
L-Mobi NL
Moving to
Self-administered
Dietary Assessment
in
National Food
Consumption
Surveys
Liangzi Zhang
> flip to unlock

Propositions

1. Intentional behavioural change in dietary intake is the hardest challenge to solve in real-time dietary assessment. (this thesis)

2. The direct use of representative standard recipes is adequate in estimating population dietary intakes. (this thesis)

3. Evidence from peer-reviewed scientific publications comes too late for containing fast-spreading novel infectious diseases.

4. The use of machine learning in cross-pollinating multidisciplinary data promotes more multidisciplinary research collaborations.

5. The better a public health system works, the less it is noticed.

6. Reasoning is an ineffective strategy to settle a family dispute.

Propositions belonging to the thesis, entitled

Moving to Self-administered Dietary Assessment in National Food Consumption Surveys

Liangzi Zhang
Wageningen, 12 August 2020

Moving to Self-administered Dietary Assessment in National Food Consumption Surveys

Liangzi Zhang

Thesis committee

Promotors
Prof. Dr Hendriek Boshuizen
Special Professor of Biostatistic modelling
Wageningen University & Research

Co-promotor
Dr Marga Ocké
Senior scientist, National institute for Public Health and the Environment (RIVM), Bilthoven
Senior Researcher, Division of Human Nutrition and Health,
Wageningen University & Research

Other members
Prof. Dr P. van 't Veer, Wageningen University & Research
Dr A. Stafleu, Netherlands Nutrition Centre, The Hague
Prof. Dr J.E. Cade, University of Leeds, UK
Dr S.J.P.M. Eussen, Maastricht University

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technologies, Agrobiotechnology, Nutrition and Health Sciences)

Moving to Self-administered Dietary Assessment in National Food Consumption Surveys

Liangzi Zhang

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 Monday 16 November 2020
at 11 a.m. in the Aula.

Liangzi Zhang
Moving to Self-administered Dietary Assessment in National Food Consumption Surveys, 208 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2020)
With references, with summary in English

ISBN: 978-94-6395-527-0
DOI: https://doi.org/10.18174/530405

Contents

Summary		7
Chapter 1	General introduction	11
Chapter 2	Importance of details in food descriptions in estimating population nutrient intake distributions	27
Chapter 3	How does a simplified recipe collection procedure in dietary assessment tools affect the food group and nutrient intake distributions of the population	91
Chapter 4	Evaluation of the recipe function in popular dietary smartphone applications, with emphasize on features relevant for nutrition assessment in large-scale studies	119
Chapter 5	A systematic review and meta-analysis of validation studies performed on food record apps	157
Chapter 6	General discussion	181
	Acknowledgements	199
	About the author	203
	List of publications	205
	Overview of training activities	206

Summary

The ever-growing findings from dietary studies have confirmed the important role of dietary intake in the development of certain non-communicable diseases (NCDs). In order to take a better control of the NCD progression in a population, primary prevention measures such as dietary guidelines and food policies are updated constantly according to the most recent scientific evidence and observed dietary patterns of a specific population. To obtain these dietary patterns, dietary intake at a national level is commonly monitored by governments in the form of food consumption surveys. However, assessing the dietary intake of a large population has been a challenging task throughout the years. With highly diversified food consumption practices and varied individual capabilities of reporting dietary intake, vast inputs (financially, physically) are required for collecting detailed dietary information. Hence, this thesis focuses on improving the methodology of dietary data collections, specifically for the Dutch National Food Consumption Surveys (DNFCS). The investigations proceeded in two parallel paths. Firstly, on how to remove burdensome procedures from current survey collection methods. Secondly, by learning from other studies and tools where dietary assessment techniques were investigated.

The current dietary assessment method in the DNFCS is the interviewer-administered 24-hour recalls (24HRs) guided by the computer software called GloboDiet. Each food item goes through a round of detailed questions extracting relevant information (also called facets). This detail acquiring step has been the most time-consuming part of the interview. Besides, undesirable answers are likely to be obtained due to limited knowledge of the respondents. In order to enhance the interview efficiency while minimising the impact on the survey results, the importance of facets in terms of predicting the nutrient outcome was determined using a prediction model called random forest. As a result, 35% of the total facets were deemed unimportant and could be omitted; this would resulted in a change of 3.7% of the foods linked to the NEVO (Dutch Food Composition Database). The majority (79.4%) of the differences between percentile estimates of the population nutrient intake distributions before and after facet deletion ranged from 0% to 1%. The reduction of facets was estimated to save 637 hours for data collection and 442 hours for the data handling for a survey conducted on 3819 participants. However, facets that are informative for other food-related issues (e.g. food safety) should be carefully examined before deletion. (**Chapter 2**)

Another complicated task in the current GloboDiet 24HR interview is the recipe pathway. Typically, mixed dishes are firstly identified with a standard recipe, then the ingredient composition and amounts are adjusted according to the available information on the real dish eaten. A replacement of the burdensome recipe modifications with the unchanged standard recipes has been simulated in this study. Comparing the simulated results and the original dataset, the average of the absolute percentage difference for the population mean intakes was 1.6% across all food groups and 0.6% for nutrients. The soup group (-6.6%) and

docosahexaenoic acid (DHA) (-2.3%) showed the largest percentage difference. The resulting small difference was mainly owning to the small proportion of energy intake consumed through mixed meals (10%) among the Dutch population according to the survey. A list with more realistic standard recipes would enable the use of a simple recipe function in a self-administered 24hR or food diary. (**Chapter 3**)

With a fast-evolving smartphone industry and an increased awareness of diet-health relationships among the general public, large varieties of dietary recording apps have been developed and were made available for download from app stores. Since most of the apps were designed as self-administered instruments, their functionality might be a useful example for developing self-administered tools for large-scale nutritional monitoring or research. Out of 57 popular food record apps, 12 apps having a recipe function were scored according to pre-defined criteria. None of the apps provided adjustable standard recipes and applied retention factors to nutrients for heat-processed raw items. Energy and nutrient content from three random recipes were compared across apps and with NEVO. The variation in food composition databases (FCDs) underlying each app contributed the most to the differences larger than 5% of Daily Reference Intake (DRI) in 49% of the micronutrients and 20% of the energy and macronutrients. Applying retention factors decreased the nutrient content for specific heat-sensitive vitamins such as B6, B12, and folate up to 45%. Overall, the components of current commercial apps vary, which might affect the accuracy of nutrient outcomes. In general, they focused more on the ease of use than getting accurate information. (**Chapter 4**)

Different from commercial apps, that have been mostly compared with each other, research-based apps have been described more in detail, reporting on their development, validity compared with a reference method, and usability or feasibility of applications in a sample population. A systematic review and meta-analysis on validation studies therefore provides insights into the general applicability and potentially the common flaws of apps. From an online search of literatures from 2013 to 2019, 14 studies were found that have validated food record apps in real-life settings. The pooled mean difference between the apps and the reference methods across studies showed a general underestimation of energy intake (-202 kcal/day) by using the apps. Studies with different FCDs for each method had the largest mean differences. The sources of variation were traced for studies that compared food group/ food item differences. A variation in study designs has been found among studies, which impedes the comparisons across studies, (e.g. use of energy-adjusted/log-transformed values). In general, most studies did not comply with the recommended procedures for conducting validation studies. (**Chapter 5**)

In **Chapter 6**, we discuss the implications and methodological reflection combining the findings from previous chapters. Other aspects related to dietary assessment, such as technology evolvement, data privacy, future directions are also discussed. Lastly, based on the evidence from our studies and other literatures, we come up with a recommended procedure for developing new self-administered methods for NFCS in general.

In conclusion, this thesis has shown that a simplification of current interviewer-administered 24HR is promising, which implies that the simplified functions might work equally well in a more cost-effective self-administered method. The advanced features and prevalence of use have made smartphones the optimal platform for monitoring dietary intakes at a population level. Still, a larger underestimation of energy intake using self-administered methods is expected compared to interviewer-administered methods, which implies the need for more guidance compared to using commercial apps, and careful interpretation of results. The validity of apps should be tested among different age groups, and a compatible option for those having difficulties in completing the survey by themselves should be considered. Moving to a self-administered method is a big step for NFCS, which requires careful considerations and large inputs during the development and validation phase. However, the lower costs and efforts required by using self-administered method could highly likely to counterbalance the initial investment, in the meanwhile, providing participants with a more flexible platform for dietary recording.

Chapter 1

General introduction

National Food Consumption Surveys: General Use and Challenges

In recent decades, the increasing prevalence of non-communicable diseases worldwide (e.g. cardiovascular disease, stroke, type 2 diabetes, and some cancers) induces enormous economic and social burden. These diseases are currently contributing to around 75% of all death worldwide (1). Overweight and obesity are one of the major cardiometabolic risk factors closely associated with unhealthy diet (2). Therefore, there has been a shift in focus of health authorities from disease treatments to primary prevention by assessing and managing dietary patterns (3, 4). The dietary patterns of a specific population can be derived from food consumption surveys that capture the detailed consumption of foods, beverages, and supplements at individual levels (5). National-scale food consumption surveys became to be the main source of information on the prevalence of dietary risk factors at a population level and have been increasingly conducted across countries worldwide (6). They are important basis for policy-making, providing insights into the dietary practices of the population, and enable evaluation of compliance with dietary guidelines, and inform on the appropriateness of food policies (7-9). Equally important, nutrition surveys can provide information on the exposure to food-related hazards and emerging risks to inform updates on food safety legislation (10).

Nutrition surveys typically consist of, firstly, collecting food consumption from a representative sample of the population using a dietary assessment instrument, secondly, obtaining nutrient information by linking reported food items to food composition databases (FCDs) (11). The data collection step is especially challenging, due to the vast varieties of available foods and unbalanced participation rates from different population groups. Besides, errors made intentionally as well as unintentionally by subjects can easily occur when perceiving and reporting the kind and the amount of food they consumed (12). Such measurement error can be divided into random or systematic. Random-errors (e.g. day-to-day variation of intakes) reduce the precision of the measurement, resulting in a loss in statistical power. Loss of power, however, could be mitigated with large-enough sample sizes and repeated measurements. Systematic errors generate bias (e.g. underreporting), can be intake-related or person-specific, and can only be identified and corrected with a reference method that is preferably free of error (13-15). Hence, a successful collection of large-scale data should take both types of measurement error into account, require substantial investment in time and cost, and has been a challenge for government institutions, researchers, and dietitians (16).

National Food Consumption Surveys in Europe

National food consumption surveys are presently carried out in many European countries and provide valuable information on dietary patterns and food safety at both national and EU level. The most frequently used dietary assessment methods in Europe for collecting national food consumption data are 24HRs and food records (2). Both open methods can provide detailed information on the intake of all foods and drinks on a specific day(s). 24HRs depend on the subjects‘ ability to recall all foods and portion sizes consumed over a reference period of one day, and were traditionally conducted in person or by telephone interviews with a trained interviewer following a structured protocol that facilitate participants in recalling (17, 18). On the other hand, food records require participants to self-report food consumption in real-time. Although this prevents errors associated with memory loss, food records suffer from behaviour change and misreporting, due to reactivity bias and social desirability bias, respectively (19). To facilitate complete and detailed recording, careful in-person training of participants before data collection and data reviewing by researchers afterwards poses additional burdens for this method (20). Both short-term methods have limited ability in capturing episodically consumed foods (21). Hence, multiple days of measurement in combination with a food frequency questionnaire (FFQ) were suggested as inputs for statistical techniques developed for usual dietary intake estimations (11, 22-25). The European Food Safety Authority advises EU Member states to collect two non-consecutive 24hRs for adults and two non-consecutive food records for children (26).

The demand for a structured and standardised collection of dietary intake data in national nutrition surveys has led to a wide application of Computer Assisted Interview (CAI) software. A computer-assisted 24HR interview software Automated Multiple-Pass Method (AMPM), developed and validated by the U.S. Department of Agriculture (USDA), was used in the National Health and Nutrition Examination Survey in the US (7). Whereas a validated (27) and standardized software GloboDiet (formerly known as Epic-Soft), was used by some European countries for the aim of collecting harmonized data among the EU Member states (28-30).

