reminders and minimize reporting burden, but also to analyses the collected, often complex, data. These technological advances are not able to fully replace existing dietary assessment technologies but could be valuable additions to existing tools such as Traqq and have the potential to decrease reporting burden and improve accuracy, especially for specific target groups.

VIDEO IMAGE ANALYSIS AND SENSORS TO IMPROVE EATING BEHAVIOUR MEASURES

Whereas cameras provide opportunities to improve dietary intake assessment, advanced video image analysis techniques offer new opportunities in terms of automated detection of eating behaviours such as emotion detection (adults and children), acceptance and rejection behaviour of infants and automated oral processing behaviours such as chews and bites in different age groups [77, 78]. Deep learning models can be based on extracted data from the video, for example facial landmarks, and training the model on annotated events and their time. Alternatively, video images can be analysed as a whole using all the available pixels as inputs to the model. This allows in context analysis, tracking all behaviours that occur when eating a family meal where interactions with the environment and all family members are taken into account [79]. The former may yield higher accuracy with simpler models and smaller datasets, and the latter may allow more comprehensive machine learning of the full complexity of eating behaviour. Besides video processing, eating behaviour can also be assessed using newly-developed sensors such as a wristband with an IMU sensor to detect the hand-movement bringing food to the mouth [80, 81]. In the future, eating events may also be detected using other sensor wearables such as headbands like Muse™, which measures EEG, PPG heartrate and accelerometery. Based on input from all these different sensors eating episodes or chewing may be detected using prediction analysis such as a neural networks (AI).

Technology to Improve Personalized Dietary Behaviour Interventions

Conversational Agents for Stimulating Dietary Behaviour Change

In addition to using conversational agents for dietary assessment, conversational agents are also used to stimulate dietary behaviour change. However, considering existing apps using this technology, it can be concluded that the level of personalized advice is limited. Users are often referred to health professionals for more detailed advice [71]. Integrating AI-based chatbots will enable the detection and visualization of dietary patterns, both graphical and textual, as well as the provision of real-time personalized suggestions and goal setting to promote small daily changes. Such changes or goals could apply to changes in dietary intake as well as eating behaviour. More specifically, in terms of eating behaviour, the chatbot could pop questions related to appetite and feelings of hunger, which could translate into portion sizes suggestions, or a timer could be initiated to motivate the user to eat more slowly.

Additionally, embodied conversational agents (ECAs) also seem promising in stimulating and maintaining health behaviour change [82]. ECAs are animated computer characters (i.e., avatars) that are able to establish and maintain a more personal relationship with the user [83]. Research indicates that face-to-face coaching seems to be more effective in establishing long lasting behaviour change [19]. As ECAs can mimic face-to-face coaching and are available 24/7, ECAs can offer coaching when it is needed most. Moreover, motivational features can stimulate the user to adhere to the recommendations and, for instance, support users with shopping in a healthy and sustainable way [84]. Additionally, avatars can also be used to communicate feelings of hunger and satiation to teach children, for instance, when to stop and start eating through imitation, establishing healthy eating behaviour from an early age [85]. Therefore, implementing an AI-based chatbot or ECA into existing tools such as Traqq can be very valuable in enabling tailored diet coaching [82, 83]. Future research should address how the use of ECAs can be optimized in order to create actual behaviour change. Regular interaction is a prerequisite, but like other eHealth tools, uptake is limited. Studies show that appearance of the ECA matters, but even when appearance and design is optimized, the dialogues need to remain persuasive and engaging over time. It might well be that conversational agents for dietary change could function best in an add-on format to other interventions (e.g., in addition to regular care or dietician’s advice) or for specific patients group that are highly motivated to adjust their diets (e.g., patients in cardiac rehabilitation) [86-88].

GAME-ELEMENTS TO IMPROVE DIETARY BEHAVIOUR INTERVENTIONS

Another promising and attractive strategy to increase intrinsic motivation and engagement of diet-related tools is via games (i.e., serious games) or game elements (i.e., gamification) [89, 90]. By harnessing the 'fun factor' of games, enjoyment in using diet-related tools can be increased and in such way users' motivation and engagement can be fostered. Moreover, digital tools can become easier to use and better understandable [90, 91]. Although research on effectiveness is still in its infancy, reviews show promising results of gamified interventions to promote healthy behaviours and, specifically, the promotion of a healthy diet, both in children and in adults [92-94]. An example of a gamified intervention is the serious game 'Squire's Quest!' that encourages children to consume more fruits and vegetables. The purpose of the game is to go from squire to knighthood and for this the squire (i.e., child) had to overcome multiple challenges. These challenges consisted of attaining real-word fruit and vegetable consumption-related goals. By successfully completing these challenges, the child earned badges and progressed towards knighthood [95]. Squire's Guest! contains a variety of integrated behaviour change techniques to promote self-efficacy and intrinsic motivations (e.g., goal setting, planning, self-monitoring, goal review and feedback), which are key mediators of behaviour change in children [96]. To optimize chances of app engagement and effectiveness it is important to match the game or gamified tool to the user's needs and preferences, therefore, it is recommended to follow a user-centred design approach when developing gamified diet-related tools [97, 98].

TARGETED DIETARY INTERVENTIONS DURING FOOD SHOPPING

Trustworthy diet-related apps also have the potential to deliver real-time dietary advice to users when food shopping. For example, users receive personal dietary advice, e.g., specific product recommendations, recipes, grocery lists, tailored to the assortment while shopping in an online grocery store or when using a smartphone or handheld scanner devices in a physical store.

To date, nutrition health apps (i.e., apps aiming at improving users' health) and nutrition information apps (i.e., apps aiming at delivering transparent nutrition information of food products) [99] exist that help make smarter food purchases. These have mostly been developed by researchers [100, 101], governmental organizations [102], retailers, or other organizations. Most of these dietary recommendations apps focus on tracking weight loss, or specific health conditions over a longer period and aim to educate users to adopt a healthier lifestyle. However, few focus on the specific context of food consumption and purchasing moments [99, 101, 103-106]. In addition, most apps do not include personalized dietary advice but are based on general nutrition guidelines.

It is assumed that app-based interventions that do not solely focus on educating but include more concrete coaching in the form of eating tips, recipes, and grocery lists (e.g., for specific food retailers) make it easier to adopt behavioural change [105]. A practical and tailored grocery shopping list can assist in healthful shopping. An example, ‘MyNutriCart’, is a smartphone app that helps users select healthier foods based on the U.S. Dietary Guidelines and on their budget [100, 106]. It offers users a practical grocery list fitting the dietary nutritional recommendations and taking into account the caloric requirements based on their weight goal (i.e., loose, gain or maintain weight) [100] and led to improvements in terms of healthy food-related behaviour [106].

Integrating such approach into a dietary assessment app like Traqq could also simplify the reporting of food intake. Recommended recipes and grocery lists can be stored in the app and consulted during food intake reporting. Consumed recipes or products can then easily be transferred to the food record or recall module of the app.

Tailored Dietary Behaviour Interventions

While most food intake recommendations follow a one-size-fits-all approach catered to a population or subpopulation, personalized dietary advice provides recommendations tailored to the individual. Therefore, personalized dietary advice fits better with individual needs and leads to a more effective approach towards long-term change in dietary behaviours [107, 108]. In order to personalize advice, various individual characteristics and their interactions can be taken into account such as educational level, social and economic status, current nutritional status (i.e., nutrient deficits and surpluses), and individual lifestyle behaviours and preferences. Provision of personalized advice may offer opportunities for tailored interventions with superior health benefits aligned to the individual’s nutritional status and a better compliance with the advice through better alignment with individual lifestyles and preferences. Examples of such a personalized intervention are Just-In-Time-Adaptive Interventions (JITAIs). JITAIs adapt their support “overtime to an individual’s changing status and contexts”, aiming to deliver support “at the moment and in the context that the person needs it most and is most likely to be receptive” [109]. Data from smartphones or wearables are used to automatically and continuously acquire information about the user and its context (e.g., environmental exposures) and deliver individualized interventions based on states of vulnerability and receptivity of the user [110, 111]. Based on these data the delivery of the intervention elements can be continuously tailored towards the specific status and context of the user [112].

Additionally, to provide tailored dietary interventions, it is crucial to adopt behaviour change theory to identify and influence key constructs related to behaviour change. Currently, the integration of behaviour change techniques (BCTs) in diet-related apps is often lacking [24-26]. Although diet-related apps vary greatly in the number of integrated BCTs, goal setting, self-monitoring and feedback are integrated most frequently [24]. These BCTs have been proven effective in general weight loss interventions [113-115]. However, these techniques all relate to behavioural control and do not focus on other constructs such as development of essential behavioural skills [24]. More recently, Villinger et al. [119] reviewed the effectiveness of app-based diet-related behavioural interventions and concluded that additional intervention components besides the app and a higher number of implemented BCTs did not necessarily improve the effectiveness of the interventions. Thus, the effectiveness of an app-based diet intervention will not be determined by the quantity of implemented BCTs, but by their quality. The design and the technical implementation of a BCT can influence effectiveness of an app [116]. Implementing a variety of BCTs will enable the user to tailor the app to their preferences and develop a personalized intervention.

DISCUSSION

This paper discusses technological opportunities to improve dietary behaviour assessment and interventions. The P&P consortium offers a unique combination of disciplines, which is needed to improve dietary behaviour assessment and subsequently tailor interventions to establish lasting dietary behaviour change. P&P's current efforts have led to the development of tools such as Traqq, mEETr and SIT. These tools allow for assessment of dietary intake, eating behaviour and social or contextual dietary behaviours.

Developing targeted dietary behaviour measures and connecting these to behaviour change interventions is key to the establishment of lasting behaviour change in order to ultimately improve health. Specific target (sub)populations have explicit (nutritional) requirements, and behaviour change efforts should be tailored to the individual. Consequently, various individual characteristics should be taken into account such as age, culture, social-economic status, personality trait and level according to the theory of behaviour change [117]. During the development of tailored interventions or tools it is imperative to take these requirements in to account. Currently, P&P efforts are focused on four main target populations.

The first target group of the P&P consortium are pregnant, lactating women and their (unborn) children. Although a healthy diet during pregnancy and lactation is important to ensure optimal supply of various nutritional sources, the diet of pregnant and lactating women is often suboptimal [118-120], which stresses the need for more effective personalized approaches. The new technologies as described in this paper offer the opportunity to develop more personalized approaches for pregnant and lactating women. Another window of opportunity to stimulate a healthy diet is childhood, which highlights the second target group of the P&P consortium, namely children. Currently, dietary assessment among young children is challenged by limited reading and writing skills as well as food knowledge of the children themselves. As a result, health care professionals and researchers depend on the caregivers. Therefore, P&P aims to develop practical technological tools to optimize dietary intake assessment and guidance of (future) mothers and their young children. The third target group of the P&P consortium are older adults, who have specific nutritional needs, often limited digital capabilities, and prefer the more personal approach [75]. Human-computer interaction design seems promising here, e.g., using ECAs that are able to mimic face-to-face coaching. A fourth target group of the P&P consortium are office workers, who spend a large proportion of their time at work where they consume about a third of their daily energy intake [121, 122]. The office environment is ideal to set up dietary interventions due to fixed work schedules and a generally limited access to food items and meals [121, 123].

To arrive at personalized diet-related tools, P&P not only focuses on end-users but also on stakeholders (e.g., health and nutrition professionals, commercial firms, government bodies). We investigate underlying values that play a role for different stakeholders; uncovering values provides insight into how people wish to live their lives and what matters most to them. Stakeholder values and any potential tensions between co-existing values need to be taken into account in the design of interventions. In part, the focus needs to be on the worth that is created for different parties [124]. For example, for users there could be worth in having to spend less time on their diet, while for health and nutrition professionals a better insight into the nutritional needs of different target groups could be of interest. However, stakeholders' internal values also need to be taken into account, such as the need for privacy, or being able to autonomously decide about food intake [124]. We feel that adopting a multi-stakeholder, value-based approach is crucial to arrive at tools that can be seamlessly integrated into people's daily lives for long-term use, and that will be supported by the ecosystem of stakeholders involved.