GloboDiet Features

The current Dutch National Food Consumption Survey (DNFCS) follows the standard protocol of GloboDiet, with adjusted food lists, probe questions and facet-descriptor system specific to the Dutch dietary culture and available food products (31). The flow diagram of the 24HR procedure is shown in Figure 1. Firstly, a quick list of all consumed foods throughout the recalled day is generated. Then, the facet-descriptor system (e.g. preservation method, fortification, etc.) enables the interviewers to collect detailed information for each food item. Sufficiently detailed dietary intake data enable more accurate nutrient estimations. They are also required for adequate exposure assessment of food contaminants, fortification and environmental impact because exposure levels vary widely due to variation in food processing, preservation, cooking, etc. (32). Meanwhile, probe questions recover food items and eating occasions not reported initially, such as common additions to foods (e.g., butter on toast) and snacks. The effectiveness of these probes is well-established and is therefore part of the interviewing protocols for all standardised high-quality 24HRs (33). An early study found that respondents with interviewer probing reported 25% higher dietary intakes than did respondents without interviewer probing (34).

Foods typically eaten as mixed dishes consist of multiple ingredients with specific food preparation and often with cooking involved (35). For respondents, it might be difficult to accurately describe the types and amounts of the various ingredients in mixed dishes, especially for those who were not involved in cooking (36). Standard recipe databases are often used in national food consumption surveys to ease the recording of mixed dishes (37). The possibility to modify the standard recipes, if a participant can report the specific recipe, is part of a comprehensive recipe function in GloboDiet which involves ingredient identification and quantity calculation with the presence of an interviewer (Figure 1). The accurate calculation of nutrients for a cooked food takes weight change and nutrient loss due to cooking and processing into account (38). The GloboDiet program calculates the cooked amount of ingredients from raw amount using pre-defined algorithms and standard food-specific coefficients (e.g., raw-to-cooked yield factors, or edible part coefficients)(29, 30). The cooked amounts are then multiplied by the nutrient values of the cooked food items found in the Dutch food composition table (NEVO)(38).

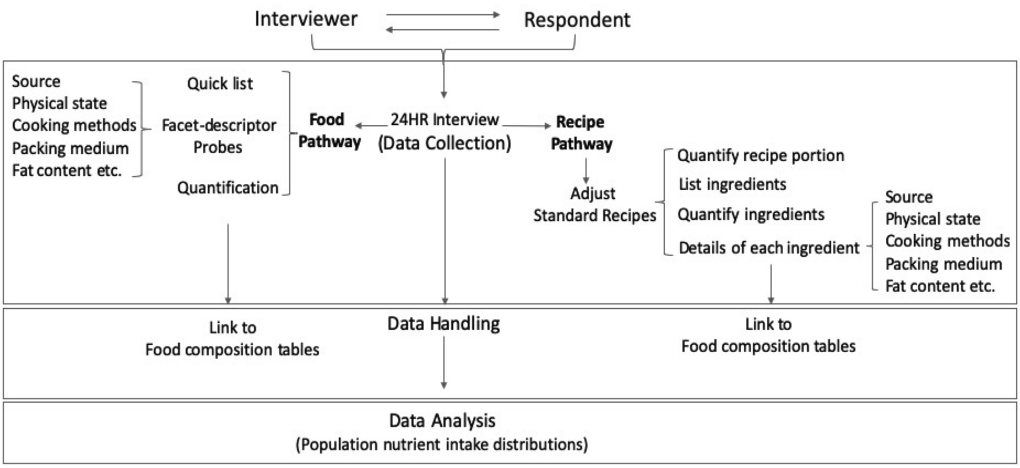


Figure 1. Main steps of the 24HR interview procedure using GloboDiet (based on Slimani et al. (39)), including data collection, handling and analysis.

Limitations of Interviewer-administered 24HR

An interviewer-administered 24-hour recall interview using the multiple-pass approach typically requires between 30 and 45 minutes (40). Interviewers must be highly trained and experienced to collect high-quality data that provide information that fits the study purposes. In addition, the procedures in the current method have limitations in time and location for data collection, all these features together induce high costs (12, 41). Besides, the requirement of knowing many details of the food/recipe consumed will lengthen the interview, which is at odds with increasing the participation rate of the survey, and might also cause intentional under-reporting. Moreover, apart from involuntary food omission, the lack of ability to conceptualise portion sizes based on memory is another main source of recall bias that could not be solved easily in current data collections (42). These limitations, together with the high cost and efforts for the survey implementation, have contributed to the initiatives to simplify the collection and handling of dietary data in national food consumption surveys (43, 44).

New Opportunities with Technologies

Continued development in ICTs and increased ubiquity of computer and portable devices led to investigations into self-administered tools to overcome the high cost and reliance on highly trained interviewers, reduce respondents' burden and increase the efficiency of data collection and handling (45). Internet-based dietary assessment software has been increasingly developed that require respondent themselves (rather than a trained nutritionist) to correctly identify and select the appropriate food or drink item that they have consumed (46). These self-administered tools usually provide tutorials, digital images for food identification and portion-size estimation, and various audio files, to facilitate the data collection procedures (12). In terms of linkage to FCDs, compared to manual linkage in some interviewer-assisted methods (e.g. GloboDiet), efforts were also reduced in self-administered methods due to the already-established linkage between food items and options in FCDs before the data collection (47).

More recently, the advent of mobile devices allowed more functionalities beyond text-based systems to be incorporated, such as barcode scanning and image capturing, which requires less time and effort on the part of respondents, and are less subjective compared to descriptions provided by respondents (16, 48, 49). Among the available mobile devices (e.g. PDA, tablets, etc.), smartphones are the most prevalent tool that reached a global penetration of 41.5% in 2020, while the Netherlands is forecasted to reach 96 percent as of 2024 (50). The accessibility and popularity of diet and healthy lifestyle applications (known as "apps") opened a new array of possibilities for innovative applications for dietary recording (51). A wide variety of food record apps became available to increase the awareness of the type of food consumed and facilitate body weight control or disease management with personalised advice provided (52). In the research domain, apps can enable the measuring of food and nutrient intakes in real-time from large populations at a relatively low cost, with automated calculation of daily food and nutrient intake and less interviewer involvement. Participants have greater flexibility and fewer time constraints to complete the survey because users usually carry smartphones with them (44, 46). The advantages of using smartphones for dietary assessment has prompted researchers investigating the opportunities for their applications in epidemiological research and nutrition monitoring (53-57).

Apart from an exclusive self-reported electronic food diary, photo-assisted food records with or without analysis by dieticians and automatic analysis of digital food images have also been actively investigated (56). The images provide objective information such as food type, volume, and leftovers, and may even record foods that were forgotten and not reported in the food registration, hence can be used as a supplement to traditional written or electronic food records (58, 59). Although image-assisted methods minimise participant and researcher burden to some extent during data collection, the amount of data influx is vast and requires additional work from researchers for data cleaning (60). On the other hand, advanced computer vision has enabled the development of automatic image recognition (61). However, computer vision methods still exhibit practical limitations, such as a shortage of food images that are representative of a specific diet for training the algorithm. Hence, a higher level of maturity is required before they can be used as the main dietary assessment method (62).

Alternative methods for detecting eating behaviour or food consumption are based on sensors. One type of sensor could detect noises of chewing or swallowing when placed on the ear or the neck (63). More recent miniaturised tooth-mounted radiofrequency sensors are capable of detecting nutrients and wirelessly communicating to a mobile device (64). There are also devices attached to the arm for detecting movement for eating behaviours using magnetic proximity and infrared sensors (65). However, these sensors are often intrusive, uncomfortable and/or cosmetically unpleasant for long-term wear. Less intrusive methods including using smart kitchen equipment (e.g., plates, spoons, and tables) to identify food items and weight before and after meal consumption (66). Another miniaturised hand-held (near-infrared) spectrometers could determine the characteristic of food matrix properties by scanning food items (67). However, the applications of these sensor-based devices were limited to controlled-settings and are still immature to be applied in larger samples of free-living individuals (67). The cost of these devices has not been established since they are still in the development phase and have not gone further to establish a market cost (9). Besides, their inability to recognize all foods and nutrients is the main impediment for current sensors being the main dietary assessment method. Some examples of available technologies in sensor-based technologies are listed in Figure 2.

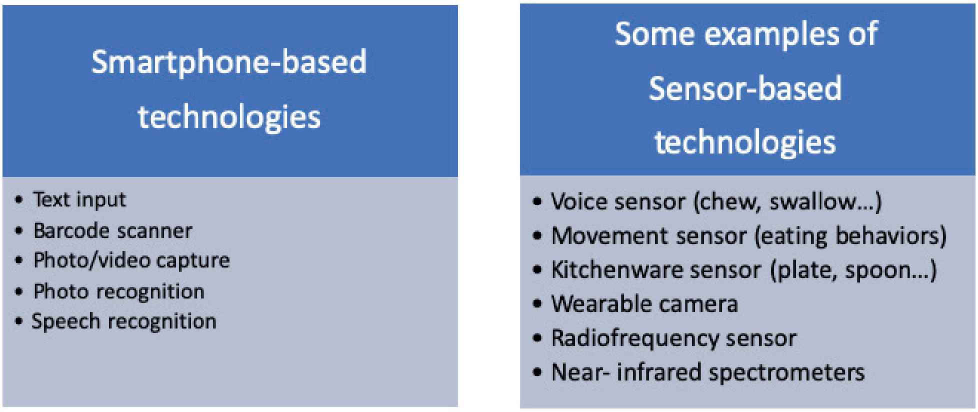


Figure 2. Examples of technologies used in the smartphone- and sensor-based dietary assessment methods.

Development and Evaluation of Smartphone Apps

With the increasing amount of smartphone dietary apps, there is a corresponding rise in the number of studies that evaluate apps in terms of their accuracy, usability and behavioural change impact (68). Owing to the large varieties of app design, more multidimensional quality assessments have been found in addition to conventional evaluations, such as on functionality, popularity, adherence to self-monitoring etc. In terms of content accuracy, the evaluations focusing on assessing the quality of the underlying food composition database and associated nutrient calculation algorithms have been published for most commercial apps (68, 69). However on some specific functionalities, like recipe functions, evaluations are missing. On the other hand, for apps developed for research purposes, the development process, feasibility, usability and validity are more commonly assessed (45, 54, 70). The level of validity refers to the degree to which the new method measures what it intends to measure quantitively (71). Investigating the validity of new methods is crucial, given the complexity of our diet and multiple sources of bias that impact the nutrient outcomes of dietary assessments. The result of using a test method, in this case the app, should be compared with a reference method that has a greater degree of demonstrated validity and has uncorrelated errors with the test method (17). A summary on the study design and nutrient comparisons of validation studies could provide useful information on the likelihood of applying a certain type of apps to a specific study purpose (e.g. in NFCS).

Aim and outline of this thesis

As discussed before, traditional dietary assessment methods are subjected to both random and systematic errors, mainly owing to self-reporting. Meanwhile, dietary assessment methods that are open in nature induce a heavy burden on both the researchers and respondents, especially for a detailed dietary data collection at a large scale. Developing a more cost-effective method taking advantages of current technologies is the foremost task for NFCS conducted in most countries, resulting in a trend of moving from an interviewer-administered to a self-administered dietary assessment method. Among currently available devices, smartphones show a great potential to be applied in DNFCS with their growing functionalities and data processing capacities. Therefore, to enhance the cost-efficiency of DNFCS in data collection and handling, a more efficient and flexible method built in smartphones has been proposed. This thesis includes investigations into the proposed component configurations and review of other evidence to deliver support for future app development.

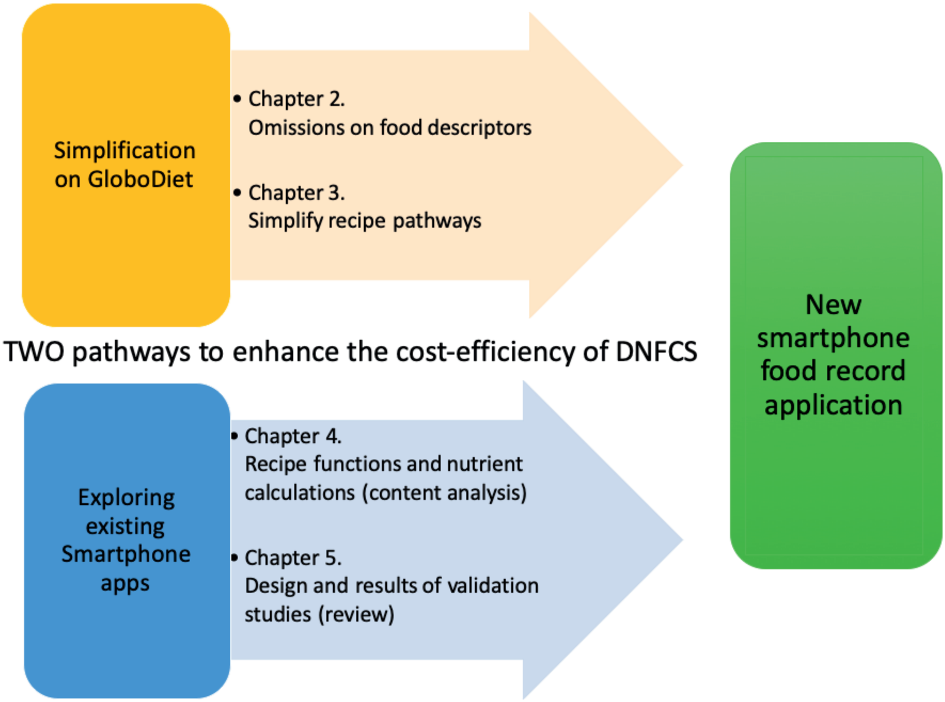


Figure 3. Two parallel paths for methodology development of switching from interviewer-assisted 24HR to self-administered smartphone food records.

Figure 3 illustrates two parallel paths assisting methodological transformations from interviewer-assisted 24HRs to self-administered smartphone food records in DNFCS. In the first path, the level of accuracy and efficiency resulting from the component simplification of existing methods were evaluated. Chapter 2 identified less important food descriptions (facets) in GloboDiet, and evaluated the influence of the facet reduction on nutrient intake distributions of the population and the extent of time-saving in a simulation study. In chapter 3, a simulation on removing complicated steps of recording consumed mixed meal composition from GloboDiet is presented, looking at the impact this has on the food group and nutrient intake distributions of the Dutch population.

In the second path, evidence on the available technologies from other studies are summarised and combined with the existing innovation propositions in DNFCS. Chapter 4 summarises the mixed meal recording features in several popular food diary apps from a research perspective and their accuracy in estimating nutrient intakes compared to the Dutch Food Composition Database (NEVO). Chapter 5 systematically reviews the existing validation studies of food record apps concerning their study designs, and pools the results of nutrient comparisons between the apps and the respective reference methods. In the final chapter of this thesis, Chapter 6, the main findings of the chapters are summarised and discussed. This chapter also includes recommendations for future research.

Reference

1. Arens-Volland AG, Spassova L, Bohn T. Promising approaches of computer-supported dietary assessment and management-Current research status and available applications. Int J Med Inform. 2015;84(12):997-1008.
2. Rippin H, Hutchinson J, Evans C, Jewell J, Breda J, Cade J. National nutrition surveys in Europe: a review on the current status in the 53 countries of the WHO European region. Food & Nutrition Research. 2018;62.
3. Mozaffarian D, Wilson PW, Kannel WB. Beyond established and novel risk factors: lifestyle risk factors for cardiovascular disease. Circulation. 2008;117(23):3031-8.
4. Organization WH. European food and nutrition action plan 2015–2020. 2015.
5. Kroes R, Muller D, Lambe J, Lowik MR, van Klaveren J, Kleiner J, et al. Assessment of intake from the diet. Food Chem Toxicol. 2002;40(2-3):327-85.
6. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2224-60.
7. Ahluwalia N, Dwyer J, Terry A, Moshfegh A, Johnson C. Update on NHANES Dietary Data: Focus on Collection, Release, Analytical Considerations, and Uses to Inform Public Policy. Advances in Nutrition. 2016;7(1):121-34.
8. Rollo ME, Williams RL, Burrows T, Kirkpatrick SI, Bucher T, Collins CE. What Are They Really Eating? A Review on New Approaches to Dietary Intake Assessment and Validation. Current Nutrition Reports. 2016;5(4):307-14.
9. Archundia Herrera MC, Chan CB. Narrative Review of New Methods for Assessing Food and Energy Intake. Nutrients. 2018;10(8).
10. Ambrus Á, Horváth Z, Farkas Z, Dorogházi E, Cseh J, Petrova S, et al. Pilot study in the view of a Pan-European dietary survey – adolescents, adults and elderly. EFSA Supporting Publications. 2013;10(11):508E-n/a.
11. Gibson RS. Principles of nutritional assessment: Oxford university press, USA; 2005.
12. Shim J-S, Oh K, Kim HC. Dietary Assessment Methods in Epidemiologic Studies. Epidemiology and Health. 2014;36:e2014009.
13. Thompson FE, Kirkpatrick SI, Subar AF, Reedy J, Schap TE, Wilson MM, et al. The National Cancer Institute's Dietary Assessment Primer: A Resource for Diet Research. J Acad Nutr Diet. 2015;115(12):1986-95.
14. Subar AF, Crafts J, Zimmerman TP, Wilson M, Mittl B, Islam NG, 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):55-64.
15. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. J Natl Cancer Inst. 2011;103(14):1086-92.

16. Tay W, Kaur B, Quek R, Lim J, Henry CJ. Current Developments in Digital Quantitative Volume Estimation for the Optimisation of Dietary Assessment. Nutrients. 2020;12(4):1167.
17. Willett W, Lenart E. Reproducibility and validity of food-frequency questionnaires. Nutritional epidemiology. 2013:96-141.
18. Zimmerman TP, Hull SG, McNutt S, Mittl B, Islam N, Guenther PM, et al. Challenges in converting an interviewer-administered food probe database to self-administration in the National Cancer Institute Automated Self-administered 24-Hour Recall (ASA24). Journal of food composition and analysis : an official publication of the United Nations University, International Network of Food Data Systems. 2009;22(Suppl 1):S48-S51.
19. Macdiarmid J, Blundell J. Assessing dietary intake: Who, what and why of under-reporting. Nutrition research reviews. 1998;11(2):231-53.
20. Kolar AS, Patterson RE, White E, Neuhouser ML, Frank LL, Standley J, et al. A practical method for collecting 3-day food records in a large cohort. Epidemiology. 2005;16(4):579-83.
21. Illner AK, Nothlings U, Wagner K, Ward H, Boeing H. The assessment of individual usual food intake in large-scale prospective studies. Ann Nutr Metab. 2010;56(2):99-105.
22. Conrad J, Nothlings U. Innovative approaches to estimate individual usual dietary intake in large-scale epidemiological studies. P Nutr Soc. 2017;76(3):213-9.
23. Satoh M, Tanno Y, Hosaka M, Metoki H, Obara T, Asayama K, et al. Salt intake and the validity of a salt intake assessment system based on a 24-h dietary recall method in pregnant Japanese women. Clinical and Experimental Hypertension. 2015;37(6):459-62.
24. Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. J Am Diet Assoc. 2006;106(10):1640-50.
25. Carriquiry AL. Understanding and Assessing Nutrition. Annu Rev Stat Appl. 2017;4:123-46.
26. Guidance on the EU Menu methodology1 European Food Safety Authority.
27. Crispim SP, Geelen A, Souverein OW, Hulshof PJ, Ruprich J, Dofkova M, et al. Biomarker-based evaluation of two 24-h recalls for comparing usual fish, fruit and vegetable intakes across European centers in the EFCOVAL Study. Eur J Clin Nutr. 2011;65 Suppl 1:S38-47.
28. Crispim SP, de Vries JH, Geelen A, Souverein OW, Hulshof PJ, Lafay L, et al. Two non-consecutive 24 h recalls using EPIC-Soft software are sufficiently valid for comparing protein and potassium intake between five European centres--results from the European Food Consumption Validation (EFCOVAL) study. Br J Nutr. 2011;105(3):447-58.

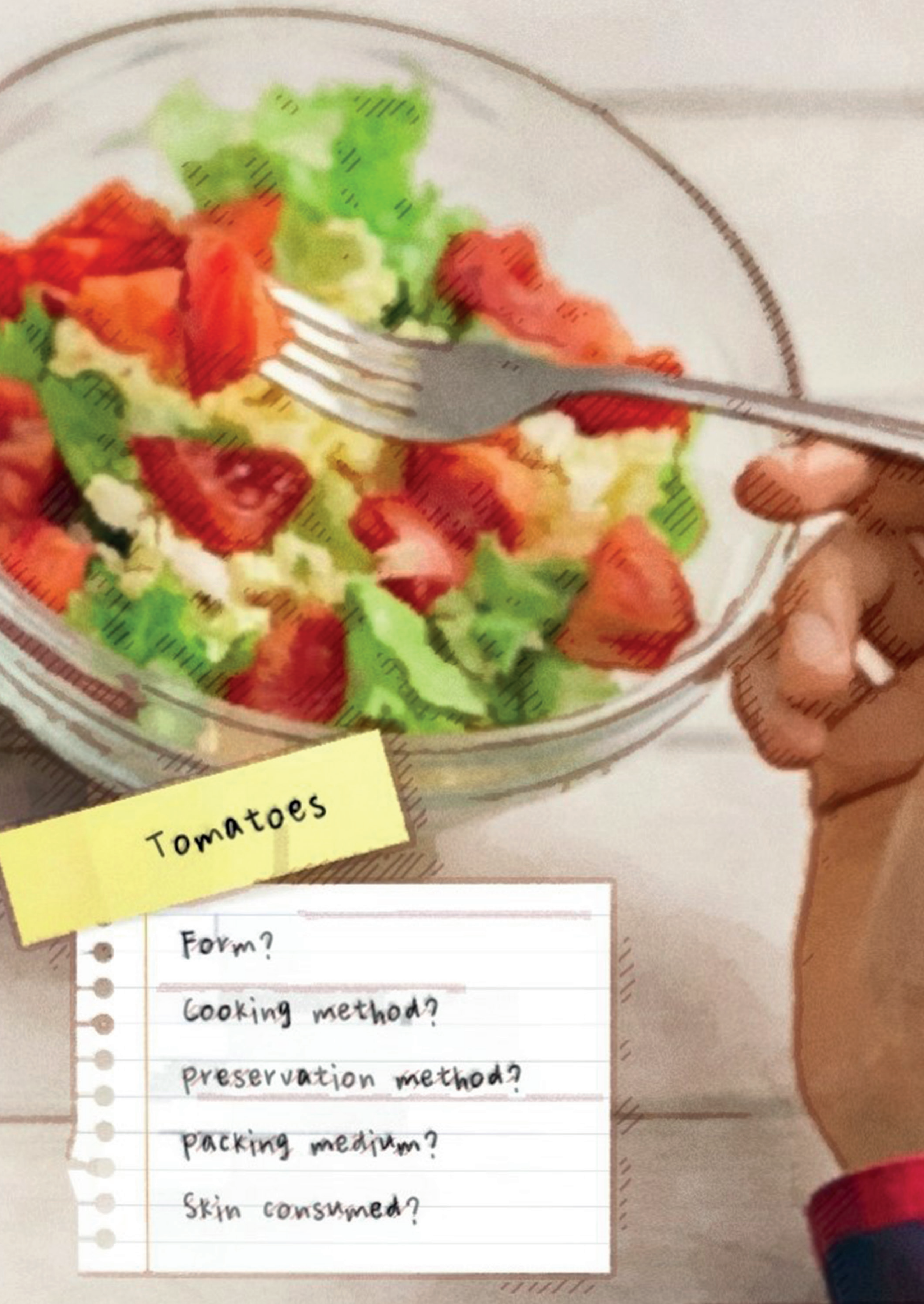
1

29. Slimani N, Deharveng G, Charrondiere RU, van Kappel AL, Ocke MC, Welch A, et al. Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. Comput Methods Programs Biomed. 1999;58(3):251-66.
30. EFSA. Guidance on the EU Menu methodology. EFSA Journal2014.
31. Slimani N, Casagrande C, Nicolas G, Freisling H, Huybrechts I, Ocke MC, et al. The standardized computerized 24-h dietary recall method EPIC-Soft adapted for pan-European dietary monitoring. Eur J Clin Nutr. 2011;65 Suppl 1:S5-15.
32. EFSA. General principles for the collection of national food consumption data in the view of a pan-European dietary survey. EFSA Journal. 2009;7(12):1435.
33. Thompson FE, Subar AF, Boushey CJ, Ferruzzi MG, Delahanty LM. Chapter 1 - Dietary Assessment Methodology A2 - Coulston, Ann M. Nutrition in the Prevention and Treatment of Disease (Fourth Edition): Academic Press; 2017. p. 5-48.
34. Campbell VA, Dodds ML. Collecting dietary information from groups of older people. J Am Diet Assoc. 1967;51(1):29-33.
35. Kong K, Zhang L, Huang L, Tao Y. Validity and practicability of smartphone-based photographic food records for estimating energy and nutrient intake. Asia Pac J Clin Nutr. 2017;26(3):396-401.
36. Schober DJ, Carpenter L, Currie V, Yaroch AL. Evaluation of the LiveWell@School Food Initiative Shows Increases in Scratch Cooking and Improvement in Nutritional Content. J School Health. 2016;86(8):604-11.
37. Imaeda N, Goto C, Fujiwara N, Kato R, Tokudome Y. Comparison of Nutrient Contents between the Standard Recipes Used in the National Health and Nutrition Survey and the Dishes Actually Consumed by Community-dwelling Elderly Individuals. Journal of Japanese Society of Nutrition & Food Science. 2016;69(5):237-48.
38. Reinivuo H, Bell S, Ovaskainen ML. Harmonisation of recipe calculation procedures in European food composition databases. J Food Compos Anal. 2009;22(5):410-3.
39. Slimani N, Ferrari P, Ocke M, Welch A, Boeing H, Liere M, et al. Standardization of the 24-hour diet recall calibration method used in the european prospective investigation into cancer and nutrition (EPIC): general concepts and preliminary results. Eur J Clin Nutr. 2000;54(12):900-17.
40. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc. 2010;110(1):48-51.
41. Thompson FE, Subar AF. Dietary assessment methodology. Nutrition in the Prevention and Treatment of Disease: Elsevier; 2017. p. 5-48.
42. Forman M, Zhang J, Gunter E, Yao S, Gross M, Qiao Y, et al. Season-specific Correlation between Dietary Intake of Fruits and Vegetables and Levels of Serum Biomarkers among Chinese Tin Miners at High Risk for Lung Cancer. Annals of the New York Academy of Sciences. 1999;889(1):230-9.