To achieve personalized technology-assisted dietary interventions, supporting technologies need access to relevant data on the individual. This implies that individuals have to be willing to engage in an information exchange process with the diet-coaching app [125]. People have to share personal, often sensitive information with the app, which induces (perceived) privacy risks for users in order to receive tailored advice. Therefore, the use of user-documented food consumption data also raises specific legal and ethical challenges [22]. The trade-off between the perceived personal benefit of this advice and the perceived privacy risk of sharing personal information will affect whether or not users adopt the diet-coaching app [125, 126]. Therefore in developing apps to facilitate personalized advice, characteristics such as the type, amount and sensitivity of personal information disclosed, advice scope (e.g., personal diet plan, personalized shopping list, religious taboos, personal allergies), the trustworthiness of the communication channel, business model, and service provider need are to be taken into account [125, 127].

Nowadays, technology advances rapidly which results in new technological opportunities to improve dietary behaviour assessment and intervention. However, it also brings along new challenges. Due to the rapid technological evolution, it is difficult to stay up-to-date. New quickly becomes outdated. To ensure the quality of the tools we develop, we focus on techniques that have been proven effective in either dietary assessment or dietary behaviour change. We do not attempt to develop new techniques but implement existing techniques. To further ensure the accuracy and effectiveness of the tools, thorough evaluation is an important aspect of the developmental process.

CONCLUSIONS

Technological innovations offer the opportunity to improve dietary behaviour assessment and interventions. Moreover, they enable targeted dietary behaviour interventions, tailored to individuals, specific target groups or situations (e.g., during food shopping). Advanced image and video processing in combination with AI and machine learning techniques will be explored to improve current dietary behaviour measures and to reduce registration burden and improve accuracy. Integration of conversational agents (i.e., chatbots, avatars) and game elements in existing systems such as Traqq show promise in tailoring dietary behaviour coaching to improve engagement. Finally, targeted dietary behaviour interventions can be improved by integrating behaviour change techniques and tailoring to the individual, target group, or situation. Utilization of these technological innovations in dietary behaviour assessment and interventions has the potential to significantly improve the healthiness of individuals' eating behaviours.

REFERENCES

1. Blakely, T., Major strides in forecasting future health. Lancet, 2018. 392(10159): p. e14-e15.
2. Gezondheidsraad, Richtlijnen Goede Voeding 2015. 2015, Gezondheidsraad: Den Haag, The Netherlands.
3. van Rossum, C.T.M., et al., The diet of the Dutch. Results of the Dutch National Food Consumption Survey 2012-2016. 2020, National Institute for Public Health and the Environment: Bilthoven, The Netherlands.
4. Berthoud, H.R., The neurobiology of food intake in an obesogenic environment. Proceedings of the Nutrition Society, 2012. 71(4): p. 478-487.
5. French, S.A., et al., Eating behavior dimensions. Associations with energy intake and body weight. A review. Appetite, 2012. 59(2): p. 541-549.
6. Viskaal-van Dongen, M., F.J. Kok, and C. de Graaf, Eating rate of commonly consumed foods promotes food and energy intake. Appetite, 2011. 56(1): p. 25-31.
7. Ohkuma, T., et al., Association between eating rate and obesity: a systematic review and meta-analysis. International journal of obesity, 2015. 39(11): p. 1589-1596.
8. Robinson, E., et al., A systematic review and meta-analysis examining the effect of eating rate on energy intake and hunger. American Journal of Clinical Nutrition, 2014. 100(1): p. 123-151.
9. Hill, J.O., et al., Obesity and the environment: where do we go from here? Science, 2003. 299(5608): p. 853-855.
10. Rolls, B.J., The relationship between dietary energy density and energy intake. Physiology & behavior, 2009. 97(5): p. 609-615.
11. De Graaf, C. and F.J. Kok, Slow food, fast food and the control of food intake. Nature Reviews Endocrinology, 2010. 6(5): p. 290.
12. Eaton, S.B. and M. Konner, A consideration of its nature and current implications. N Engl j Med, 1985. 312(5): p. 283-9.
13. Cordain, L., et al., Origins and evolution of the Western diet: health implications for the 21st century. The American journal of clinical nutrition, 2005. 81(2): p. 341-354.
14. Krop, E.M., et al., Influence of oral processing on appetite and food intake - A systematic review and meta-analysis. Appetite, 2018. 125: p. 253-269.
15. Lasschuijt, M.P., et al., Comparison of oro-sensory exposure duration and intensity manipulations on satiation. Physiol Behav, 2017. 176: p. 76-83.
16. Zijlstra, N., et al., The effect of viscosity on ad libitum food intake. International journal of obesity, 2008. 32(4): p. 676-683.

17. Bolhuis, D.P. and C.G. Forde, Application of food texture to moderate oral processing behaviors and energy intake. Trends in Food Science & Technology, 2020. 106: p. 445-456.
18. Kelly, M.P. and M. Barker, Why is changing health-related behaviour so difficult? Public Health, 2016. 136: p. 109-116.
19. Al-Awadhi, B., et al., Insights Into the Delivery of Personalized Nutrition: Evidence From Face-To-Face and Web-Based Dietary Interventions. Front Nutr, 2020. 7: p. 570531.
20. Payne, H.E., et al., Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth, 2015. 3(1): p. e20.
21. Eldridge, A.L., et al., Evaluation of New Technology-Based Tools for Dietary Intake Assessment—An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients, 2019. 11(1): p. 55.
22. Maringer, M., et al., User-documented food consumption data from publicly available apps: an analysis of opportunities and challenges for nutrition research. Nutrition Journal, 2018. 17(1): p. 59.
23. Brouwer-Brolsma, E.M., et al. Dietary intake assessment: From traditional paper-pencil questionnaires to technology-based tools. in International Symposium on Environmental Software Systems. 2020. Springer.
24. Bardus, M., et al., A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. International Journal of Behavioral Nutrition and Physical Activity, 2016. 13(1): p. 1-9.
25. DiFilippo, K.N., et al., The use of mobile apps to improve nutrition outcomes: a systematic literature review. Journal of telemedicine and telecare, 2015. 21(5): p. 243-253.
26. Ferrara, G., et al., A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. JMIR mHealth and uHealth, 2019. 7(5): p. e9232.
27. König, L.M., et al., Barriers to and Facilitators for Using Nutrition Apps: Systematic Review and Conceptual Framework. JMIR mHealth and uHealth, 2021. 9(6): p. e20037.
28. Chang, T.-R., E. Kaasinen, and K. Kaipainen. What influences users' decisions to take apps into use? A framework for evaluating persuasive and engaging design in mobile apps for well-being. in Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia. 2012. Ulm, Germany.
29. Peng, W., et al., A qualitative study of user perceptions of mobile health apps. BMC Public Health, 2016. 16(1): p. 1158.
30. Preece, J., H. Sharp, and Y. Rogers, Interaction design: beyond human-computer interaction. 2015, Hoboken, NJ, USA: John Wiley & Sons.

31. Chen, J., et al., Designing Health Apps to Support Dietetic Professional Practice and Their Patients: Qualitative Results From an International Survey. JMIR Mhealth Uhealth, 2017. 5(3): p. e40.
32. Cambon, L., Health smart devices and applications…towards a new model of prevention? European Journal of Public Health, 2017. 27(3): p. 390-391.
33. McClung, H.L., et al., Dietary Intake and Physical Activity Assessment: Current Tools, Techniques, and Technologies for Use in Adult Populations. Am J Prev Med, 2018. 55(4): p. e93-e104.
34. Vu, T., et al., Wearable Food Intake Monitoring Technologies: A Comprehensive Review. Computers, 2017. 6(1).
35. Forster, H., et al., Personalised nutrition: the role of new dietary assessment methods. Proc Nutr Soc, 2016. 75(1): p. 96-105.
36. Lucassen, D.A., et al., Iterative Development of an Innovative Smartphone-Based Dietary Assessment Tool: Traqq. Journal of visualized experiments: JoVE, 2021(169).
37. Oinas-Kukkonen, H. and M. Harjumaa, Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems, 2009. 24(1): p. 28.
38. Subar, A.F., et al., Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. J Nutr, 2015. 145(12): p. 2639-45.
39. Thompson, F.E. and A.F. Subar, Dietary Assessment Methodology, in Nutrition in the Prevention and Treatment of Disease. 2017, Elsevier Inc.: Amsterdam, The Netherlands. p. 5-48.
40. Ioakimidis, I., et al., Description of chewing and food intake over the course of a meal. Physiol Behav, 2011. 104(5): p. 761-9.
41. Papapanagiotou, V., et al., A Novel Chewing Detection System Based on PPG, Audio, and Accelerometry. IEEE J Biomed Health Inform, 2017. 21(3): p. 607-618.
42. Lausberg, H. and H. Sloetjes, Coding gestural behavior with the NEUROGES-ELAN system. Behavior Research Methods, 2009. 41(3): p. 841-849.
43. Noldus Information Technology. Observer XT. 2021 29 April 2021]; Available from: https://www.noldus.com/observer-xt?gclid=Cj0KCQjwsqmEBhDiARIsANV8H3ZF1xKvZD3hx2eYpPV-rw75WQDKl1FCT7DTWeL6TZHFywrMdnXrY5AaAva9EALw_wcB.
44. McCrickerd, K., et al., Texture-Based Differences in Eating Rate Reduce the Impact of Increased Energy Density and Large Portions on Meal Size in Adults. Journal of Nutrition, 2017. 147(6): p. 1208-1217.
45. Sazonov, E., et al., Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. Physiological Measurement, 2008. 29(5): p. 525-541.

46. Puig-Divi, A., et al., Validity and reliability of the Kinovea program in obtaining angles and distances using coordinates in 4 perspectives. Plos One, 2019. 14(6): p. e0216448.
47. Aguayo-Mendoza, M.G., et al., Oral processing behavior of drinkable, spoonable and chewable foods is primarily determined by rheological and mechanical food properties. Food Quality and Preference, 2019. 71: p. 87-95.
48. Kohyama, K., et al., Natural eating behavior of two types of hydrocolloid gels as measured by electromyography: Quantitative analysis of mouthful size effects. Food Hydrocolloids, 2016. 52: p. 243-252.
49. Fuentes, R., et al., Systematic Standardized and Individualized Assessment of Masticatory Cycles Using Electromagnetic 3D Articulography and Computer Scripts. Biomed Research International, 2017. 2017.
50. Okada, A., et al., Oral behavior from food intake until terminal swallow. Physiology & Behavior, 2007. 90(1): p. 172-179.
51. Lee, J., et al., Semi-automatic segmentation for 3D motion analysis of the tongue with dynamic MRI. Computerized Medical Imaging and Graphics, 2014. 38(8): p. 714-724.
52. de Lavergne, M.D., F. van de Velde, and M. Stieger, Bolus matters: the influence of food oral breakdown on dynamic texture perception. Food & Function, 2017. 8(2): p. 464-480.
53. Lewinski, P., T.M. den Uyl, and C. Butler, Automated Facial Coding: Validation of Basic Emotions and FACS AUs in FaceReader. Journal of Neuroscience Psychology and Economics, 2014. 7(4): p. 227-236.
54. Liu, C., et al., A New Deep Learning-Based Food Recognition System for Dietary Assessment on An Edge Computing Service Infrastructure. Ieee Transactions on Services Computing, 2018. 11(2): p. 249-261.
55. Mezgec, S. and B.K. Seljak, NutriNet: A Deep Learning Food and Drink Image Recognition System for Dietary Assessment. Nutrients, 2017. 9(7): p. 657.
56. De Vries, R.A., et al. The Sensory Interactive Table: Exploring the Social Space of Eating. in Proceedings of the 2020 International Conference on Multimodal Interaction. 2020. Amsterdam, The Netherlands.
57. Pesch, M.H. and J.C. Lumeng, Methodological considerations for observational coding of eating and feeding behaviors in children and their families. International Journal of Behavioral Nutrition and Physical Activity, 2017. 14(1): p. 1-14.
58. Haarman, J.A., et al., Sensory Interactive Table (SIT)—Development of a Measurement Instrument to Support Healthy Eating in a Social Dining Setting. Sensors, 2020. 20(9): p. 2636.
59. Kim, Y., M. Singh, and S.E. Kays, Near-infrared spectroscopic analysis of macronutrients and energy in homogenized meals. Food Chemistry, 2007. 105(3): p. 1248-1255.

60. Boushey, C.J., et al., New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proceedings of the Nutrition Society, 2017. 76(3): p. 283-294.
61. Ptomey, L.T., et al., Validity of energy intake estimated by digital photography plus recall in overweight and obese young adults. Journal of the Academy of Nutrition and Dietetics, 2015. 115(9): p. 1392-1399.
62. Elinder, L., et al., Validation of personal digital photography to assess dietary quality among people with intellectual disabilities. Journal of intellectual disability research, 2012. 56(2): p. 221-226.
63. Ptomey, L.T., et al., Photo-assisted recall increases estimates of energy and macronutrient intake in adults with intellectual and developmental disabilities. Journal of the Academy of Nutrition and Dietetics, 2013. 113(12): p. 1704-1709.
64. Ahmad, Z., et al., A Mobile Food Record For Integrated Dietary Assessment. MADiMa16 (2016), 2016. 2016: p. 53-62.
65. Lo, F.P.W., et al., Image-Based Food Classification and Volume Estimation for Dietary Assessment: A Review. IEEE J Biomed Health Inform, 2020. 24(7): p. 1926-1939.
66. Chang, X., et al., Image-Based Food Volume Estimation. CEA13 (2013), 2013. 2013: p. 75-80.
67. Wang, W. and J. Paliwal, Near-infrared spectroscopy and imaging in food quality and safety. Sensing and Instrumentation for Food Quality and Safety, 2007. 1(4): p. 193-207.
68. Pallav, P., et al., A Near-Infrared (NIR) Technique for Imaging Food Materials. Journal of Food Science, 2009. 74(1): p. E23-E33.
69. Bickmore, T.W., et al. Managing chronic conditions with a smartphone-based conversational virtual agent. in Proceedings of the 18th International Conference on Intelligent Virtual Agents. 2018. Sydney, Australia.
70. Bol, N., N. Helberger, and J.C.M. Weert, Differences in mobile health app use: A source of new digital inequalities? The Information Society, 2018. 34(3): p. 183-193.
71. Fadhil, A., Can a chatbot determine my diet?: Addressing challenges of chatbot application for meal recommendation. arXiv preprint arXiv:1802.09100, 2018.
72. Franco, R.Z., et al., Popular Nutrition-Related Mobile Apps: A Feature Assessment. JMIR Mhealth Uhealth, 2016. 4(3): p. e85.
73. Graf, B., et al. Nombot: simplify food tracking. in Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia. 2015. Singapore.
74. Casas, J., E. Mugellini, and O.A. Khaled. Food diary coaching chatbot. in Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. 2018.

75. Rusu, A., et al., Aspects Influencing Food Intake and Approaches towards Personalising Nutrition in the Elderly. Journal of Population Ageing, 2020. 13(2): p. 239-256.
76. Forksy website. 26 February 2021]; Available from: https://getforksy.com/.
77. Bell, B.M., et al., Automatic, wearable-based, in-field eating detection approaches for public health research: a scoping review. Npj Digital Medicine, 2020. 3(1): p. 1-14.
78. Hossain, D., T. Ghosh, and E. Sazonov, Automatic Count of Bites and Chews From Videos of Eating Episodes. Ieee Access, 2020. 8: p. 101934-101945.
79. Farooq, M., et al., Validation of Sensor-Based Food Intake Detection by Multicamera Video Observation in an Unconstrained Environment. Nutrients, 2019. 11(3): p. 609.
80. Kyritsis, K., et al. Automated analysis of in meal eating behavior using a commercial wristband IMU sensor. in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2017. Jeju, Korea: IEEE.
81. Lin, B. and A. Hoover. A Comparison of Finger and Wrist Motion Tracking to Detect Bites during Food Consumption. in 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN). 2019. Chicago, IL, USA.
82. Kramer, L.L., et al., Developing Embodied Conversational Agents for Coaching People in a Healthy Lifestyle: Scoping Review. J Med Internet Res, 2020. 22(2): p. e14058.
83. Ruttkay, Z. and C. Pelachaud, From brows to trust: Evaluating embodied conversational agents. Vol. 7. 2006, Berlin/Heidelberg, Germany: Springer Science & Business Media.
84. Fadhil, A. and S. Gabrielli. Addressing challenges in promoting healthy lifestyles: the al-chatbot approach. in Proceedings of the 11th EAI international conference on pervasive computing technologies for healthcare. 2017.
85. Fox, J., J. Bailenson, and J. Binney, Virtual Experiences, Physical Behaviors: The Effect of Presence on Imitation of an Eating Avatar. Presence-Teleoperators and Virtual Environments, 2009. 18(4): p. 294-303.
86. ter Stal, S., et al., Design Features of Embodied Conversational Agents in eHealth: a Literature Review. International Journal of Human-Computer Studies, 2020. 138: p. 102409.
87. Kramer, L.L., et al., Supporting eating behaviour of community-dwelling older adults: co-design of an embodied conversational agent. Design for Health, 2021: p. 1-20.
88. Kramer, L.L., et al., Use and Effect of Web-Based Embodied Conversational Agents for Improving Eating Behavior and Decreasing Loneliness Among Community-Dwelling Older Adults: Protocol for a Randomized Controlled Trial. JMIR Research Protocols, 2021. 10(1): p. e22186.
89. Schoeppe, S., et al., Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. International Journal of Behavioral Nutrition and Physical Activity, 2017. 14(1): p. 1-10.

90. Baranowski, T., et al., Nutrition education and dietary behavior change games: A scoping review. Games for health journal, 2019. 8(3): p. 153-176.
91. Sardi, L., A. Idri, and J.L. Fernández-Alemán, A systematic review of gamification in e-Health. Journal of biomedical informatics, 2017. 71: p. 31-48.
92. Alahäivälä, T. and H. Oinas-Kukkonen, Understanding persuasion contexts in health gamification: A systematic analysis of gamified health behavior change support systems literature. International journal of medical informatics, 2016. 96: p. 62-70.
93. Johnson, D., et al., Gamification for health and wellbeing: A systematic review of the literature. Internet interventions, 2016. 6: p. 89-106.
94. Nour, M., et al., A narrative review of social media and game-based nutrition interventions targeted at young adults. Journal of the Academy of Nutrition and Dietetics, 2017. 117(5): p. 735-752. e10.
95. Thompson, D., et al., A serious video game to increase fruit and vegetable consumption among elementary aged youth (Squire's Quest! II): Rationale, design, and methods. JMIR research protocols, 2012. 1(2): p. e2348.
96. Thompson, D., Incorporating behavioral techniques into a serious videogame for children. Games for health journal, 2017. 6(2): p. 75-86.
97. Schnall, R., et al., A user-centered model for designing consumer mobile health (mHealth) applications (apps). Journal of biomedical informatics, 2016. 60: p. 243-251.
98. McCurdie, T., et al., mHealth consumer apps: the case for user-centered design. Biomedical instrumentation & technology, 2012. 46(s2): p. 49-56.
99. Ahmed, M., et al., A randomized controlled trial examining consumers' perceptions and opinions on using different versions of a FoodFlip© smartphone application for delivery of nutrition information. International Journal of Behavioral Nutrition and Physical Activity, 2020. 17(1): p. 1-16.
100. López, D., et al., Development and evaluation of a nutritional smartphone application for making smart and healthy choices in grocery shopping. Healthcare informatics research, 2017. 23(1): p. 16.
101. Samoggia, A. and B. Riedel, Assessment of nutrition-focused mobile apps' influence on consumers' healthy food behaviour and nutrition knowledge. Food Research International, 2020. 128: p. 108766.
102. Voedingscentrum. Voedingscentrum-app 'Kies Ik Gezond?'. 2020 [cited 2021; Available from: https://www.voedingscentrum.nl/nl/thema/apps-en-tools-voedingscentrum/kies-ik-gezond-.aspx.
103. Flaherty, S.-J., et al., Can existing mobile apps support healthier food purchasing behaviour? Content analysis of nutrition content, behaviour change theory and user quality integration. Public Health Nutrition, 2018. 21(2): p. 288-298.

104. Flaherty, S.J., et al., A different perspective on consumer engagement: exploring the experience of using health apps to support healthier food purchasing. Journal of Marketing Management, 2019. 35(3-4): p. 310-337.
105. Gilliland, J., et al., Using a smartphone application to promote healthy dietary behaviours and local food consumption. BioMed research international, 2015. 2015.
106. Palacios, C., et al., Effectiveness of the nutritional app "mynutricart" on food choices related to purchase and dietary behavior: A pilot randomized controlled trial. Nutrients, 2018. 10(12): p. 1967.
107. Adams, S.H., et al., Perspective: guiding principles for the implementation of personalized nutrition approaches that benefit health and function. Advances in Nutrition, 2020. 11(1): p. 25-34.
108. Celis-Morales, C., et al., Effect of personalized nutrition on health-related behaviour change: evidence from the Food4me European randomized controlled trial. International journal of epidemiology, 2017. 46(2): p. 578-588.
109. Spruijt-Metz, D., et al., Innovations in the use of interactive technology to support weight management. Current obesity reports, 2015. 4(4): p. 510-519.
110. Terzimehić, N., C. Schneegass, and H. Hußmann. Exploring Challenges in Automated Just-In-Time Adaptive Food Choice Interventions. in Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care. 2017. Mountain View, CA< USA.
111. Nahum-Shani, I., E.B. Hekler, and D. Spruijt-Metz, Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. Health Psychology, 2015. 34(S): p. 1209.
112. Wang, L. and L.C. Miller, Just-in-the-moment adaptive interventions (JITAI): a meta-analytical review. Health communication, 2020. 35(12): p. 1531-1544.
113. Lyzwinski, L.N., A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. Journal of personalized medicine, 2014. 4(3): p. 311-385.
114. Hartmann-Boyce, J., et al., Behavioural Weight Management Review Group. Effect of behavioural techniques and delivery mode on effectiveness of weight management: systematic review, meta-analysis and meta-regression. Obes Rev, 2014. 15(7): p. 598-609.
115. Dombrowski, S.U., et al., Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review. Health Psychology Review, 2012. 6(1): p. 7-32.

116. Sporrel, K., et al., Unraveling Mobile Health Exercise Interventions for Adults: Scoping Review on the Implementations and Designs of Persuasive Strategies. JMIR mHealth and uHealth, 2021. 9(1): p. e16282.
117. Lippke, S. and J.P. Ziegelmann, Theory-based health behavior change: Developing, testing, and applying theories for evidence-based interventions. Applied Psychology, 2008. 57(4): p. 698-716.
118. Looman, M., et al., Supplement use and dietary sources of folate, vitamin D, and n-3 fatty acids during preconception: the GLIMP2 study. Nutrients, 2018. 10(8): p. 962.
119. Ter Borg, S., N. Koopman, and J. Verkaik-Kloosterman, Food consumption, nutrient intake and status during the first 1000 days of life in the netherlands: A systematic review. Nutrients, 2019. 11(4): p. 860.
120. Kind, K.L., V.M. Moore, and M.J. Davies, Diet around conception and during pregnancy–effects on fetal and neonatal outcomes. Reproductive biomedicine online, 2006. 12(5): p. 532-541.
121. Clohessy, S., L. Walasek, and C. Meyer, Factors influencing employees' eating behaviours in the office-based workplace: A systematic review. Obesity Reviews, 2019. 20(12): p. 1771-1780.
122. Jastran, M.M., et al., Eating routines. Embedded, value based, modifiable, and reflective. Appetite, 2009. 52(1): p. 127-136.
123. Lassen, A.D., et al., The impact of worksite interventions promoting healthier food and/or physical activity habits among employees working ‘around the clock’hours: a systematic review. Food & nutrition research, 2018. 62.
124. Bos-de Vos, M. A framework for designing for divergent values. in Proceedings of DRS2020 International Conference: Synergy. 2020. Online: Design Reseach Society.
125. Berezowska, A., et al., Understanding consumer evaluations of personalised nutrition services in terms of the privacy calculus: a qualitative study. Public health genomics, 2014. 17(3): p. 127-140.
126. Berezowska, A., et al., Consumer adoption of personalised nutrition services from the perspective of a risk–benefit trade-off. Genes & nutrition, 2015. 10(6): p. 1-14.
127. Ronteltap, A., et al., Nutrigenomics-based personalised nutritional advice: in search of a business model? Genes & nutrition, 2013. 8(2): p. 153-163.