43. Ocke MC, Slimani N, Brants H, Buurma-Rethans E, Casagrande C, Nicolas G, et al. Potential and requirements for a standardized pan-European food consumption survey using the EPIC-Soft software. European Journal of Clinical Nutrition. 2011;65:S48-S57.
44. Touvier M, Kesse-Guyot E, Méjean C, Pollet C, Malon A, Castetbon K, 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. 2010;105(7):1055-64.
45. Subar AF, Thompson FE, Potischman N, Forsyth BH, Buday R, Richards D, et al. Formative research of a quick list for an automated self-administered 24-hour dietary recall. J Am Diet Assoc. 2007;107(6):1002-7.
46. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr. 2013;109(3):539-46.
47. Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Research. 2017;6:926.
48. Maringer M, Wisse-Voorwinden N, Veer PV, Geelen A. Food identification by barcode scanning in the Netherlands: a quality assessment of labelled food product databases underlying popular nutrition applications. Public Health Nutr. 2018:1-8.
49. Cade JE. Measuring Diet in the 21st Century: Use of New Technologies. P Nutr Soc. 2017;76(3):276-82.
50. Statista. Forecast of smartphone user numbers in the Netherlands 2018-2024 (in millions) 2020 [Available from: https://www.statista.com/statistics/494636/smartphone-users-in-netherlands/.
51. Long JD, Littlefield LA, Estep G, Martin H, Rogers TJ, Boswell C, et al. Evidence review of technology and dietary assessment. Worldviews on Evidence-Based Nursing. 2010;7(4):191-204.
52. Foster E, Lee C, Imamura F, Hollidge SE, Westgate KL, Venables MC, 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:e29.
53. Ambrosini GL, 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(1):5.
54. Eldridge AL, Piernas C, Illner AK, Gibney MJ, Gurinovic MA, de Vries JHM, 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).
55. Fallaize R, Zenun Franco R, Pasang J, Hwang F, Lovegrove JA. Popular Nutrition-Related Mobile Apps: An Agreement Assessment Against a UK Reference Method. JMIR Mhealth Uhealth. 2019;7(2):e9838.
56. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition. 2014;30(11-12):1257-66.

57. Allman-Farinelli M, Gemming L. Technology Interventions to Manage Food Intake: Where Are We Now? Current Diabetes Reports. 2017;17(11):103.
58. O'Loughlin G, Cullen SJ, McGoldrick A, O'Connor S, Blain R, O'Malley S, et al. Using a Wearable Camera to Increase the Accuracy of Dietary Analysis. American journal of preventive medicine. 2013;44(3):297-301.
59. Gemming L, Utter J, Mhurchu CN. Image-Assisted Dietary Assessment: A Systematic Review of the Evidence. Journal of the Academy of Nutrition and Dietetics. 2015;115(1):64-77.
60. McClung HL, Champagne CM, Allen HR, McGraw SM, Young AJ, Montain SJ, et al. Digital food photography technology improves efficiency and feasibility of dietary intake assessments in large populations eating ad libitum in collective dining facilities. Appetite. 2017;116:389-94.
61. Ming Z-Y, Chen J, Cao Y, Forde C, Ngo C-W, Chua TS, editors. Food Photo Recognition for Dietary Tracking: System and Experiment. International Conference on Multimedia Modeling; 2018: Springer.
62. Mezgec S, Korousic Seljak B. NutriNet: A Deep Learning Food and Drink Image Recognition System for Dietary Assessment. Nutrients. 2017;9(7).
63. Passler S, Wolff M, Fischer WJ. Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food. Physiol Meas. 2012;33(6):1073-93.
64. Tseng P, Napier B, Garbarini L, Kaplan DL, Omenetto FG. Functional, RF-Trilayer Sensors for Tooth-Mounted, Wireless Monitoring of the Oral Cavity and Food Consumption. Adv Mater. 2018;30(18):e1703257.
65. Jia W, Li Y, Qu R, Baranowski T, Burke LE, Zhang H, et al. Automatic food detection in egocentric images using artificial intelligence technology. Public health nutrition. 2019;22(7):1168-79.
66. Xu X, Wang J, Hou L, Guo Z, Li J. A Dietary Management System Using Radio-Frequency Identification Technology to Collect Information on Chinese Food Consumption: Development and Feasibility Study. JMIR Mhealth Uhealth. 2018;6(8):e166.
67. McClung HL, Ptomey LT, Shook RP, Aggarwal A, Gorczyca AM, Sazonov ES, et al. Dietary Intake and physical activity assessment: current tools, techniques, and technologies for use in adult populations. American journal of preventive medicine. 2018;55(4):e93-e104.
68. Grundy QH, Wang Z, Bero LA. Challenges in Assessing Mobile Health App Quality: A Systematic Review of Prevalent and Innovative Methods. American journal of preventive medicine. 2016;51(6):1051-9.
69. Griffiths C, Harnack L, Pereira MA. Assessment of the accuracy of nutrient calculations of five popular nutrition tracking applications. Public Health Nutr. 2018;21(8):1495-502.

70. Six BL, Schap TE, Zhu FM, Mariappan A, Bosch M, Delp EJ, et al. Evidence-based development of a mobile telephone food record. J Am Diet Assoc. 2010;110(1):74-9.
71. Noor Hafizah Y, Ang LC, Yap F, Nurul Najwa W, Cheah WL, Ruzita AT, et al. Validity and Reliability of a Food Frequency Questionnaire (FFQ) to Assess Dietary Intake of Preschool Children. Int J Environ Res Public Health. 2019;16(23).


Tomatoes
Form?
Cooking method?
preservation method?
packing medium?
Skin consumed?

Chapter 2

Importance of Details in Food Descriptions in Estimating Population Nutrient Intake Distributions

Liangzi Zhang, Anouk Geelen, Hendriek Boshuizen, José Ferreira, Marga Ocké

Nutrition Journal 2019, 18(1), 17

Abstract:
Background: Food consumption data with much detail in food descriptions enable their use for many purposes. However, the collection and handling of such data also require huge efforts. Our aim was to improve the efficiency of data collection and handling in 24-h dietary recalls (24HRs), by identifying less important characteristics of food descriptions (facets) and assessing the impact of ignoring them on energy and nutrient intake distributions.

Methods: In the Dutch National Food Consumption Survey 2007-2010, food consumption data was collected through 24HRs using GloboDiet software in 3819 persons. Questions on each food characteristic were asked according to the applicable facets. Food consumption data were subsequently linked to the food composition database. The importance of facets for predicting energy and each of the 33 nutrients was estimated by food group, using the random forest algorithm. Then a simulation study was performed to determine the influence of the deletion of the least important facets on population nutrient intake distributions.

Results: After 35% of facet descriptors were identified as unimportant, they were deleted from the total food consumption database. The majority (79.4%) of the percent difference between percentile estimates of the population nutrient intake distributions before and after facet deletion ranged from 0% to 1%, while 20% cases ranged from 1% to 5% and 0.6% cases more than 10%.

Conclusion: We conclude that our procedure was successful in identifying less important characteristics of food description for estimation of population nutrient intake distributions. This has the potential to reduce the time needed for conducting interviews and data handling.

Background

National food consumption surveys are important policy instruments and have been carried out successively in many countries [1, 2]. They serve many purposes, such as identification of nutrient inadequacies at the population level, risk assessment of hazardous substances, and development of dietary guidelines [1, 3].

In many national food consumption surveys, food consumption data are collected through 24-hour dietary recalls (24HRs) [4, 5]. This method allows the collection of abundant food consumption data, while it is less likely to alter diet behaviour and has fewer literacy requirements of the participants than food records [6, 7]. The 24HR methodology is an open-ended and retrospective method. Traditionally, interviewers collect information about the foods consumed during the preceding day or the previous 24 hours by triggering the participant's memory using different cues to increase the completeness of the survey [8]. This way of detail collection enables the survey to serve multiple purposes, but in the meantime increases the complexity and duration of the interview, data handling and linkage to databases [9, 10].

With the advent of computers, several comprehensive dietary assessment protocols have been incorporated into computer-assisted 24HR interview software and have been used in large-scale studies [5, 11, 12]. These protocols standardize the dietary data collection procedure and help the respondents recall their food intake to the maximum extent [13]. Examples include the Automated Multiple-Pass Method (AMPM), developed by the U.S. Department of Agriculture (USDA) to conduct the dietary interview for the National Health and Nutrition Examination Survey [14]. In Europe, the International Agency for Research on Cancer (IARC) has developed the menu-driven 24HR software GloboDiet (previously known as EPIC-Soft), which was validated to be used in food consumption surveys in European countries [15, 16].

In the multiple-pass protocol of GloboDiet, the most time-consuming step is the collection of detailed information on each consumed food (i.e. food description). Details of each food are collected through prompt windows of facets (various characteristics of a food) comparable to the probing questions in AMPM and descriptors (predefined answers on these questions). Examples of facets are fat content, cooking method, brand name, etc. Examples of descriptors are full fat, semi-skimmed, etc. [17]. Facets and descriptors standardize the procedure among different interviewers [18, 19]. In addition, the use of facets and descriptors enables characterization of the consumed foods in terms of their content of nutrients and potentially hazardous chemicals [20].

While applying a large number of facets and descriptors provides a high level of detail, it also increases the interview and data handling duration and thus the survey costs [7]. Furthermore, some food characteristics that require reading food labels (e.g. fortification) or knowledge about the preparation of the food (e.g. type of fat used) are difficult to answer for many of the participants [1, 21]. Also, linkage of consumed foods to the generic food composition database is more complex given the detailed information, because more available details increases the number of unique food-descriptor combinations that need to be linked to the food composition database [10, 22, 23]. To improve the cost-effectiveness of the survey while maintaining the quality of the data, there is a need to find a balance between the level of details in the data collection and the burden laid on participants and researchers.

The aim of the current study is to evaluate facet importance in predicting nutrient contents of foods, and the impact on population nutrient intake distributions of deleting less important facets from the data collection procedure.

Methods

Data collection

In the Netherlands, food consumption of the general Dutch population is monitored in Dutch National Food Consumption Surveys (DNFCS). The data used in this study came from the DNFCS performed from 2007 to 2010 on the diet of children and adults aged 7 to 69 years. Study design, recruitment, and results have been described elsewhere [24]. Subjects were excluded if they were pregnant, lactating, institutionalized or did not speak adequate Dutch. In total, 3819 participants (69%) were qualified and responded to the survey.

Dietary intake of participants was collected through two 24HRs on non-consecutive days with 2-6 weeks in between. The 24HRs for 2522 persons aged 16 and older were conducted by trained dieticians through telephone interviews. The 24HRs for 1297 children between 7 to 15 years old were collected by face-to-face interviews with the children and their care takers during home visits. All interviews were conducted following a same data collection and handling protocol.