Chapter 8

General discussion

MAIN FINDINGS

The overarching objective of this thesis was to develop a flexible smartphone-based dietary assessment tool that can be tailored to specific research objectives while also exploring the potential of ecological momentary dietary assessment (EMDA) principles by 2-hour recalls (2hRs): Traqq®. The development of Traqq® was an iterative process that included both expert and user evaluations (**Chapters 2 & 3**), which led to in an innovative app with many tailoring options, see Table 1 for an overview.

Table 1. Overview of Traqq®'s flexible features

Feature	Standard option	Tailoring options
Dietary assessment method	Full-day food record	Food record: - Dayparts (e.g., only office hours) Recall with flexible time periods, e.g.: - 2hR - 4hR - 8hR
Sampling scheme	Manual food record scheduling	Automated recall schemes: - Full-day scheme (i.e., automatic scheduling of predefined number of recall days & rescheduling when predetermined number of prompts is missed on a recall day) - Random scheme (i.e., automatic random scheduling of a predefined number of recalls & rescheduling on same time different day in case of none-response)
Food list	Cleaned version of NEVO 2021	Tailored food lists, e.g.: - Adding of specific items of interest (e.g., sport nutrition, infant foods) - Adding of provided intervention foods - Deleting of irrelevant food items (e.g., focus on snacking, focus on fruit/vegetable intake)
Portion size estimation	Household measures, standard portion sizes, grams	- Only household measures and standard portion sizes - Only grams - No portion size assessment
Additional questions	None	Send additional questions via online survey tool (only for recalls), e.g.: - Food choice motive questions - Contextual/food environment questions - Mood questions - Symptoms related to intake of intervention foods

We assumed that repeated 2hRs instead of a 24-hour recall (24hR) or a food frequency questionnaire (FFQ) would decrease memory-related bias and reporting burden, and increase the accuracy of dietary assessment. In **Chapter 4** we describe the study design and evaluation approach of the use of consecutive 2hRs on one day for assessing *actual* intake (i.e., full 2hR-days), and single 2hRs spread over multiple weeks for assessing *habitual* intake (i.e., random 2hRs). **Chapter 5** shows that three 2hR-days are a reliable approach to assess actual intake of energy, nutrients, and food groups. Differences with validated 24hRs were small and biomarker comparisons showed smaller underestimations of protein and potassium intake for the 2hR-days as compared to 24hRs. **Chapter 6** suggests that random 2hRs (i.e., equivalent of three full days; three times each timepoint) can be a good alternative for regular 24hRs to assess habitual dietary intake as shown by a good ranking ability for energy, most nutrients, and most frequently consumed foods when compared to a FFQ and concentration markers. More variation was seen for episodically consumed foods and nutrients, which may be solved by a more extensive sampling scheme. **Chapters 5 and 6** also showed that the majority of the participants favoured using either 2hR-days or random 2hRs over using traditional methods (i.e., 24hRs, FFQ). Although the current version of Traqq® is well received by respondents (**Chapters 3-6**), points for improvement in terms of accuracy remain (**Chapters 5-6**). Therefore, in **Chapter 7** we explored emerging technologies that may help to improve Traqq®.

CURRENT STATE OF THE ART

Traqq® distinguishes itself from other dietary assessment research-apps by its sound scientific basis, transparent development process, extensive validation, and its wide-ranging tailoring options. Comparable available dietary assessment research apps all have rather fixed designs (Chapter 1). To illustrate, e-CA [1], e-DIA [2, 3], Eat and Track [4, 5], PIQNIQ [6], and Research Food Diary [7] can only be used as a food record, whereas e-12HR [8, 9] can only be used as recall. Moreover, developmental processes of comparable text-entry based apps are often not or minimally described [1-3, 6, 8-10], which hampers their quality assessment [11]. As an exception, the development of the Australian 'Eat and Track (EaT)' app shows important resemblances to the approach for Traqq® (**Chapters 2 & 3**). To illustrate, EaT has also been developed guided by usability testing including think aloud interviews, and assessment of the system usability scale. Moreover, EaT also includes a tailored food list that is compiled after thorough review of the national food consumption database and development of a portion size database [5]. As the development of reliable dietary assessment tools is challenging, it is important to provide insight in tool development and features such data entry, food list, source of food composition data, food quantification options, tailoring options, outputs, and pre-testing [11].

In addition, it is important to validate new tools against established and, preferably, also objective measures to determine whether the new tool or methodology assesses what it is intended to asses. Only a limited number of apps underwent some sort of evaluation [11-14]. Strikingly, none of the comparable apps was validated against objective measures for nutrient or food intake (e.g., recovery biomarkers) [1-4, 6, 8, 9]. PIQNIQ was evaluated during a randomized control trial; participants were asked to report their food intake through PIQNIQ while consuming study menus and handing-in any leftovers (i.e., to determine true intake) [6]. However, although true intake was assessed, this evaluation was conducted in a controlled setting where participants where especially focussed on their intake and reporting [15]. The Traqq®-2hR method was extensively validated in a real-world setting, including objective markers (**Chapters 4-6**). Although validation studies using objective markers are lacking, most available food record apps are evaluated against dietitian-guided 24hRs (i.e., relative validity) [14]. Dietitian-guided 24hRs are perceived as a high quality dietary assessment method, but they are not a true gold standard [16, 17]. Evaluation studies of food record apps mostly show lower intake estimates of energy and macronutrients compared to 24hRs [1, 3, 4, 18-20]. In contrast, our results show mostly higher intake estimates of the 2hR-days as compared to 24hRs. This may relate to the fact that our approach limits memory-related bias owing to the relatively short reporting window of the recall-method (i.e., 2-hours vs. 24-hours). Specifically, with the 2hR, participants register their food intake every two hours of the day and immediately send it to an external server. This also minimizes reactivity bias as recorded intakes are not visible anymore for the participant. With regular food records, food intake reports remain accessible throughout the day, which increases the likelihood of introducing reactivity bias. All in all, 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 to more traditional food record based-apps.

Yet, as Traqq® is extremely flexible, it is important continue with the evaluations following major upgrades related to for instance the food list, or when using other recall variants (e.g., 4hR, 8hR). In addition, as mentioned in **Chapter 7**, new technologies have the potential to improve the current version of Traqq®, and also to make the app more suitable for specific populations (e.g., chatbot integration for individuals with a lower e/mHealth literacy). When exploring new technologies, it is important to involve relevant stakeholders (e.g., target users) and disciplines (e.g., nutrition experts, design experts, IT experts) throughout the development process, to ensure that the final tool is tailored to the needs and requirements of the target population.

METHODOLOGICAL CONSIDERATIONS

DEVELOPMENT PROCESS

We followed a user-centred co-design approach, where both intended end-users and experts were involved throughout the design process. However, the 'user' in this process was more the expert instead of the actual end-user of the app. As a result the design of Traqq® was mostly based on expert input, where end-users evaluated the design via pre-testing and usability evaluations after which the design was adjusted according to their feedback (**Chapters 2 & 3**). This resulted in an simple and easy-to-use app for the general Dutch adult population, with many flexible features in the backend tailored to the needs and preferences of (nutrition) researchers. Still, the app might have been even more accurate if potential end-users would have been able to indicate their needs and preferences as well [21]. It also needs to be stressed that the current version of Traqq® might be less appropriate for specific target groups. Similar to widely-used commercially available food tracker apps, Traqq® relies on an individual's ability to search and enter consumed foods via the food list (i.e., text-entry), [22]. However, a chatbot or barcode scanner might be a more suitable data entry method than text-entry for e.g., (older) adults with functional impairments (e.g., visual and/or motor impairment) and individuals with limited (E-)health literacy [23-26], (**Chapter 7**). Therefore, in future studies aiming to tailor Traqq®, a user-centred co-design approach is recommended that actually puts the targeted end-user in the centre [27, 28]. This approach is not only assumed to benefit app quality, but also trustworthiness of the app and the prospect of adoption of the app by the target group [28-30].

SAMPLING SCHEMES

Although both the 2hRs-days and random 2hRs resulted in a clear overview of frequently consumed foods and nutrients, they lacked details on episodically consumed foods and nutrients, such as fish, legumes, EPA, DHA, and vitamin A (**Chapters 5 & 6**). Both sampling schemes were based on the premises that three full recall days are sufficient to gain insight in habitual intake of commonly consumed foods [31]. In line with previous research, our results show that a more extensive sampling scheme is required to capture the day-to-day variation of episodically consumed foods. However, including more recall days may not automatically result in better coverage of episodically consumed foods [32]. In addition, the required number of days is also variable between individuals as individuals with a stable food pattern have less day-to-day variation than individuals with a highly variable food pattern [31]. Thus even with more recall days, there is still a large possibility to 'miss' episodically consumed foods due to their infrequent consumption. Yet, it certainly will result in an increased burden for the respondents. Another option might be to supplement the recalls with a short food propensity questionnaire (FPQ) that focusses on the intake of episodically consumed food groups (e.g., fish and legumes) [32-34]. Although the FQP is an added burden for the respondent, it can be administered once while an potentially large number or 2hRs is required to ensure coverage of episodically consumed foods. Still, it remains essential to tailor the sampling scheme and dietary assessment method(s) to the research questions and the study population.

PROMPTS

In contrast to web-based 24hR tools, the 2hR approach is not based on the automated multiple-pass method (AMPM). The AMPM is a validated five-step method developed to systematically conduct 24hRs and provides standardized questions and response options [35, 36]. Although this is a valid approach, completing all five steps for short recall periods, including only a limited number of foods per time slot, is too burdensome. However, the AMPM does contain a certain aspect that is now lacking in the 2hR method, namely "the forgotten-foods list", where the respondents is prompted to report frequently omitted foods (e.g., sugar/milk in coffee, fat with cooked meal). Our results show that, similar to other studies, these items are often not reported (**Chapters 5 & 6**) [37, 38]. Integrating additional prompts linked to certain food items and/or eating occasion could remind the respondent to report such items. In turn, this will increase the completeness of reported intake.

GENERALIZATION

Despite extensive validation of our 2hR methodology, where we compared its accuracy against both established self-report methods and independent biological markers (**Chapters 4-6**), these results need to be interpreted with some caution. We aimed to include a representative sample of Dutch adults (**Chapter 4**). However, in our sample most participants were highly educated (58%) and/or women (72%) whereas in 2019, 41% of the general Dutch population was highly educated [39] and 50% of the population were women [40]. Several determinants are known to affect reporting of food intake, i.e., body mass index (BMI), age, sex, socio-economic status, educational level, health-related activities (e.g., smoking, dieting), psychological factors, and eating habits (e.g., high vs. low intake) [41]. Of these determinants, BMI seems to be the most important factor related to misreporting [41-43]. Still, findings are inconclusive and misreporting may occur in all individuals. To illustrate, not all overweight individuals underreport, whereas not all normal weight individuals provide accurate reports [41]. It is better to assume that all self-reported dietary intake data is affected by some degree of measurement error, and to check, and if needed, correct for it by using statistical methods [15, 41]. Examples of such methods are the Goldberg cut-offs [44] or the Willet cut-offs [45] to identify (and exclude) under- and over-reporters, which we both used in our validation studies (**Chapters 5 & 6**). Another statistical method that we used are attenuation factors which give insight in the extent to which diet-disease relations are attenuated by measurement error, e.g., using self-report food intake data instead of true intake. Furthermore, attenuation factors can also be used to correct attenuated diet-disease associations assessed with that specific method [46]. In the end, validity of self-reported dietary assessment depends most on participant instruction, literacy, and commitment to accurate report their food intake, instead of on participant determinants [47].

SELECTION OF DIETARY ASSESSMENT METHOD

The EMDA-based 2hR method has been proven effective to capture both actual and habitual intake of energy, most nutrients and frequently consumed food groups (**Chapters 5 & 6**). Still, 2hR-days may not always be the most appropriate dietary assessment method. When selecting a dietary assessment method there are many factors to take into account, i.e., main objective, required type of information (e.g., actual intake, usual intake), study population (e.g., children, adults, patients), reference period (e.g., week, month, year), comparability to other studies, resources (e.g., budget, expertise) [34]. The 2hR method is flexible and can be tailored to many of these factors. Yet, 2hRs might not be appropriate to all study populations. Older adults are often faced with memory-related limitations which mainly affect short-term memory, making a 2hR less appropriate. Another example are children, who need their parents help to report food intake. Thus, although the 2hR method is well-validated, it does not replace existing methods and is merely an addition. For each new research questions, it is important to select the most appropriate dietary assessment method taking all previously mentioned factors in account.

IMPLICATIONS FOR RESEARCH

TRAQQ®

Traqq®'s main strength lies in its flexibility e.g., of the food list to specific research aims, portion size estimation options, dietary assessment method, and monitoring of the dietary assessment data collection process in the backend. Our evaluation studies already showed that this flexibility significantly decreases researcher burden and participant burden [48], which makes Traqq® a popular new tool for a variety of study types. To illustrate, Traqq® has been used as a food record in a study focussing on shift workers, who followed a nutrition protocol during two or three consecutive night shifts. Provided intervention foods were added to food list to ensure correct reporting of these items, and reporting windows were altered to fit the shift workers unregular schedules. The nutrition protocol started with dinner before the first night shift (i.e., food record opened at 17:00) and lasted until breakfast after either the second or third night shift (food record closed at 9:30), depending on the night shift schedule of the participant.

Traqq® has also been used in an observational study amongst millennials that focussed on the motivations for certain food choices, particularly focussing on eating occasions other than the three main meals (e.g., snacks). As the focus of this study was not on what and how much food was consumed but on the what, where and why, this study benefitted from the flexible Traqq® option to include additional questions related to food intake such as contextual factors. Moreover, Traqq®'s flexibility also allowed tailoring of the food list to also include popular foods that are frequently consumed by millennials (e.g., plant-based drinks, vegan foods, super foods), which were not included in the Dutch food composition database of 2016. A list of important foods, frequently consumed by millennials, was created based on literature research. This list was compared with available foods in Dutch supermarket and on-the-go food vendors (e.g., at train stations). Identified items were added to Traqq®'s food list. As the focus of this study was on what and why certain foods were consumed, we did not need elaborate food composition data. Therefore, we used data from supermarkets to supplement the food list.

The potential of Traqq® is not only limited to nutrition research. Traqq® is also very suitable for use in food environment research, which focusses on understanding how the food environment affects dietary intake [49]. Moreover, although the original version of Traqq® is developed for use in the Netherlands, its flexible nature allows easy tailoring to other cultures. This is in contrast to most tools which are only available in 1 maybe 2 languages, except for myfood24 which is available in English, Danish, French, German, Norwegian, with an Arabic version in development [50]. Currently, we also have a Polish, Swiss-German, Spanish and Greek version of the app (i.e., and an English, French and Belgium version in development). The development of these country-specific versions occurred according to a simplified version of the developmental protocol (**Chapter 3**). As the app itself is finished, and contains limited fixed text, these were easily translated. The main focus was on the development of country-specific food lists with accompanying portion sizes. The original cleaning protocol was used with slight project-specific adjustments. We made use of national databases and the cleaning was conducted by local nutrition experts according to the protocol. Issues were discussed during expert meetings to ensure comparability of final food lists. The quality (i.e., publication date) and extensiveness (i.e., number of foods, nutrients) of national food composition databases (FCDB) varied widely over countries (Table 2) [51]. To ensure comparability of food intake data over countries, high-quality and up-to-date FCDBs are essential.

Table 2. Overview of number of foods in FCDB and publication year per country.

Country	Number of foods	Publication year
Greece	305	2013
Poland	1045	2017
Spain	967	2009
Switzerland	1152	2021
The Netherlands	2207	2021

These five apps (i.e., also including the Dutch version) have been embedded in a large European project that uses 4hRs with additional questions to gain insights in the food environment, which makes Traqq® one of the first dietary assessment tools adapted to multiple cultures [13]. As the collected data needs to be comparable over countries, differences in FCDBs were challenging. To ensure coverage of all foods of interest for this specific project, missing items were added from the Dutch FCDB (e.g., plant-based drinks, vegetarian foods). Moreover, with the exception of Poland, the aforementioned countries did not have country-specific portion size databases available. Therefore, we used a combination of the portions described in the national food-based dietary guidelines [52-54], the Dutch food portion size database [55], and to input of local nutrition experts to compile country-specific food portion sizes.

These are just a few examples of current uses of Traqq®. Yet, Traqq®'s flexible nature allows for application in many research projects and is currently used or being used in close to 20 national and international studies, more requests pending. Most of these studies requested some form of tailoring, mostly regarding the food list, which shows the need for such a tool.

2hRs

In addition to the app Traqq® itself, the 2hR method has great implications for research as well. Our validation studies showed that this EMDA-based method results in more accurate food intake data as compared to established self-report methods. Moreover, EMDA has the potential to also capture additional (intake-related) information to understand determinants of dietary behaviour [56]. EMDA's potential is not limited to short-term nutrition-related research, but can also be applied in longitudinal studies or national food consumption surveys for monitoring long-term dietary intake. Applying the 2hR methodology in such studies has the potential to increase the accuracy of collected food intake data while minimizing participant burden, as compared to frequently used approaches; repeated 24hRs and FFQs [15, 48]. As these methods only capture part of the diet, while a random 2hR scheme has the potential to obtain a full overview of habitual food intake. Full-day recordings, and associated burden, are spread over a long time period. Therefore, we can collect more days of food intake, while the perceived participant burden remains low. However, it needs to be stressed that an adequate sampling scheme is crucial, i.e., ensuring equal coverage of all eating occasions, allowing the assessment of habitual dietary intake. To illustrate, to monitor a participant's food intake over a year, we could do the equivalent of two full 2hR-days per season (i.e., eight full days over a year). To obtain a full day of 2hRs you need approximately eight single 2hRs, i.e., 8 full days results in 64 random 2hRs spread over 365 days, this is less than 6 random 2hRs per month. The sampling scheme would run automatically, striving to complete all scheduled recalls (i.e., each timepoint 8 times). Afterwards, the data can easily by processed for further analyses with minimal burden for both researcher and participant.

In addition to previous example, the 2hR methodology also has potential for use in citizen science-based projects [57]. An example of such a project is the Horizon2020 funded *Big Data Against Childhood Obesity (BigO)* project. BigO is a citizen science project that collect data on children's health behavioural patterns and their living environment via an app [58, 59]. Currently, food intake data is only collected via pictures. Yet, these images are proven to be difficult to process and do not provide accurate insights in these children's dietary behaviours. A modified version of the 2hR methodology with additional questions could allow for a more detailed account of their dietary behaviours and the food environment, which in turn has great potential for nutritional epidemiology [60].

Finally, the 2hR-method also has potential for use in personalized health initiatives. Personalized health, or in this case personalized nutrition, uses data on individual characteristics (e.g., biological, behavioural) to provide targeted nutritional advice to attain lasting behaviour change [61, 62]. Using 2hRs in such personalized nutrition initiatives allows for easy assessment of current dietary behaviour and immediate communication with the respondent [48], thus tailored feedback can be provided [63]. Utilization of other smartphone features (e.g., GPS) or linkage to wearables (e.g., smartwatch), in combination with 2hRs, would allow for more extensive personalized interventions. To illustrate, smartwatches can be used as in signal-contingent EMDA as they are able to detect eating-gestures, which in turn, if connected to Traqq®, could prompt an invite to report current food intake (e.g., shortly after the repeated movement has stopped) [64-66]. Obtained data can be used to automatically and continuously acquire information about the user (e.g., 2hRs for dietary intake) and its context (e.g., GPS for environmental exposures) and deliver individualized interventions based on states of vulnerability and receptivity of the user [67, 68]. Such a personalized intervention is also known as a Just-In-Time-Adaptive Intervention (JITAI). JITAIs adapt their support "overtime to an individual's changing status and contexts", aiming to deliver support "at the moment and in the context that the person needs it most and is most likely to be receptive" [69]. The 2hRs method could be a valuable addition to diet-related JITAIs for the assessment of current dietary behaviours. However, more research is needed to bring this into practice.

FUTURE RESEARCH

More research is needed to further improve dietary assessment efforts, especially with regard to EMDA. We focused on 2hRs to assess dietary intake. However, there are many different recall intervals that may be more appropriate in general, or for specific purposes (e.g., 3hRs, 4hRs, 6hRs). In addition, for each of these options, more research is needed in the development of tailored sampling schemes to capture an entire diet, i.e., include more recall days [31] or add an additional measure such as the FPQ [32]. In addition, more research is needed in use of both Traqq® and EMDA for specific target populations [11, 13].

However, even a perfectly tailored dietary assessment tool, with personalized sampling scheme, still relies on self-report and will remain sensitive for measurement errors. One major source of measurement error in current dietary assessment tools is portion size estimation [48, 70, 71]. New technologies have the potential to decrease these errors. For Traqq®, the integration of a visual portion size estimation aid could be a valuable addition. An example of such an aid is already integrated in the PIQNIQ app which includes a visual portion size selector, i.e., the slider alongside the portion size image can be moved to increase or decrease portion size, and the amount on the plate will change accordingly [6].

Another reporting error relates to unintendedly omitting certain foods (e.g., sugar/milk in coffee, fat in cooked meals). Integration of additional prompts to remind respondents to also report these items could increase completeness of the intake recording. At first, previously collected intake data can be reviewed to identify frequently omitted foods [72]. These results can then be used to integrate prompts for these specific foods, e.g., respondent reports the eating occasion 'dinner'; Traqq® prompts the user to report fat or oil. Eventually Traqq® should be able to learn from previous reports and provide personalized recommendations to ensure data completeness, by using artificial intelligence.

As self-reported dietary intake data will always contain measurement error, ideally, we would like to move away from self-report dietary assessment towards objective food intake monitoring. However, at the moment there is no sensor technology available that is able to automatically and objectively monitor food intake [73, 74]. Still, sensor-based technologies can be a valuable addition to self-report measures [47]. An opportunity worth exploring is linkage to sensor-based wearables such as smartwatches to collect additional data (e.g., heartrate, accelerometry, GPS). The accelerometry data, for instance, can be used to detect eating gestures which in turn can be used to further tailor sampling schemes [64, 65]. This way, dietary intake is assessed shortly after consumption (reducing memory bias) and only when the smartwatch suspects eating behaviour (avoiding unnecessary prompts) [66].

Recommendations for future projects

The main lesson learned from this thesis relate to the design process of a new dietary assessment and/or coaching tools. Although we used an iterative process and included both experts and targeted end-users from early on, improvements for future, similar research remain. In addition, to the guidelines for reporting on new dietary assessment tools as proposed by Eldridge and colleagues [11], we propose a user-centered, multistakeholder approach for the actual development of future tools. The developmental process is guided by four main themes: 1) explore, 2) develop, 3) evaluate, and, if appropriate, 4) implement. This process should be supported by a highly interdisciplinary approach to cover all research needed to ultimately realize technology-driven, personalized, and effective dietary solutions. To illustrate this process, we will describe the further development of Traqq® to improve dietary assessment efforts and also integrate a personalized dietary coaching module. Objectives then include:

1. **Explore:** Determining needs and preferences of targeted end-users and other important stakeholders in terms of required support, tools, and design to establish long-term dietary behaviour changes.
2. **Develop:** Co-create, test, and integrate new techniques to improve dietary assessment and develop novel personalized dietary support tools with underlying ICT platforms.
3. **Evaluate:** Conduct evaluation studies to:
   a. validate the improved dietary assessment tools.
   b. evaluate the effect of integrated behaviour change interventions.
4. **Implement:** Map the steps to be taken for successful implementation of the developed tools in daily life, healthcare, and nutrition and sustainability research.

Developing, testing, and (possible) implementing novel, technology-driven, solutions to support individuals to improve and maintain a healthy diet, required an interdisciplinary team of researchers. This team should include researchers in nutrition, dietetics, AI, ICT, design, behavioural and social sciences.

Theme 1 - Explore needs to be driven by an interdisciplinary team of designers, nutrition scientists, and behavioural scientists. Perspectives, needs and preferences of targeted end-users need to be mapped. To facilitate the adoption of the dietary support tool and empower end-users towards long-term dietary behaviour change, perspectives, needs, and preferences of other relevant stakeholders need to be mapped as well (e.g., dieticians, health care providers, supermarkets, municipal health institutes, health insurance companies). Thus, theme 1 should consist of a mix of quantitative research (i.e., self-administered (online) questionnaires) and qualitative research (e.g., focus groups, in-depth interviews, journey mapping and/or Delphi rounds) to gain insight in various design directions to improve Traqq® and develop dietary behaviour change interventions.