During both face-to-face and telephone 24HR interview, dieticians used the multi-step computer-based interview software GloboDiet to guide the interview and to enter the data in the computer. The average time needed to complete one face-to-face 24HR interview was 41 minutes and 46 minutes for telephone interviews. The GloboDiet interview consists of the following five steps: 1. Collection of the general information, 2. Listing of foods and recipes consumed throughout the day, 3. Specification of details of foods by choosing descriptors of relevant facets and consumed amounts, 4. Quality check of inaccurate input, and 5. Dietary supplement intake [15]. The collection of details in step three took about 15 minutes. IARC provided for countries that used Globodiet as their data collection software with the common facets and descriptors. The actual selection of facets and descriptors could be adjusted according to country-specific situations. For the Dutch version of the software, a total of 16 facets with varying numbers of descriptors were selected by experienced dieticians based on knowledge of the food market and insight in the purposes for which the data were collected (Table 1).

Table 1. The list of facets and the examples of the corresponding descriptors in Globodiet for DNFCS 2007-2010.

	FACET NAMES	*NUMBER OF DESCRIPTORS*	*EXAMPLES OF DESCRIPTORS*
1	Source	21	beef, goat, pork…
2	Physical state/form as quantified	28	liquid, reconstituted from powder, minced …
3	Cooking method	28	cooked, baked, barbecued…
4	Preservation method	13	canned, frozen, dried…
5	Packing medium	22	canned in oil, canned in water…
6	Flavoured component	37	nuts, spices, mint…
7	Sugar content	6	non sweetened, sweetened, sugar reduced…
8	Fat content	39	whole, partially skimmed, skimmed…
9	Type of packing	4	in box, in paper, in bottle…
10	Food production	12	homemade fat used known, commercial fat used unknown…
11	Enriched/fortified	11	vitamins, mineral components, dietary fibre...
12	Brand name (yes/no)[a]	2	yes, no
13	Skin consumed	3	undefined, without skin, with skin
14	Visible fat consumed	3	undefined, without visible fat, with visible fat
15	Type of fat used	2	no fat used, choose from food list
16	Type of milk/liquid used	13	milk, whole milk, skimmed milk...

[a] A brand name would be entered if participants chose the descriptor 'yes', entered brand names were not put in the random forest analysis in this study.

2

Data handling

The total collected consumption data from all participants for the two 24HRs has 219,006 food records, with 350,369 descriptors ranging from 0 to a maximum of 8 for individual foods. This results in a number of 26,679 unique combinations of foods with descriptors. All food records were linked to 1599 most appropriate food codes in the Dutch National Food Composition Database (NEVO table 2011/3.0) by trained dieticians. NEVO 3.0 contains the energy, macro- and micronutrient contents of 2,389 food codes in total [25].

Statistical analysis

To assess the importance of the GloboDiet facets in predicting the nutrient contents of foods within a specific food group consumed in DNFCS, random forest prediction modelling was used [26]. Random forest is a prediction model that consists of a multitude of decision trees. Each tree is trained on different subsets of training data, and the remaining data (not used for the training) are used to estimate prediction error and variable importance. In our study, foods consumed by all participants in both 24HRs were used for predicting facet importance, the number of randomly selected variables to be considered when splitting the tree at each node was set to its default value (mtry = Total number of predictor variables/3); the number of trees for each nutrient was set at 10,000. Stratified by food group, the importance of a facet (denoted by %IncMSE), was calculated as the percentage increase in prediction error, when data for that facet were permuted in the dataset, while keeping data for the other facets unchanged. The random forest algorithm was applied through the randomForest package in Rstudio 1.1.383.

The 24HR variables of 16 facets, food IDs (a series of numbers identifying food items) and food subgroups (elements of main food groups) were regarded as predictor variables. The detailed food group information can be found in Addition file 1. The energy and 33 macro- and micronutrients were regarded as response variables and were predicted one by one with the prediction variables. Food IDs were treated as continuous variables, because it exceeds the limit of 32 levels allowed to categorical variables in the implementation of random forest. As comparable foods are numbered sequentially, treating food ID as continuous is reasonable. Facets were treated as categorical variables. Facet "Flavoured/added components" was separated into three sub-groups based on the category (nuts, sugary, savoury) of its descriptors, since the number of descriptors also exceeded the allowed 32 for categorical variables like in food IDs. The variable brand name was not included as predictor, as this consists of a free text field, yielding many unordered categories that were difficult to separate into sub-groups. Instead, we included the facet "Brand name (yes/no)" that indicated whether this brand name field was filled in or not.

In order to facilitate the comparison of the relative importance of facets between nutrients, within each food group and each nutrient, %IncMSEs were normalized by dividing them by the highest %IncMSE over the facets. The maximum normalized %IncMSE for the facet across all nutrients would be retained for each food group. After deleting facets with a max. normalized %IncMSE lower than 0.80 in each food group, trivial effects on population nutrient intake distributions were observed, therefore a cut-off point at 1.00 was chosen for greater effects. Hence, in each food group, facets with a normalized value below 1.00 for all nutrients were considered unimportant.

Simulation study

A simulation study was conducted to evaluate whether the distributions of population nutrient intake change significantly when the less important facets would not have been asked for during the 24HRs. The average nutrient intake calculated from two 24 HRs of each participant was used in the simulation study for estimating population nutrient intake distributions.

The simulation study consisted of two steps. Firstly, food-descriptor combinations with one or more facets which were considered unimportant were identified in the dataset with unique food-descriptor combinations. These food-descriptor combinations were relinked to the national food composition database NEVO considering only important facets. As illustrated in Figure 1, a NEVO code reassignment protocol was developed to identify NEVO codes of the most similar food-descriptor combinations with the combination of facets that needed to be relinked considering only important facets. For foods-descriptor combinations in the dataset with same food IDs, combinations received a positive score for each identical pair of descriptors (equal to the maximum normalized %IncMSEs) and a penalty for descriptors that were different (equal to the negative maximum normalized %IncMSEs). The scores were summed and the NEVO code of the food-descriptor combination with the highest score was assigned to the combination that needed to be relinked. In case there were more than one NEVO codes with the same highest score, or when no descriptors were left for a food, the NEVO code of a food-descriptor combination with a higher consumed quantity would be selected. In case the consumed quantities were also the same (occurred in 38 cases), the decision on NEVO code selection was made by a researcher.

Secondly, the energy and nutrient contents for 100 grams of foods in NEVO were multiplied with the quantities consumed in DNFCS 2007-2010, summarised by person by day and averaged over two days in both the dataset with original linkage to the NEVO database and the newly linked dataset. All results were weighted for small deviances in sociodemographic characteristics (age, sex, region, degree of urbanisation and educational level), day of the week and season of data collection, in order to give results that are representative for the Dutch population and representative for all days of the week and all seasons. The mean, median, 5th, 25th, 75th, 95th percentile and the percent differences of consumption per nutrient between the original and newly linked dataset were calculated for the total population and stratified by gender and age group (7-18 years old and 19-69 years old). The population nutrient intake distributions were conducted using the SAS 9.4 and the percent difference between the original and newly linked dataset were calculated using Excel 2016 software.

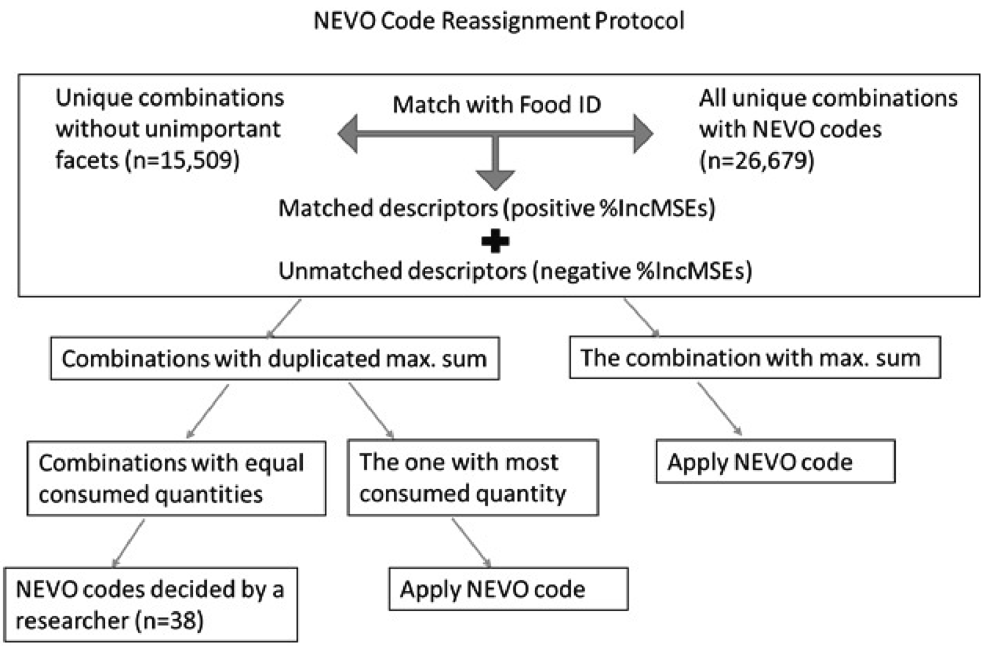


Figure 1. Flow chart of the NEVO code Reassignment Protocol. A NEVO code was assigned to each relinking combination according to the NEVO codes of the same food with the most similar descriptor combinations that have been linked by dieticians during the survey period. The combinations received a positive score for each identical pair of descriptors (equal to the maximum normalized %IncMSEs) and a penalty for descriptors that were different (equal to the negative maximum normalized %IncMSEs). The scores were summed, and the NEVO code of the food-descriptor combination with the highest score was assigned to the combination that needed to be relinked. In case there were more than one NEVO codes with the same highest score, or when no descriptors were left for a food item, the NEVO code of a food-descriptor combination with a higher consumed quantity would be selected. In case the consumed quantities were also the same (occurred in 38 cases), a researcher decided on NEVO code selection.

Results

Table 2 shows the normalized maximum importance (%IncMSEs) of each of 16 facets in predicting the nutrient contents of food items within each of 17 food groups using the cut-off point of 1.00, whereas results for the cut-off of 0.80 are shown in Additional File 2. Using the cut-off level of 0.80, a total of 50 out of 112 facets across food groups were considered unimportant. When the unimportant facets were deleted from the total food consumption database, 22% of the 350,369 facet descriptors were omitted. The majority of the percent difference between percentile estimates of the population nutrient intake distributions before and after facet deletion ranged from 0% to 1%, while only 2% cases ranged from 1% to 5%.

Across food groups, a total of 64 out of 112 facets fell below the cut-off point at 1.00, and have been deleted from the corresponding food groups in the simulation study. In the food groups 'Fats and oils' and 'Alcoholic beverages', no facets were unimportant, whereas all facets were unimportant for 'Cakes and sweet biscuits'. The food group 'Miscellaneous' has the largest amount of unimportant facets than the rest of the food groups. In the 'Meat' group, most facets had zero effect in predicting food groups, including 'Source', 'Packing medium', 'Fat content', 'Brand name (yes/no)', 'skin consumed, and 'visible fat consumed'.

From the perspective of the facets, 'Brand name (yes/no)' and 'Packing medium' were unimportant for the most of the food groups (10 and 7 food groups, respectively). The number of deletions ranged from 1 to 5 times for the rest of the facets. 'Source' and 'Visible fat consumed' were unimportant for all the food groups for which they are relevant (3 and 1 food groups, respectively). On the other hand, 'Physical state' and 'Cooking method' were strong predictors (importance of 1.00) for the largest number of food groups. Facet 'Type of packing' was only available for food group 'Fats and oils' and was a strong predictor for that food group. Despite that 'Brand name (yes/no)' was unimportant for most of the food groups, it was a strong predictor for food group 'Cereals', 'Fats and oils', 'Alcoholic' and 'Non-alcoholic beverages'. Full results of the facet importance for each nutrient in each food group can be found in Additional File 3.

In the original total food consumption database, 35% (121,015 out of 350,369) of the total number of descriptors used were identified as unimportant, which has resulted a NEVO code change of 11% (2,923 out of 26,679) of the combinations in the unique food dataset and 3.7% (8,196 out of 219,006) of the combinations in the total food consumption dataset.

After the NEVO codes had been reassigned, the population means and percentiles of two days' average energy and nutrient intakes in DNFCS 2007-2010 were calculated, as well as the percent difference between them. Table 3 shows the results of energy and ten nutrients that were mostly found in nutrition facts label. The results of all nutrients can be found in Additional File 4. The majority (79.4%) of the percent difference between distribution percentiles before and after facet deletion ranged from 0% to 1%, while 20% cases ranged from 1% to 5% and 0.6% cases more than 10%. Percent difference larger than 1% were mainly found in vitamins. Differences more than 10% appeared mostly in vitamins for 7-18 year olds and in the extreme percentiles P5 and P95. Some of the differences that were larger than 10% were very small as absolute difference. For example, the largest differences of 14.1% was for the P95 of vitamin B6; but the absolute difference of the two scenarios was 0.5 mg (rounded to mg). No general patterns were found on nutrient over- and underestimation after facet deletion for most nutrients. However, less vitamin C was found in each percentile after facet deletion for all age groups, whereas higher amounts of vitamin B group were found after facet deletion.