Theme 2 - Develop needs to be driven by designers, software developers, data scientists, nutrition scientists, and behaviour scientists, and focus on the improvement of Traqq® by developing, testing, and implementing novel techniques. Promising techniques need to be explored that can improve dietary assessment efforts in terms of accuracy, reporting burden, and motivation for long-term use (e.g., use of images, AI/machine learning, smartwatch integration, persuasive game design). Additional focus is needed on the development of novel behaviour change interventions using desk-top research and co-creation sessions with target users. Behaviour change solutions need to be able to be integrated directly in Traqq®, which will result in an all-round dietary support tool including dietary assessment module and customizable dietary behaviour change features. In parallel, a supporting ICT infrastructure needs to be created to facilitate use of the tool.

Theme 3 - Evaluate needs to be driven by nutrition, behavioural, and social scientists. A solid dietary intake assessment is crucial for well-founded personalized dietary behaviour change interventions, dietary monitoring efforts, and nutrition research. Therefore, quantitively evaluation (e.g., usability, self-report dietary intake) and validation (biochemical markers by metabolomics) of the upgraded version of Traqq® is essential. Additional focus should be on examining whether the integrated dietary support tool(s) effectively empower citizens to obtain and maintain a healthy diet. Preferably, this is evaluated in a randomized control trial.

Theme 4 - Implement needs to be driven by nutrition, behavioural, social, and design scientists. All partners need to work together to optimize the chance for successful adoption of the tool by both end-users and health care professionals. To achieve this, the steps to be taken to ensure a successful implementation of the dietary support tool in daily life and health care need to be mapped. Relevant stakeholders as determined in theme 1 need to be interviewed and required steps need to be incorporated in an implementation plan.

The described process will result in a full-fledged tool, tailored to the needs and requirements of a specific target population. To ensure transparency of the development of future tools, publication of the developmental process in addition to evaluation results is eminent.

OVERALL CONCLUSIONS

The work described in this thesis showed the development of the dietary assessment app Traqq® and the potential of the EMDA-based 2hR methodology for accurate assessment of dietary behaviours. The iterative developmental process resulted in a flexible dietary assessment app that can be tailored to different research aims. In addition, Traqq® was used to evaluate the 2hR methodology against established dietary assessment methods and independent markers of intake. We showed that 2hR-days result in a more accurate assessment of food intake as compared to traditional 24hRs and that random 2hRs are able to rank participants according to intake of frequently consumed foods, similar to an FFQ. Overall, we showed that both Traqq® as well as the 2hR methodology have great potential for use in nutrition-related research.

REFERENCES

1. Bucher Della Torre, S., et al., Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. Nutrients, 2017. 9(1).
2. Rangan, A.M., 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 and uHealth, 2015. 3(4): p. e4613.
3. Rangan, A.M., et al., Electronic Dietary Intake Assessment (e-DIA): relative validity of a mobile phone application to measure intake of food groups. British Journal of Nutrition, 2016. 115(12): p. 2219-2226.
4. Wellard-Cole, L., 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, 2019. 11(3): p. 621.
5. Wellard-Cole, L., et al., A Tool to Measure Young Adults' Food Intake: Design and Development of an Australian Database of Foods for the Eat and Track Smartphone App. JMIR Mhealth Uhealth, 2018. 6(11): p. e12136.
6. Blanchard, C.M., et al., Evaluation of PIQNIQ, a Novel Mobile Application for Capturing Dietary Intake. J Nutr, 2021. 151(5): p. 1347-1356.
7. Ambrosini, G.L., et al., Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr J, 2018. 17(1): p. 5.
8. Bejar, L.M., et al., Relative Validity of a Method Based on a Smartphone App (Electronic 12-Hour Dietary Recall) to Estimate Habitual Dietary Intake in Adults. JMIR Mhealth Uhealth, 2019. 7(4): p. e11531.
9. Bejar, L.M., O.A. Reyes, and M.D. Garcia-Perea, Electronic 12-Hour Dietary Recall (e-12HR): Comparison of a Mobile Phone App for Dietary Intake Assessment With a Food Frequency Questionnaire and Four Dietary Records. JMIR Mhealth Uhealth, 2018. 6(6): p. e10409.
10. Barchitta, M., et al., Development of a Web-App for the Ecological Momentary Assessment of Dietary Habits among College Students: The HEALTHY-UNICT Project. Nutrients, 2022. 14(2).
11. Eldridge, A.L., et al., Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients, 2018. 11(1).
12. Schembre, S.M., et al., Mobile ecological momentary diet assessment methods for behavioral research: systematic review. JMIR mHealth and uHealth, 2018. 6(11): p. e11170.

13. Khazen, W., et al., Rethinking the Use of Mobile Apps for Dietary Assessment in Medical Research. J Med Internet Res, 2020. 22(6): p. e15619.
14. Zhang, L., et al., A systematic review and meta-analysis of validation studies performed on dietary record apps. Advances in Nutrition, 2021. 12(6): p. 2321-2332.
15. Subar, A.F., et al., Addressing current criticism regarding the value of self-report dietary data. The Journal of nutrition, 2015. 145(12): p. 2639-2645.
16. Thompson, F.E. and A.F. Subar, Dietary assessment methodology. Nutrition in the Prevention and Treatment of Disease, 2017: p. 5-48.
17. Touvier, M., et al., Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. British Journal of Nutrition, 2011. 105(7): p. 1055-1064.
18. Carter, M.C., et al., 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr, 2013. 109(3): p. 539-46.
19. Ocké, M., et al., Relative Validity of mijneetmeter: A food diary app for self-monitoring of dietary intake. Nutrients, 2021. 13(4): p. 1135.
20. Lemacks, J.L., K. Adams, and A. Lovetere, Dietary intake reporting accuracy of the bridge2u mobile application food log compared to control meal and dietary recall methods. Nutrients, 2019. 11(1): p. 199.
21. Chang, T.-R., E. Kaasinen, and K. Kaipainen. What influences users' decisions to take apps into use? A framework for evaluating persuasive and engaging design in mobile apps for well-being. in Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia. 2012.
22. Franco, R.Z., et al., Popular nutrition-related mobile apps: a feature assessment. JMIR mHealth and uHealth, 2016. 4(3): p. e5846.
23. Illner, A.K., et al., Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol, 2012. 41(4): p. 1187-203.
24. Peng, W., et al., A qualitative study of user perceptions of mobile health apps. BMC Public Health, 2016. 16(1): p. 1158.
25. Bickmore, T.W., et al. Managing chronic conditions with a smartphone-based conversational virtual agent. in Proceedings of the 18th International Conference on Intelligent Virtual Agents. 2018.
26. Bol, N., N. Helberger, and J.C. Weert, Differences in mobile health app use: a source of new digital inequalities? The Information Society, 2018. 34(3): p. 183-193.
27. Preece, J., H. Sharp, and Y. Rogers, Interaction design: beyond human-computer interaction. 2015, Hoboken, NJ, USA: John Wiley & Sons.

28. Harrington, C.N., et al. Designing health and fitness apps with older adults: Examining the value of experience-based co-design. in Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare. 2018.
29. Chen, J., et al., Designing health apps to support dietetic professional practice and their patients: qualitative results from an international survey. JMIR mHealth and uHealth, 2017. 5(3): p. e6945.
30. Flaherty, S.J., et al., Exploring engagement with health apps: the emerging importance of situational involvement and individual characteristics. European Journal of Marketing, 2021.
31. Willett, W.C., Nutritional Epidemiology. 2013, New York: Oxford University Press Inc
32. Subar, A.F., et al., The food propensity questionnaire: concept, development, and validation for use as a covariate in a model to estimate usual food intake. J Am Diet Assoc, 2006. 106(10): p. 1556-63.
33. Smiliotopoulos, T., E. Magriplis, and A. Zampelas, Validation of a Food Propensity Questionnaire for the Hellenic National Nutrition and Health Survey (HNNHS) and Results on This Population's Adherence to Key Food-Group Nutritional Guidelines. Nutrients, 2020. 12(6): p. 1808.
34. Ocké, M.C., J.H.M. de Vries, and P.J.M. Hulshof, Assessment of dietary intake by self-reports and biological markers, in Present Knowledge in Nutrition. 2020. p. 249-265.
35. Blanton, C.A., et al., The USDA Automated Multiple-Pass Method accurately estimates group total energy and nutrient intake. The Journal of nutrition, 2006. 136(10): p. 2594-2599.
36. Moshfegh, A.J., et al., The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. The American journal of clinical nutrition, 2008. 88(2): p. 324-332.
37. Chan, V., et al., Using wearable cameras to assess foods and beverages omitted in 24 hour dietary recalls and a text entry food record app. Nutrients, 2021. 13(6): p. 1806.
38. Kirkpatrick, S.I., et al., Performance of the Automated Self-Administered 24-hour Recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. The American journal of clinical nutrition, 2014. 100(1): p. 233-240.
39. CBS. Onderwijs. 2020 [cited 2022 12 August]; Available from: https://digitaal.scp.nl/ssn2020/onderwijs/.
40. CBS. Mannen en vrouwen. 2020 [cited 2022 12 August]; Available from: https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/verdeling/#:~:text=Dit%20loopt%20op%20met%20de,op%20elke%20100%20vrouwen%20zijn.

41. Poslusna, K., et al., Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. British Journal of Nutrition, 2009. 101(S2): p. S73-S85.
42. Trijsburg, L., et al., BMI was found to be a consistent determinant related to misreporting of energy, protein and potassium intake using self-report and duplicate portion methods. Public health nutrition, 2017. 20(4): p. 598-607.
43. Freisling, H., et al., Dietary reporting errors on 24 h recalls and dietary questionnaires are associated with BMI across six European countries as evaluated with recovery biomarkers for protein and potassium intake. British journal of nutrition, 2012. 107(6): p. 910-920.
44. Black, A.E., 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. International journal of obesity, 2000. 24(9): p. 1119-1130.
45. Rhee, J.J., et al., Comparison of methods to account for implausible reporting of energy intake in epidemiologic studies. American journal of epidemiology, 2015. 181(4): p. 225-233.
46. Kipnis, V., et al., Bias in dietary-report instruments and its implications for nutritional epidemiology. Public health nutrition, 2002. 5(6a): p. 915-923.
47. Mattes, R.D., et al., Valuing the diversity of research methods to advance nutrition science. Advances in Nutrition, 2022. 13(4): p. 1324-1393.
48. Das, S.K., et al., Perspective: Opportunities and Challenges of Technology Tools in Dietary and Activity Assessment: Bridging Stakeholder Viewpoints. Adv Nutr, 2022. 13(1): p. 1-15.
49. Kirkpatrick, S.I., et al., Dietary assessment in food environment research: a systematic review. Am J Prev Med, 2014. 46(1): p. 94-102.
50. myfood24. Frequently Asked Questions - What languages is myfood24 available in? 2022 [cited 2022 16 August]; Available from: https://www.myfood24.org/faqs.
51. Finglas, P.M., R. Berry, and S. Astley, Assessing and improving the quality of food composition databases for nutrition and health applications in Europe: the contribution of EuroFIR. Advances in Nutrition, 2014. 5(5): p. 608S-614S.
52. Montagnese, C., et al., European food-based dietary guidelines: a comparison and update. Nutrition, 2015. 31(7-8): p. 908-915.
53. Herforth, A., et al., A global review of food-based dietary guidelines. Advances in Nutrition, 2019. 10(4): p. 590-605.
54. FAO. Food-based dietary guidelines. 2022 [cited 2022 30 Agust]; Available from: https://www.fao.org/nutrition/education/food-dietary-guidelines/home/en/.
55. RIVM. Portion-Online. 2020 04-02-2020 [cited 2022 30 August]; Available from: https://portie-online.rivm.nl/.