Table 2. The maximum normalized %IncMSEs of the existing facets in each food group.

		Food groups																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
	Facet names	Potatoes	Vegetables	Legumes	Fruits, nuts, olives	Dairy (products)	Cereal (products)	Meat (products)	Fish and shellfish	Egg (products)	Fats and oils	Sugar and confectionery	Cakes and sweet biscuits	Non-alcoholic beverages	Alcoholic beverages	Condiments and sauces	Soups	Miscellaneous	# of Omitted/# of original
1	Source					0.53*		0.00*		0.12*									3/3
2	Physical state/form quantified	1.00	1.00	0.34*	0.73*	1.00			0.80*					1.00		1.00	1.00	1.00	3/10
3	Cooking method	0.77*	1.00	0.31*	1.00	0.26*	1.00	1.00	1.00	1.00								1.00	3/10
4	Preservation method	1.00	1.00	0.61*	1.00	0.35*	0.51*	0.70*	1.00									0.52*	5/9
5	Packing medium	0.02*	0.87*	1.00	0.61*	0.00*	0.40*	0.00*	1.00									0.00*	7/9
6 A	Flavoured component A											1.00		0.50*				0.74*	2/3
6 B	Flavoured component B					1.00	0.96*							0.50*				0.63*	3/4
6 C	Flavoured component C											1.00	0.83*	0.67*				0.65*	3/4
7	Sugar content				0.57*	1.00	1.00					1.00		1.00				0.60*	2/6
8	Fat content				1.00	1.00	1.00	0.00*			1.00							0.60*	2/6
9	Type of packing										1.00								0/1
10	Food production					0.68*	0.68*						0.89*	1.00		1.00	0.92*	0.41*	5/7
11	Enriched/fortified					0.89*	1.00					1.00	0.50*	0.85*				1.00	3/6

12	Brand name(yes/no)	0.69*			0.73*	0.85*	1.00	0.00*	0.00*		1.00	0.96*	0.84*	1.00	1.00	0.87*	0.82*	0.74*	10/14
13	Skin consumed	1.00	1.00		0.82*			0.00*	0.40*										3/5
14	Visible fat consumed							0.00*											1/1
15	Type of fat used	1.00		0.31*			0.06*						0.45*			0.90*	0.38*		5/6
16	Type of milk/liquid used	1.00		0.30*		0.63*	1.00						0.60*			1.00	1.00	0.70*	4/8
	# of omitted/ # of original	3/8	1/5	5/6	5/8	8/12	5/11	7/8	3/6	1/2	0/3	1/5	6/6	4/8	0/1	2/5	3/5	10/13	

*Facets to be omitted for the corresponding food group with the maximum normalized %IncMSEs (among all nutrients) below the cut-off point of 1.00.

[a]Random forests require categorical predictors to have no more than 32 levels. Facet 6 (flavoured component) has more than 32 levels and was categorized into three sub-groups (6A, 6B, 6C) based on the category of flavours (nuts, sugary, savoury).

Table 3. The population means and percentiles of two days' average energy and ten nutrients' intake distributions before and after facets' deletion at cut-off at 1.00.

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Before	After	% D[a]	Before	After	% D	Before	After	% D	Before	After	% D	Before	After	% D	Before	After	% D
Energy (kcal)																		
P5	1457	1456	0.0	1314	1318	-0.3	1482	1481	0.0	1121	1134	-1.1	1481	1475	0.4	1163	1165	-0.2
P25	1859	1857	0.1	1671	1669	0.1	2048	2042	0.3	1539	1539	0.0	2010	2009	0.1	1565	1567	-0.1
P50	2226	2226	0.0	1915	1920	-0.2	2483	2487	-0.2	1837	1834	0.2	2421	2427	-0.2	1853	1851	0.1
P75	2653	2657	-0.1	2200	2199	0.0	2964	2965	0.0	2223	2225	-0.1	2907	2906	0.0	2220	2225	-0.2
P95	3676	3654	0.6	2694	2694	0.0	3749	3764	-0.4	2834	2830	0.2	3742	3745	-0.1	2813	2806	0.2
Mean	2326	2328	-0.1	1951	1953	-0.1	2534	2539	-0.2	1902	1905	-0.1	2496	2500	-0.2	1911	1914	-0.1
Protein (g)																		
P5	41	41	-0.4	37	37	0.0	55	55	0.0	42	42	0.0	49	49	0.0	41	41	-0.2
P25	56	56	-0.3	50	50	-0.1	74	74	-0.1	57	57	-0.1	70	70	-0.4	55	55	-0.3
P50	69	70	-0.8	59	59	0.2	87	88	-0.2	68	68	0.0	84	84	0.3	66	66	-0.1
P75	86	86	0.2	70	70	-0.2	105	105	0.4	82	82	0.1	102	102	-0.1	80	80	0.0
P95	114	114	0.0	91	91	0.4	133	132	0.5	102	102	0.0	131	130	0.3	101	101	0.2
Mean	72	72	0.0	61	61	0.0	90	90	0.0	70	70	0.1	87	87	0.0	69	68	0.1
Fat (g)																		
P5	45	45	0.1	40	41	-0.9	47	47	-0.1	34	34	0.2	47	47	-0.4	35	35	-0.8
P25	63	63	0.1	56	56	0.1	73	73	-0.3	52	53	-0.4	71	71	0.4	53	53	-0.4
P50	83	83	-0.9	70	70	-0.8	95	95	-0.3	68	69	-0.1	92	93	-0.3	69	69	-0.1
P75	102	103	-0.9	86	87	-1.2	118	119	-1.0	90	90	-0.5	116	117	-1.3	89	89	-0.4
P95	151	151	0.0	112	112	0.0	158	158	-0.3	122	123	-0.7	158	158	-0.1	120	121	-0.9
Mean	87	87	-0.3	72	73	-0.4	98	98	-0.4	73	73	-0.4	96	96	-0.4	73	73	-0.4
Saturated fatty acids (g)																		
P5	16	16	-0.9	14	14	0.2	16	16	-0.1	12	12	-0.1	16	16	-0.1	12	13	-4.1
P25	23	23	0.6	21	21	0.5	27	27	-0.4	20	20	-1.7	26	26	-0.3	20	20	-1.8
P50	30	30	-0.7	27	27	-1.3	34	35	-0.3	26	27	-1.4	33	34	-0.5	26	27	-1.1
P75	39	39	-0.4	33	33	-0.3	44	44	-0.5	35	35	-0.1	43	43	-0.5	34	34	-0.5
P95	54	54	-0.7	44	44	-0.2	60	61	-0.6	49	49	0.6	58	59	-0.8	48	48	-0.1
Mean	32	32	-0.3	27	28	-0.4	36	36	-0.7	28	28	-0.7	35	35	-0.6	28	28	-0.6
Carbohydrates (g)																		
P5	187	188	-0.1	160	161	-0.3	147	147	0.0	117	117	0.4	151	150	0.1	121	121	-0.1
P25	238	237	0.4	213	212	0.5	214	213	0.4	171	170	0.1	218	219	-0.1	176	177	-0.5
P50	287	288	0.0	246	246	0.0	266	265	0.3	208	209	-0.3	270	270	-0.2	216	216	0.0
P75	344	344	0.0	286	286	0.2	330	330	0.0	251	251	0.1	334	334	-0.1	260	260	0.1
P95	454	454	0.0	359	361	-0.4	436	436	0.1	338	337	0.2	439	438	0.2	341	340	0.2

2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Before	After	% D[a]	Before	After	% D	Before	After	% D	Before	After	% D	Before	After	% D	Before	After	% D
Mean	299	299	0.0	252	251	0.1	277	277	-0.1	215	215	0.0	281	281	-0.1	222	222	0.0
Monosaccharides (g)																		
P5	79	79	0.9	65	65	-0.2	42	43	-0.4	38	38	-1.1	46	45	0.4	41	41	0.3
P25	120	120	0.1	103	102	0.8	82	81	0.8	73	72	0.4	86	86	0.0	76	76	0.3
P50	148	148	0.3	131	130	0.2	114	114	-0.1	96	96	0.0	122	122	0.3	103	102	0.1
P75	185	185	0.3	161	160	1.0	159	159	0.2	129	129	0.2	166	165	0.3	137	137	0.0
P95	263	263	0.0	218	218	-0.1	243	243	0.2	193	194	-1.0	246	246	0.0	200	201	-0.6
Mean	156	156	0.2	134	133	0.4	125	125	0.1	104	104	-0.1	131	130	0.1	109	109	0.0
Fibre (g)																		
P5	9.8	9.79	0.0	9.0	9.1	-0.3	11.6	11.6	0.1	9.6	9.6	0.0	10.9	10.9	0.0	9.5	9.5	0.0
P25	13.9	14.2	-2.2	12.6	12.7	-0.8	16.9	16.9	-0.1	14.2	14.1	0.5	16.3	16.3	-0.3	13.9	13.9	-0.2
P50	17.5	17.8	-1.9	15.5	15.4	0.6	21.5	21.6	-0.2	17.6	17.5	0.1	21.0	21.0	-0.1	17.2	17.2	0.0
P75	21.7	21.8	-0.4	18.6	18.8	-0.7	26.4	26.6	-0.8	22.0	22.0	0.1	25.8	26.0	-0.7	21.3	21.3	-0.1
P95	29.9	29.9	0.0	24.6	25.0	-1.5	35.6	35.3	0.8	28.6	28.6	-0.1	35.3	35.1	0.7	28.4	28.3	0.3
Mean	18.5	18.5	-0.4	16.0	16.1	-0.5	22.3	22.3	0.0	18.4	18.4	0.2	21.6	21.6	0.0	18.0	18.0	0.0
Sodium (mg)																		
P5	1447	1444	0.2	1273	1270	0.2	1625	1622	0.2	1226	1219	0.6	1533	1559	-1.7	1234	1228	0.5
P25	1951	1953	-0.1	1748	1754	-0.3	2357	2364	-0.3	1773	1789	-0.9	2270	2272	-0.1	1767	1785	-1.0
P50	2473	2492	-0.8	2139	2134	0.2	2970	2986	-0.5	2236	2239	-0.1	2873	2890	-0.6	2225	2224	0.1
P75	3147	3152	-0.2	2567	2568	0.0	3630	3644	-0.4	2812	2825	-0.5	3540	3561	-0.6	2782	2782	0.0
P95	4256	4260	-0.1	3398	3409	-0.3	4792	4830	-0.8	3772	3780	-0.2	4739	4745	-0.1	3668	3674	-0.2
Mean	2611	2619	-0.3	2195	2198	-0.1	3058	3065	-0.2	2336	2346	-0.4	2977	2984	-0.2	2311	2320	-0.4
Calcium (mg)																		
P5	391	391	-0.1	332	332	0.0	470	469	0.2	429	429	0.0	448	449	-0.2	404	404	0.0
P25	686	690	-0.5	616	627	-1.9	791	792	-0.1	708	710	-0.3	769	769	0.0	695	696	-0.2
P50	922	917	0.5	828	834	-0.7	1092	1091	0.1	944	941	0.3	1049	1055	-0.6	924	924	0.0
P75	1232	1238	-0.5	1072	1080	-0.7	1448	1441	0.4	1195	1199	-0.3	1403	1401	0.1	1176	1179	-0.2
P95	1832	1865	-1.8	1545	1545	0.0	1999	1999	0.0	1702	1702	0.0	1998	1998	0.0	1677	1681	-0.2
Mean	995	1000	-0.4	874	880	-0.6	1148	1150	-0.1	992	992	0.0	1121	1122	-0.1	971	973	-0.1
Vitamin C (mg)																		
P5	24	24	1.3	23	21	8.3	26	27	-1.2	24	24	0.3	26	26	0.0	24	23	2.0
P25	48	46	3.6	47	44	5.2	50	50	-0.1	51	51	0.0	50	49	1.2	50	48	2.8
P50	76	72	5.4	73	67	8.4	84	83	1.1	83	81	2.1	82	80	1.7	81	79	2.2
P75	115	109	4.9	110	106	4.1	129	127	2.0	126	124	1.6	127	123	3.3	122	119	2.3
P95	176	167	5.2	169	165	2.6	218	218	0.0	209	206	1.3	210	208	1.0	204	201	1.9
Mean	86	82	4.8	84	79	5.3	98	96	1.5	95	94	1.6	96	94	2.1	93	91	2.2
Vitamin B6 (mg)																		
P5	0.7	0.7	-4.6	0.7	0.7	-3.0	0.9	0.9	-0.2	0.7	0.7	0.2	0.9	0.9	-1.5	0.7	0.7	-2.1

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Before	After	% D[a]	Before	After	% D	Before	After	% D	Before	After	% D	Before	After	% D	Before	After	% D
P25	1.1	1.2	-3.7	1.0	1.0	-4.1	1.5	1.5	-1.9	1.1	1.1	-1.0	1.4	1.4	-2.2	1.1	1.1	-2.0
P50	1.6	1.6	-4.2	1.3	1.4	-6.2	2.0	2.0	-1.5	1.5	1.6	-1.2	1.9	1.9	-2.6	1.5	1.5	-2.8
P75	2.2	2.4	-10.4	1.8	1.9	-5.1	2.5	2.6	-3.0	2.1	2.1	-1.0	2.5	2.6	-4.4	2.0	2.1	-2.2
P95	3.8	4.3	-14.1	3.0	3.4	-13.5	4.1	4.3	-2.6	3.4	3.5	-3.1	4.1	4.3	-4.1	3.3	3.5	-6.8
Mean	1.8	1.9	-7.5	1.5	1.6	-7.6	2.2	2.2	-2.5	1.7	1.8	-1.9	2.1	2.2	-3.3	1.7	1.7	-2.8

[a] % D represents the percent difference of nutrient intake distributions before and after facets' deletion for the Dutch population aged 7 to 69 years.
Percent difference larger than 5% are shown in bold.

2

Discussion

From the perspective of enhancing the efficiency of data collection and handling of GloboDiet 24HRs, we explored the option of deleting less important characteristics (facets) of food descriptions from the interview. At the food group level, the importance of each facet in predicting nutrient contents in foods was determined by the random forest algorithm. When the 35% least predictive facets were deleted from the dataset of the Dutch national food consumption survey 2007-2010, the difference between recalculated and originally calculated population nutrient intake distributions was small for the majority of the nutrients.

There are several possible explanations for certain facets to be less or more predictive in certain food groups. One reason for less predictive facets is that some facets were only applicable to few food items in certain food groups, and those food items were rarely consumed. An example of this is the facet ‘Enriched/fortified’ in the food group ‘Cakes and sweet biscuits’. A second reason is a lack of variation in the chosen descriptors within a facet. An example of this is the facet ‘source’ in dairy products since in the Netherlands cow milk is the basis for the majority of the consumed dairy products. Another possible explanation for the less predictive facets is the use of a generic food composition database NEVO [27]. Some facets might have been important for predicting true nutrient levels but not for averaged nutrient levels of a generic food. An example of this could be brand name for predicting salt levels of industrially processed foods [28]. In contrast, some facets showed strong predictive power in estimating nutrient contents in certain food groups. The facet ‘Type of packing’ was predicts strongly for the ‘Fats and oils’ group, because the type of packing materials distinguishes solid from liquid fat, which results in different nutrient contents, specifically for fat content. Similarly, as can be expected from a nutrition point of view, facet ‘Physical state’, ‘Sugar’ and ‘Fat content’ were strong predictors for most of their allocated food groups, except for unprocessed products (e.g. fruit, meat, and fish).

In terms of comparing nutrient intake distributions before and after the facets had been deleted, difference of less than 10% was found for most nutrients. This could be explained by the fact that 96.3% of the combinations were relinked to a same food code in the food composition database. Apparently, the food name and remaining facets provided sufficient information to link to the same food item in the Dutch National Food Composition Database. For those combinations with deleted facets that were linked to different food codes in the food composition database, the difference in nutrient contents of the original and alternative food codes may have been small, or the foods were consumed by few persons or in small amounts and therefore did not influence population nutrient intake distributions substantially. A similar finding was observed in a study that investigated the effect of a concise versus an extensive food list in a self-administered web-based 24HR tool. They observed that the differences between population nutrient intakes assessed by two methods were less than 6%

[29], which is consistent with our study that the majority of the differences fell below 5% before and after facet deletion.

Specifically, a large decrease in the amount of vitamin C was found for children in our study, the reason was speculated to be the deletion of the facet ‘Enriched/fortified’ in the food group ‘Non-alcoholic beverages’. According to the report of 2007-2010 survey, ‘Non-alcoholic beverages’ and ‘Meat and meat products’ together, contribute for one third to the total vitamin C intake partly due to food fortification and processing [24]. Hence, beverages with fortification were linked to NEVO codes for products without fortification and resulted a lower vitamin C content. On the other hand, a large increase in the amount of vitamin B group was found for children. A possible explanation would be the deletion of ‘Flavoured component’ in the food group of ‘Cereal’, which may have caused the linkage of NEVO codes high in vitamin B contents (i.e. whole wheat cereals) with flavoured regular cereals. A closer investigation should be conducted before deleting facets in the real setting.

To our knowledge, this is the first study investigating the impact of reducing food descriptions in interview-based 24HRs for the estimation of population nutrient intake distributions. A strength of our approach is that both the identification of facets to be deleted as well as assessing its impact was data driven. Until now decisions on the facets that were included in the 24HR interview of DNFCS were based on expert-judgment. Another strength is the use of the random forest for the identification of unimportant facets. This prediction model is more efficient in large datasets, has a lower risk of overfitting and is better in dealing with correlated predictors than multiple linear regression [30]. However, the applied random forest implementation only allows nominal variables with a limited number of levels as predictors. Therefore, the nominal variable “food ID” was treated as a continuous variable, and the importance of the information on the full brand and product name of each food could not be evaluated. In addition, importance of the facet “Cooking method” could not fully be assessed, since the added fat in case of frying was not included in the nutrient content of the food, but became a separate food item in the food consumption database. Another limitation of our study was the use of a semi-automated protocol of reassigning a different NEVO code to combinations with deleted facets rather than applying the original approach of ‘manual’ linkage by dieticians. Manual matching, however, would only have further decreased the effect of facet deletion, so we do not think our conclusions would have been different. Finally, the impact of facet reduction on respondents’ answers during the food description part of the interview was not assessed. Although a face-to-face or telephone 24HR interview has generally smaller self-reporting error than other methods, measurement error is likely to be present (i.e. rely on memory, underreporting) [6]. However, we assume that the effect of facet reduction on self-reporting error will be small.

The scope of our analyses focussed on the importance of facets on nutrient intake distribution. In addition, other aspects are also important in deciding which facets to be deleted. One example is that the facet ’Physical state’ is important during the interview to determine the options for quantifying the consumed foods, e.g. coffee powder is quantified differently than coffee as a beverage. Moreover, the effects of facet omission on the estimation of exposure to food chemicals that are potentially hazardous should be considered for the DNFCS. In principle the procedure described in this manuscript can also be applied to evaluate facet importance for food chemical distributions.

The objective of looking at the reduction in food characteristics was to enhance efficiency in conducting future surveys. Less extensive food description would result in a shorter time needed for the 24HR interview and to match the reported foods to the food composition database. The time needed to complete the facet collection procedure of a 44 minutes 24HR interview was estimated to be 15 minutes. Without 35% of the unimportant facets, the time saved for one interview would on average be 5 minutes. In a survey with 3819 participants that are interviewed twice, a total of 637 hours would be saved. In terms of linking reported foods to the food composition database, time would be saved due to a reduction in unique food-descriptor combinations. The average time needed to link a combination to the food composition table was estimated to be 5 to10 minutes. After deleting less important facets from the unique food-descriptor combination list, the number of unique combinations reduced with 3534 (from 26679 to 23145). In the data handling of DNFCS, only new food-descriptor combinations needed to be linked to the food composition database manually. Therefore, around 442 hours would be saved for data handling to link each unique food-descriptor to the food composition code. To sum up, we estimated that around 1079 hours would be saved for both data collection and handling if facet deletion would be applied.

The current study focused on reducing the number of facets as a potential efficiency measure for a national food consumption survey. For taking final decisions, advantages and disadvantages of alternative efficiency options could be taken into consideration. One alternative is to use 24-h dietary recall software to guide the interviews in which the food list is directly related to the foods in the national food composition database [9, 31]. The reason why GloboDiet did not choose for this option was to give flexibility for new foods that enter the food market (and are not included in food composition databases yet), to standardize food description across different countries that use the same software, and to be able to collect characteristics of food relevant for other purposes than nutrient intake estimations [17]. A more cost-efficient alternative regarding dietary assessment is to use self-administered methods. However, the accuracy and reliability of those tools needs to be further evaluated to be applied in large-scale surveys, due to self-reporting errors, and various levels of acceptance by different age-groups [32]. Thirdly, the matching of food consumption and food composition data could be made more efficient through automatic or semi-automatic linkages. In this study, decisions on NEVO code reassignment for food-descriptor combinations were made based on a simple algorithm with the results of the random forest algorithm. For matching future food consumption data automatically or semi-automatically, random forest prediction models using available previously matched food consumption and food composition data as training dataset could be developed. Two studies have developed a semi-automatic food matching technique using machine-learning and a natural language processing approach. Both studies have tested the effectiveness of the approach as compared to manually link the food items to the food codes by experts, and have shown the approaches to be effective [33, 34]. These procedures need further validation before they can be implemented in a large-scale survey.

Conclusion

In conclusion, the data-driven procedure that combined random forest prediction with a simulation study was successful in identifying less important characteristics of food description. When deleting those less important characteristics, there was little impact on the estimation of the distributions of population nutrient intake for most nutrients, thus yielding a promising approach for saving labour and costs.

Reference

1. Ocke MC, Van Rossum CTM, de Boer EJ. Dietary Surveys: National Food Intake. In: Caballero B, Finglas P, Toldrà F, editors. The Encyclopedia of Food and Health. 2. Oxford: Academic Press; 2016. p. 432-8.
2. Rippin H, Hutchinson J, Evans C, Jewell J, Breda J, Cade J. National nutrition surveys in Europe: a review on the current status in the 53 countries of the WHO European region. Food & Nutrition Research [Internet]. 2018;62. https://doi.org/10.29219/fnr.v62.1362
3. Ahluwalia N, Dwyer J, Terry A, Moshfegh A, Johnson C. Update on NHANES Dietary Data: Focus on Collection, Release, Analytical Considerations, and Uses to Inform Public Policy. Advances in nutrition. 2016;7(1):121-34. https://doi.org/10.3945/an.115.009258
4. de Boer EJ, Slimani N, van 't Veer P, Boeing H, Feinberg M, Leclercq C, et al. Rationale and methods of the European Food Consumption Validation (EFCOVAL) Project. Eur J Clin Nutr. 2011;65 Suppl 1:S1-4. https://doi.org/10.1038/ejcn.2011.82
5. De Keyzer W, Bracke T, McNaughton SA, Parnell W, Moshfegh AJ, Pereira RA, et al. Cross-continental comparison of national food consumption survey methods--a narrative review. Nutrients. 2015;7(5):3587-620. https://doi.org/10.3390/nu7053587
6. EFSA. General principles for the collection of national food consumption data in the view of a pan-European dietary survey. EFSA Journal. 2009;7(12):1435. https://doi.org/10.2903/j.efsa.2009.1435
7. Grandjean AC. Dietary intake data collection: challenges and limitations. Nutrition reviews. 2012;70 Suppl 2:S101-4. https://doi.org/10.1111/j.1753-4887.2012.00545.x
8. Carriquiry AL. Understanding and Assessing Nutrition. Annu Rev Stat Appl. 2017;4:123-46. https://doi.org/10.1146/annurev-statistics-041715-033615
9. Zimmerman TP, Hull SG, McNutt S, Mittl B, Islam N, Guenther PM, et al. Challenges in converting an interviewer-administered food probe database to self-administration in the National Cancer Institute Automated Self-administered 24-Hour Recall (ASA24). Journal of food composition and analysis : an official publication of the United Nations University, International Network of Food Data Systems. 2009;22(Suppl 1):S48-S51. https://doi.org/10.1016/j.jfca.2009.02.003
10. Ocke MC, Slimani N, Brants H, Buurma-Rethans E, Casagrande C, Nicolas G, et al. Potential and requirements for a standardized pan-European food consumption survey using the EPIC-Soft software. Eur J Clin Nutr. 2011;65:S48-S57. https://doi.org/10.1038/ejcn.2011.87
11. Timon CM, van den Barg R, Blain RJ, Kehoe L, Evans K, Walton J, 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):268-80. https://doi.org/10.1017/S0954422416000172
12. Raper N, Perloff B, Ingwersen L, Steinfeldt L, Anand J. An overview of USDA's dietary intake data system. J Food Compos Anal. 2004;17(3-4):545-55. https://doi.org/10.1016/j.jfca.2004.02.013