56. Maugeri, A. and M. Barchitta, A Systematic Review of Ecological Momentary Assessment of Diet: Implications and Perspectives for Nutritional Epidemiology. Nutrients, 2019. 11(11).
57. Dijksterhuis, G., R. de Wijk, and M. Onwezen, New Consumer Research Technology for Food Behaviour: Overview and Validity. Foods, 2022. 11(5): p. 767.
58. Delopoulos, A. Big Data Against Childhood Obesity, the BigO Project. in 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS). 2019. IEEE.
59. Tufford, A.R., et al., Towards systems models for obesity prevention: a big role for big data. Current Developments in Nutrition, 2022.
60. Morgenstern, J.D., et al., Perspective: big data and machine learning could help advance nutritional epidemiology. Advances in Nutrition, 2021. 12(3): p. 621-631.
61. Ordovas, J.M., et al., Personalised nutrition and health. Bmj, 2018. 361.
62. Kraemer, K., et al., Personalized Nutrition: Paving the way to better population health, in Good Nutrition: Perspectives for the 21st Century. 2016, Karger Publishers. p. 235-248.
63. Sharp, D.B. and M. Allman-Farinelli, Feasibility and validity of mobile phones to assess dietary intake. Nutrition, 2014. 30(11-12): p. 1257-66.
64. Bertrand, L., N. Cleyet-Marrel, and Z. Liang, Recognizing Eating Activities in Free-Living Environment Using Consumer Wearable Devices. Engineering Proceedings, 2021. 6(1): p. 58.
65. Stankoski, S., et al., Smartwatch-based eating detection: Data selection for machine learning from imbalanced data with imperfect labels. Sensors, 2021. 21(5): p. 1902.
66. Goldstein, S.P., et al., Combining ecological momentary assessment, wrist-based eating detection, and dietary assessment to characterize dietary lapse: A multi-method study protocol. Digital Health, 2021. 7: p. 2055207620988212.
67. Terzimehić, N., C. Schneegass, and H. Hußmann. Exploring Challenges in Automated Just-In-Time Adaptive Food Choice Interventions. in Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care. 2017.
68. Nahum-Shani, I., E.B. Hekler, and D. Spruijt-Metz, Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. Health psychology, 2015. 34(S): p. 1209.
69. Spruijt-Metz, D., et al., Innovations in the use of interactive technology to support weight management. Current obesity reports, 2015. 4(4): p. 510-519.
70. Cade, J.E., et al., DIET@ NET: Best Practice Guidelines for dietary assessment in health research. BMC medicine, 2017. 15(1): p. 1-15.
71. Timon, C.M., et al., A comparison of food portion size estimation by older adults, young adults and nutritionists. The journal of nutrition, health & aging, 2018. 22(2): p. 230-236.

72. Chen, J., et al., The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. Nutrition, 2019. 57: p. 208-216.
73. Usman, M. and H. Chen, Recent trends in food intake monitoring using wearable sensors. arXiv preprint arXiv:2101.01378, 2021.
74. Skinner, A., et al., Future directions for integrative objective assessment of eating using wearable sensing technology. Frontiers in Nutrition, 2020: p. 80.

Summary

Accurate dietary assessment is an essential aspect of nutrition research and health care. Yet, current methods have a range of drawbacks. **Chapter 1** provides an overview of most commonly-used dietary assessment methods, their limitations and sources of measurement error. In addition, promising innovations are described, including dietary assessment apps and ecological momentary dietary assessment (EMDA). This thesis describes the development and evaluation of a flexible smartphone-based EMDA tool 'Traqq®' that can be tailored to specific research objectives.

The first step in the development of Traqq® was the evaluation of portion size estimation aids that were suitable for integration in a smartphone app (**Chapter 2**). Literature review resulted in two possible options: text-based portion size assessment and image-based portion size assessment. Text-based portion size assessment refers to the use of textual descriptions of portion sizes (e.g., household measures, standard portion sizes), whereas image-based portion size assessment refers to the use of a series of food images to assess portion size. We evaluated the accuracy of both portion size estimation aids in a pilot study. Participants (n=40) consumed an ad libitum lunch (i.e., unlimited amount) after which true intake was assessed. Next, participants estimated consumed amounts using both methods. The results of this study clearly indicated that, of the two, text-based portion size assessment resulted in a more accurate estimation of consumed amounts. These findings led to the integration of text-based portion size estimation options in the app, namely estimation in household measures (e.g., cups, spoons) and standard portion sizes (small, large).

The steps described in the protocol and the results of the evaluation studies resulted in a final design for the app and the backend (**Chapter 3**). We strived to develop a flexible dietary assessment app that can be tailored to different research purposes. A first wish was flexibility in dietary assessment methods. As a result, Traqq® can be used as a food record as well as a recall. The food record-module is similar as previously described apps whereas the recall-module is completely new (**Chapter 1**). Moreover, based on EMDA principles, Traqq® not only facilitates traditional recalls (i.e., 24hRs) but shorter recall periods as well (e.g., 2h, 4h, 8h). Moreover, reporting periods are flexible for both methods, therefore, the food record-module also allows reporting during parts of the day (e.g., during office hours). To further increase flexibility, for the recall-module, it is possible to create different sampling schemes to facilitate tailored data collection. Push notifications invite respondents to record their food intake, and, to ensure complete data collection, invitations are automatically rescheduled in case of non-response. Unfortunately, automatic rescheduling is not possible for the food record-module as one recording (e.g., one cup of coffee) already counts as a response, whereas this is far from being a complete recording. Hence, for food records, manual monitoring of data collection remains needed. In nutrition research, research questions often go beyond merely what is being consumed (e.g., food choice motivation, mood during eating, environmental context). Therefore, we also included an option to connect with online survey tools to send additional questions after a recall. These questions can be general (e.g., where, with whom) or related to specific eating occasions or food items (e.g., motivation, symptoms). Another flexible aspect is the integration of the food list. The standard food list is a cleaned version of the Dutch food composition database (i.e., originally NEVO 2016, currently updated to NEVO 2021), optimized for use in an app. However, as flexibility is important, the backend allows for creating a new food list for each project. Thus, food lists can be tailored to specific research questions and/or target populations.

Following the development of the app, we designed an extensive validation study that allowed for validation of both actual and habitual intake, against established dietary assessment methods and independent biological markers: 'DIASS' (**Chapter 4**). The DIASS study (*n*=215) had a cross-over design and lasted a total of 12 weeks: week 1 – randomization, week 2-5 – study period 1, week 6-7 – wash-out, week 8-11 – study period 2, and week 12 – evaluation. DIASS had two experimental conditions; i.e., measuring *actual* intake and *habitual* intake. We assessed actual food intake on random non-consecutive days over one of the four-week study periods. Within this period, participants were invited to complete three 2hR-days (i.e., full days of consecutive 2hRs) and three web-based or telephone-based 24hRs. In addition, a random subsample of participants also provided four 24-h urine samples and two fasting blood samples. The urine collections were intentionally coupled to the recall days (i.e., 2× to 2hR-day and 2× to 24hR-day). Blood sampling occurred following two of the urine collections while the participants were at the study centre to hand in their urine containers. During the other four-week study period, we assessed habitual food intake by random 2hRs. The same number of 2hRs was used as for the 2hR-days. However, to assess habitual intake, the 2hRs were randomly distributed over the four-week period (i.e., different days and times, 3× each time slot). In addition, at the end of the study period, we invited participants to complete a web-based FFQ. In the final study week, we invited participants to complete the Eetscore™, to assess overall diet quality.

Data from the DIASS study (*n*=146) was used to explore the ability of repeated 2hRs on one day to assess *actual* intake, i.e., consecutive 2hRs from waking to sleeping, with a night-time recall the following morning. We found that, compared to traditional 24hRs, 2hR-days resulted in slightly higher intake estimates of energy, most nutrients, and most food groups (**Chapter 5**). As it is well-known that 24hRs tend to underestimate intake, this was perceived as a positive result. Comparisons of both methods to urinary biomarkers for protein and potassium intake, showed less underestimation of the 2hR-days compared to the 24hRs: protein: -14% vs. -18%; potassium: -11% vs. -16%.

In addition, data from DIASS (*n*=141) was also explored the ability of repeated random 2hRs for to assess *habitual* food intake, i.e., repeated 2hRs on random days and times, with a night-time recall, linked to a final evening 2hR, the following morning (**Chapter 6**). We found that, random 2hRs showed good ranking ability for energy, most nutrients, and most frequently consumed foods, compared to a traditional validated FFQ and blood concentration markers (carotenoids n-3 polyunsaturated fatty acids).

Although the current version of Traqq® is well received by respondents (**Chapters 3-6**), improvements in terms of accuracy remain (**Chapters 5-6**). Therefore, we explored emerging technologies that can help further improve Traqq®, but also other existing tools (**Chapter 7**). In this chapter, we distinguish between tools and technologies focussed on dietary intake (i.e., what is consumed) and eating behaviour (i.e., how it is consumed). Moreover, we not only focus on assessment of these dietary behaviours, but also explore technologies that have the potential to improve personalized dietary behaviour interventions. Potential technologies to improve dietary intake assessment include images (i.e., food identification, portion size assessment), spectroscopy (i.e., food identification, chemical fingerprint), and conversational agents (i.e., alternative reporting, personalized reminders). Conversational agents also have potential to improve personalized dietary behaviour interventions (i.e., 24/7 coaching), as well as game-elements (i.e., enjoyment), targeted interventions (e.g., during food shopping), and tailoring (i.e., personalized advice).

In **Chapter 8**, the main findings of this thesis were summarized and discussed. Methodological aspects related to generalization of findings, development process, sampling schemes and selection of a dietary assessment method/tool are discussed. Furthermore, implications for research are given. The chapter ends with lessons learned translated to recommendations for future research.

To conclude, this thesis describes the development of the flexible dietary assessment app Traqq® and the validation of the 2hR methodology for assessment of dietary intake. We found that the app was well received by participants and favoured over traditional dietary assessment methods. Moreover, we showed that both 2hR-days and random 2hRs are a viable method to assess dietary intake. It is important to mention that measurement errors decreased but remain due to the self-report nature of the 2hR recall method. Therefore, more research is needed to further improve measures of dietary intake.

Acknowledgements

Thank you!

Of course all of this work would not have been possible without a supportive environment. So many people have contributed to this thesis: colleagues, participants, friends and family. Yet, this thesis is already so long, and I could add many more pages thanking everyone in person and then probably still forget some. Therefore, to all of you, I am so grateful for your contribution these past years, thank you so much, I could not have done this without your support. Still, there are some persons that made a special contribution which I would like to mention in particular:

First of all, I would like to thank my promotor Edith Feskens and co-promotors Elske Brouwer-Brolsma and Monica Mars. Thank you for your support these last years, not just professionally but also personally. Thank you for giving me the freedom to make this project my own and supporting me throughout. Each of you has been, and still is, a great mentor and I have learned so much from you.

Additionally, Els, Corine, Karin, Hanne, Anne W, Mariëlle, Janneke, and Jeanne, with whom I started 10 years ago in Biotechnion. Because of you, returning to the division in 2017, felt like coming home. Thank you for your support throughout the years. Each of you has always been willing to think along and helped me out so much (and still do!). In addition, I want to give special thanks to Anne S, who took over from me when I was busy becoming a mom. I can honestly say that I could not have done this without you all!

Also, Arvind, you are the one who made Traqq come to life. Thank you for the smooth collaboration and making our (crazy) ideas come to life. Traqq already has become so much more than we could have ever imagined, and we are not even finished.

Finally, a special thanks to my greatest loves, Mila and Maurice. Maurice jij bent mijn grootste steun en toeverlaat geweest, meteen vanaf het moment dat we elkaar ontmoet hebben. Bedankt voor je liefde, geduld en begrip de laatste jaren. Mila, ik ben zo blij dat jij in ons leven gekomen bent. Je bent zo'n lieve kleine wijsneus, en je weet altijd een glimlach op mijn gezicht te toveren. Als laatste de vrouw die mijn grootste supporter is geweest vanaf het moment dat ik geboren ben:

Oma, deze is voor jou

About the author

CURRICULUM VITAE

Desiree Lucassen was born on March 11, 1986 in Zevenaar, the Netherlands. From 2007-2011 Desiree studied Nutrition and Dietetics at the HAN University of Applied Sciences in Nijmegen, where she specialized in sports nutrition followed by an internship with Anja van Geel at SMC Papendal and Anouska van der Zee at Voedjesport. Yet, she found her true passion during her BSc thesis for which she worked on the B-PROOF study at the division of Human Nutrition and Health and investigated the association between metformin use and the vitamin B12- and folic acid status of elderly. After finalizing her BSc, Desiree started working at the division of Human Nutrition and Health as a research dietician. From 2012 to 2015 Desiree worked on multiple dietary controlled trials, dietary intervention studies and food consumption studies. In 2014, she started with the MSc program Health Education and Promotion at Maastricht University. For her MSc thesis, she focused on influencing feelings of colorectal cancer risk using narrative risk information and self-affirmation. In 2016, she obtained her MSc degree with the distinction cum laude. After a 2-year break from the university she returned in 2017 as a PhD researcher. During her PhD she focused on the development and evaluation of a flexible ecological momentary dietary assessment app, of which the results are presented in this thesis. In 2021, she was the winner of the Foppe ten Hoor award during the Dutch Nutritional Science Days. Alongside her PhD project, Desiree was involved in many other projects related to dietary assessment (innovations). Moreover, since 2018, Desiree has been a member of the Dietary Assessment and Eating Behaviour (DAEB) group, for which she advices both internal and external researchers on the integration of dietary assessment in their research projects. From 2020-2022 Desiree also joined the Nutri_Lab team of the Department of Industrial Design of Technical University Eindhoven (TU/e) as a postdoctoral researcher. This position was created to advance the collaboration between WUR and TU/e to further innovate dietary assessment efforts by exploring new technological opportunities. Currently, Desiree is still working at the division of Human Nutrition and Health as a nutritional researcher and continues her work on innovations for dietary assessment and also improve dietary behaviours.

LIST OF PUBLICATIONS

PUBLISHED ARTICLES

Tufford, A.R., Diou, C., **Lucassen, D.A.**, Ioakimidis, I., O'Malley, G., Alagialoglou, L., Charmandari, E., Doyle, G., Filis, K., Kassari, P., Kechadi, T., Kilintzis, V., Kok, E., Lekka, I., Maglaveras, N., Pagkalos, I., Papapanagiotou, V., Sarafis, I., Shahid, A., Van 't Veer, P., Delopolous, A., Mars, M. (2022) Towards Systems Models for Obesity Prevention: A Big Role for Big Data. *Current Developments in Nutrition*, 6(9), nzac123, https://doi.org/10.1093/cdn/nzac123

Faessen, J.P.M., **Lucassen, D.A.,** Buso, M.E.C., Camp, G., Feskens, E.J.M., Brouwer-Brolsma, E.M. (2022) Eating for Two: A Systematic Review of Dutch App Stores for Apps Promoting a Healthy Diet During Pregnancy. *Current Developments in Nutrition.* doi.org/10.1093/cdn/nzac087

Lucassen, D.A., Brouwer-Brolsma, E.M., Slotegraaf, A., Kok, E., Feskens, E.J.M. (2022) DIetary ASSessment study (DIASS): An of an Evaluation study to Assess Validity, Usability and Perceived Burden of an Innovative Smartphone-based Methodology. *Nutrients.* 14(6), 1156. doi.org/10.3390/nu14061156

Lucassen, D.A., Lasschuijt, M.P., Camps, G., Van Loo, E.J., Fischer, A.R.H., de Vries, R.A.J., Haarman, J.A.M., Simons, M., de Vet, E., Bos-de Vos, M., Pan, S., Ren, X., de Graaf, K., Lu, Y., Feskens, E.J.M., Brouwer-Brolsma, E.M. (2021) Short and Long-Term Innovations on Dietary Behavior Assessment and Coaching: Present Efforts and Vision of the Pride and Prejudice Consortium. *International Journal of Environmental Research and Public Health.* 18(15), 7877. doi.org/10.3390/ijerph18157877

Wit, R.F., **Lucassen, D.A.,** Beulen, Y.H., van Dongen, J.M., Bouwer-Brolsma, E.M. (2021) Midwives' experiences with and perspectives on (online) nutritional coaching and mHealth apps in pregnant women; an explorative qualitative study. *International Journal of Environmental Research and Public Health*. 18(13), 6733. doi:10.3390/ijerph18136733

Lucassen, D.A., Willemsen, R.F., Geelen, A., Brouwer-Brolsma, E.M., Feskens, E.J.M. (2021) The accuracy of portion size estimation using food images and textual descriptions of portion sizes. *Journal of Human Nutrition and Dietetics*. 34(6), 945-952. doi:10.1111/jhn.12878

Lucassen, D. A., Brouwer-Brolsma, E. M., van de Wiel, A. M., Siebelink, E., Feskens, E. J. M. (2021) Iterative Development of an Innovative Smartphone-Based Dietary Assessment Tool: Traqq. *Journal of Visualized Experiments*, (169), e62032, doi:10.3791/62032

Brouwer-Brolsma, E.M., **Lucassen, D.**, de Rijk, M.G., Slotegraaf, A., Perenboom, C., Borgonjen, K., Siebelink, E., Feskens, E.J.M., de Vries, J.H.M. (2020) Dietary Intake Assessment: From Traditional Paper-Pencil Questionnaires to Technology-Based Tools. In: Athanasiadis I., Frysinger S., Schimak G., Knibbe W. (eds) *Environmental Software Systems. Data Science in Action. ISESS 2020. IFIP Advances in Information and Communication Technology*, vol 554. Springer, Cham. doi:10.1007/978-3-030-39815-6_2

EXPECTED ARTICLES

Lucassen, D.A., Brouwer-Brolsma, E.M., Boshuizen H.C., Mars, M., de Vogel-Van den Bosch, J., Feskens, E.J.M. Validation of the smartphone-based dietary assessment tool 'Traqq' for assessing actual dietary intake by repeated 2-hour recalls in adults: comparison with 24h recalls and urinary biomarkers. *Submitted*

Lucassen, D.A., Brouwer-Brolsma, E.M., Boshuizen H.C., Balvers M., Feskens, E.J.M. Evaluation of the smartphone-based dietary assessment tool 'Traqq' for assessing habitual dietary intake by random 2-hour recalls in adults: comparison with FFQ and blood concentration biomarkers. *In preparation*

Lucassen, D.A., Vaes, A., Kennes, L., Wagemakers, A., Kalinauskaite, I., Feskens, E.J.M., Brouwer-Brolsma, E.M. Traqq®-Z. An evaluation of dietary intake assessment, usability and user perspectives of the smartphone application "Traqq®" among Dutch adolescents aged 12-18 years: Study Design. *In preparation*

van der Heijden, Z.S., de Gooijer, F.J., Camps, G., **Lucassen, D.A.,** Feskens, E.J.M., Lasschuijt, M.P., Brouwer-Brolsma, E.M. Identifying key criteria for a novel dietary assessment tool for use among children (5-6 years): an exploratory study. *In preparation*

van der Heijden, Z.S., **Lucassen, D.A.,** Faessen, J.P.M., Schipper, H.S., Nijhof, S.L., Lu, Y., Birk, M.V., Colombo, S., Feskens, E.J.M., Brouwer-Brolsma, E.M. Using digital technologies to promote a healthy diet among children and adolescents: a systematic review. *In preparation*

OVERVIEW OF COMPLETED TRAINING ACTIVITIES

Discipline specific activities	Organizer (location)	Year
Courses		
Exposure Assessment in Nutrition Research	VLAG (Wageningen, NL)	2018
Introduction to Big Data	Coursera (online)	2018
FFQ-tool training	WUR (Wageningen, NL)	2018
SPADE workshop “Modelling of habitual dietary intake”	RIVM (Wageningen, NL)	2020
Workshop “Market validation for digital products and services”	WDCC (online)	2021
Workshop “Applying Data Science and Artificial Intelligence in your research”	WDCC (online)	2021
Workshop “Hands-on with machine learning/artificial intelligence”	WDCC (online)	2022
Conferences and meetings		
Seminar Personalized Nutrition and Health “Hoe ver zijn we?”	WUR/TNO (Wageningen, NL)	2018
Interactive lecture “How the food environment influences food choices?”	Edema Steernberg Foundation (Wageningen, NL)	2018
Digital Society Conference	VSNU (Amersfoort, NL)	2018
Food Intake meetings “IDAT”	TiFN (Wageningen, NL)	2018
BigO consortium meetings	AUTH - BigO consortium, (Athens & Thessaloniki, GR)	2018-20
Smart Food Intake consortium meetings	SFI consortium (Wageningen, Utrecht & Den Haag, NL)	2018-22
Advancing Behavior Data Science Conference	ISBNPA (Praag, CZ)	2019
Webinar SIG e- & mHealth “Behavioural implications of using smartphone apps with food image recognition capability”	ISBNPA (online)	2020
Webinar “Haal alles uit je beweeg- of slaapstudie met Actigraph accelerometers”	Pro-Care (online)	2020
European and International Congress on Obesity	ECOICO (online)	2020

Webinar “Ontdek als MedTech startup waarom wetenschappelijk onderzoek belangrijke rol speelt bij CE markering en market access”	Health Valley (online)	2020
NutriLab meetings	TU/e (Eindhoven, NL)	2020-21
Digital Society Conference “AI in Healthcare - Preparing for the future”	VSNU (online)	2020
International Conference on Diet and Activity Methods	ICDAM (online)	2021
Nutrition 2021 Live	ASN (online)	2021
Dutch Nutritional Science Days	NAV (online)	2021
Dutch Design Week	TU/e (Eindhoven, NL)	2021
EWUU Conference	TU/e (Eindhoven, NL)	2022
Zoom Forward - 29th annual European Congress on Obesity	EASO/IFSO (Maastricht, NL)	2022
Presentations		
MORe consortium “Dietary Assessment Tools”	Danone Nutricia Research (Utrecht, NL)	2018
Medworq “Dietary Assessment Tools”	Medworq (Zeist, NL)	2019
Personalized Nutrition and Health consortium “Dietary Assessment App”	TiFN (Wageningen, NL)	2019
Christelijke Hogeschool Ede “Introduction of WUR & DAEB”	CHE (Ede, NL)	2019
Course: Methodologie voor voedselconsumptieonderzoek “Innovatieve methoden om voedingsinname/gedrag te meten”	Wageningen Academy (Wageningen, NL)	2020
ICDAM “Validity of an innovative 2-hour recall smartphone app: first results”	ICDAM (online)	2021
O’BRAIN lab “Dietary Assessment in Diet-Related Research”	O’BRAIN lab (online)	2021
NutriLab “2-day workshop on dietary assessment”	WUR (online)	2021
Nutricia Science cockpit “Traqq: a flexible dietary behaviour assessment tool”	Danone Nutricia Research (online)	2021

Nutritional Science Days “Development and evaluation of an innovative 2-hour recall smartphone app for (near) real-time dietary assessment”	NAV (online)	2021
EASO/IFSO “Innovation in Food Intake Registration”	EASO/IFSO (Maastricht, NL)	2022

General courses		
Student supervision “skills and tools”	WGS (Wageningen, NL)	2018
Introduction to R	VLAG (Wageningen, NL)	2019
Scientific writing	WGS (online)	2020
Critical thinking and argumentation	WGS (online)	2020
Efficient writing strategies	WGS (online)	2020
Writing grant proposals	WGS (online)	2021

Assisting in teaching and supervision activities		
HNH-52306 Quantified Self		2018-21
HNH-32806 Exposure Assessment in Nutrition and Health Research		2018-22
Supervising MSc students		2018-22

Other activities		
Preparation of research proposal	WUR (Wageningen, NL)	2018
Biweekly chair group meetings	WUR (Wageningen, NL)	2018-22
Weekly DAEB klein meetings	WUR (Wageningen, NL)	2018-22
Bimonthly DAEB groot meetings	WUR (Wageningen, NL)	2018-22
Monthly LIV Leefstijl interventie meetings	WUR (Wageningen, NL)	2021-22
Organization of webinar “BigO, an innovative platform for supporting public health via local evidence”	BigO consortium (online)	2020
Organization of “BigO closing webinar”	BigO consortium (online)	2021
Reviewing papers	WUR (Wageningen, NL)	2018-22

Colophon

The research described in this thesis was financially supported 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.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover	Stefanie van den Herik
Lay-out	Desiree A. Lucassen
Printed by	Digiforce | Proefschriftmaken.nl

Copyright © Desiree A. Lucassen, 2022