13. Shim J-S, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiology and Health. 2014;36:e2014009. https://doi.org/10.4178/epih/e2014009
14. Moshfegh AJ, Rhodes DG, Baer DJ, Murayi T, Clemens JC, Rumpler WV, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. American Journal of Clinical Nutrition. 2008;88(2):324-32. https://doi.org/10.1093/ajcn/88.2.324
15. Slimani N, Ferrari P, Ocke M, Welch A, Boeing H, Liere M, et al. Standardization of the 24-hour diet recall calibration method used in the european prospective investigation into cancer and nutrition (EPIC): general concepts and preliminary results. Eur J Clin Nutr. 2000;54(12):900-17.
16. de Boer EJ, Slimani N, van 't Veer P, Boeing H, Feinberg M, Leclercq C, et al. The European Food Consumption Validation Project: conclusions and recommendations. Eur J Clin Nutr. 2011;65 Suppl 1:S102-7. https://doi.org/10.1038/ejcn.2011.94
17. Slimani N, Casagrande C, Nicolas G, Freisling H, Huybrechts I, Ocke MC, et al. The standardized computerized 24-h dietary recall method EPIC-Soft adapted for pan-European dietary monitoring. Eur J Clin Nutr. 2011;65 Suppl 1:S5-15. https://doi.org/10.1038/ejcn.2011.83
18. Slimani N, Deharveng G, Charrondiere RU, van Kappel AL, Ocke MC, Welch A, et al. Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. Computer methods and programs in biomedicine. 1999;58(3):251-66.
19. Steinfeldt L, Anand J, Murayi T. Food Reporting Patterns in the USDA Automated Multiple-Pass Method. Procedia Food Science. 2013;2:145-56. https://doi.org/10.1016/j.profoo.2013.04.022
20. Bakker M, Fransen H, Ocke M, Slob W. Evaluation of the Dutch National Food Consumption Survey with respect to dietary exposure assessment of chemical substances. 2010.
21. Huybrechts I, Geelen A, de Vries JH, Casagrande C, Nicolas G, De Keyzer W, et al. Respondents' evaluation of the 24-h dietary recall method (EPIC-Soft) in the EFCOVAL Project. Eur J Clin Nutr. 2011;65 Suppl 1:S29-37. https://doi.org/10.1038/ejcn.2011.85
22. Zimmerman TP, Potischman N, Douglass D, Dixit-Joshi S, Kirkpatrick SI, Subar AF, et al. The Effect of Editing Open-Ended Text Responses on Nutrient and Food Group Estimates from the Automated Self-Administered 24-Hour Dietary Recall (ASA24). Procedia Food Science. 2015;4:160-72. http://dx.doi.org/10.1016/j.profoo.2015.06.021
23. Stadlmayr B, Wijesinha-Bettoni R, Haytowitz D, Rittenschober D, Cunningham J, Sobolewski R, et al. FAO/INFOODS Guidelines on Food Matching version 1.2. Rome: FAO; 2012.

24. van Rossum CTM, Fransen HP, Verkaik-Kloosterman J, Buurma EM, Ocke MC. Dutch National Food Consumption Survey 2007-2010: Diet of children and adults aged 7 to 69 years. Bilthoven: RIVM; 2011. Report No.: RIVM-report 350070006.
25. NEVO-table [Internet]. Nutrition Center. 2011. https://nevo-online.rivm.nl/. Accessed 23 July 2018.
26. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32. https://doi.org/10.1023/a:1010933404324
27. NEVO-online 2016: background information. In: Westenbrink S, Vliet MJ-vd, editors. Bilthoven: RIVM; 2016.
28. Temme EHM, Hendriksen MAH, Milder IEJ, Toxopeus IB, Westenbrink S, Brants HAM, et al. Salt Reductions in Some Foods in The Netherlands: Monitoring of Food Composition and Salt Intake. Nutrients. 2017;9(7):791. https://doi.org/10.3390/nu9070791
29. Evans K, Hennessy Á, Walton J, Timon C, Gibney E, Flynn A. 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:e46. https://doi.org/10.1017/jns.2017.49
30. Epifanio I. Intervention in prediction measure: a new approach to assessing variable importance for random forests. BMC Bioinformatics. 2017;18:230. 10.1186/s12859-017-1650-8
31. Gavrieli A, Naska A, Konstantinidi C, Berry R, Roe M, Harvey L, et al. Dietary Monitoring Tools for Risk Assessment. EFSA Supporting Publications [Internet]. 2014 01; 11(5). Available from: http://onlinelibrary.wiley.com/doi/10.2903/sp.efsa.2014.EN-607/abstract.
32. Conrad J, Nothlings U. Innovative approaches to estimate individual usual dietary intake in large-scale epidemiological studies. P Nutr Soc. 2017;76(3):213-9. https://doi.org/10.1017/S0029665116003025
33. Koroušić Seljak B, Korošec P, Eftimov T, Ocke M, van der Laan J, Roe M, et al. Identification of Requirements for Computer-Supported Matching of Food Consumption Data with Food Composition Data. Nutrients. 2018;10(4):433. https://doi.org/10.3390/nu10040433
34. Eftimov T, Korosec P, Seljak BK. StandFood: Standardization of Foods Using a Semi-Automatic System for Classifying and Describing Foods According to FoodEx2. Nutrients. 2017;9(6). https://doi.org/10.3390/nu9060542

Additional File 1. Dutch National Food Consumption Survey 2007–2010 EPIC-Soft group classification.

	Food Groups	Food sub-groups
1	Potatoes and other tubers	Unclassified, mixed and other tubers
		Potatoes
2	Vegetables	Unclassified, mixed salad/vegetables
		Leafy vegetables
		Fruiting vegetables
		Root vegetables
		Cabbages
		Mushrooms
		Grain and pod vegetables
		Leek, onion, garlic
		Stalk vegetables, sprouts
3	Legumes	Legumes
4	Fruits, nuts and seeds, olives	Fruits
		Fruit compote
		Nuts, peanuts, seeds
		Peanut butter, nut/seeds spread
		Olives
5	Dairy products and substitutes	Non fermented milk and milk beverages
		Fermented milk, milk beverages and yogurt
		Milk substitutes
		Yoghurt
		Fromage blanc, petits suisses
		Cheeses (including spread cheeses)
		Unclassified creams
		Dairy creams and creamers
		Nondairy creams and creamers
		Ice cream
6	Cereals and cereal products	Flours, starches, flakes, semolina
		Pasta, rice, other grain
		Bread
		Crispbread, rusks
		Breakfast cereals
		Dough and pastry

7	Meat, meat products and substitutes	Unclassified and combined meat
		Unclassified, mixed and other mammals
		Beef
		Veal
		Pork
		Mutton/lamb
		Chicken, hen
		Turkey, young turkey
		Game
		Hot processed meat
		Cold processed meat
		Hot meat substitutes
		Cold meat substitutes
8	Fish, shellfish and amphibians	Unclassified and combined fish products
		Fish
		Crustaceans, mollusks
		Fish products
9	Eggs and egg products	Eggs
10	Fats and oils	Unclassified and combined fats
		Vegetable oils
		Butter
		Margarines and cooking fats
		Other animal fats (including fish oils)
11	Sugar and confectionery	Sugar
		Jam, jelly, marmalade
		Honey
		Other sweet spread
		Syrup
		Unclassified and other chocolate
		Chocolate spread and chocolate powder
		Confectionery non chocolate
12	Cakes and sweet biscuits	Cakes, pies, pastries, puddings
13	Non alcoholic beverages	Unclassified and combined non alc. Drinks
		Fruit and vegetable juices
		Carbonated/soft/isotonic drinks
		Waters
14	Alcoholic beverages	Wine

		Beer
		Spirits
		Liqueurs
		Cocktails
15	Condiments, spices, sauces and yeast	Unclassified or combined condiments
		Other and mixed sauces
		Tomato sauces
		Dressing sauces, mayonnaises and similar
		Mayonnaise based spreads
		Spices, herbs and flavourings
		Unclassified and combined condiments
		Vinegar
16	Soups and stocks	Soups
		Stocks
17	Miscellaneous	Artificial sweeteners
		Meal substitutes
		Savoury snacks, biscuits and crisps
		Savoury filled buns, croissants

Additional File 2. The population two days' average of energy and nutrient intake distributions before and after facets' deletion at cut-off at 0.8. The population means and percentiles of two days' average energy, macro- and micronutrient intake distributions before and after facets' deletion and percent difference (% D) for the Dutch population aged 7 to 69 years (DNFCS 2007–2010) (n = 3819).

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d
Energy(kcal)																		
P5	1457	1457	0.0	1314	1314	0.0	1482	1482	0.0	1121	1122	0.0	1481	1481	0.0	1163	1165	0.2
P25	1859	1849	0.5	1671	1674	0.2	2048	2048	0.0	1539	1539	0.0	2010	2009	0.1	1565	1565	0.0
P50	2226	2226	0.0	1915	1915	0.0	2483	2480	0.1	1837	1837	0.0	2421	2416	0.2	1853	1852	0.0
P75	2653	2653	0.0	2200	2198	0.1	2964	2964	0.0	2223	2225	0.1	2907	2907	0.0	2220	2222	0.1
P95	3676	3667	0.2	2694	2694	0.0	3749	3749	0.0	2834	2834	0.0	3742	3742	0.0	2813	2813	0.0
Mean	2326	2326	0.0	1951	1951	0.0	2534	2535	0.0	1902	1902	0.0	2496	2497	0.0	1911	1911	0.0
Protein(g)																		
P5	41	41	0.1	37	37	0.0	55	55	0.4	42	42	0.0	49	49	0.0	41	41	0.0
P25	56	56	0.0	50	50	0.0	74	74	0.0	57	57	0.0	70	70	0.0	55	55	0.0
P50	69	69	0.0	59	59	0.1	87	87	0.0	68	68	0.0	84	84	0.1	66	66	0.1
P75	86	86	0.0	70	70	0.0	105	105	0.0	82	82	0.0	102	102	0.0	80	80	0.0
P95	114	114	0.0	91	91	0.0	133	133	0.0	102	102	0.0	131	131	0.0	101	101	0.1
Mean	72	72	0.0	61	61	0.0	90	90	0.0	70	70	0.0	87	87	0.0	69	69	0.0
Fat(g)																		
P5	45	45	0.0	40	40	0.4	47	47	0.0	34	34	0.0	47	47	0.0	35	35	0.0
P25	63	63	0.1	56	56	0.1	73	73	0.0	52	52	0.0	71	71	0.1	53	53	0.1
P50	83	83	0.0	70	70	0.5	95	95	0.0	68	69	0.1	92	92	0.0	69	69	0.0
P75	102	102	0.0	86	86	0.0	118	118	0.2	90	90	0.0	116	116	0.0	89	89	0.0
P95	151	151	0.0	112	112	0.0	158	158	0.3	122	123	0.6	158	158	0.0	120	121	0.9
Mean	87	87	0.0	72	72	0.0	98	98	0.0	73	73	0.0	96	96	0.0	73	73	0.0
SFA(g)																		
P5	16	16	0.4	14	14	0.0	16	16	0.0	12	12	0.0	16	16	0.0	12	12	1.0
P25	23	23	0.3	21	21	0.0	27	27	0.1	20	20	0.0	26	26	0.1	20	20	0.0
P50	30	30	0.3	27	27	0.3	34	35	0.3	26	26	0.2	33	34	0.2	26	26	0.1
P75	39	39	0.0	33	33	0.0	44	44	0.2	35	34	0.1	43	43	0.0	34	34	0.0
P95	54	54	0.0	44	44	0.2	60	60	0.0	49	49	0.0	58	58	0.0	48	48	0.0

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
Mean	32	32	0.0	27	27	0.0	36	36	0.0	28	28	0.0	35	35	0.0	28	28	0.0
MUFA(g)																		
P5	14	14	0.0	13	13	0.0	15	15	0.0	10	10	0.0	15	15	0.0	11	11	0.6
P25	22	22	0.0	19	19	0.0	24	24	0.0	17	17	0.0	24	24	0.0	17	17	0.0
P50	29	29	0.0	25	25	0.3	32	32	0.0	23	23	0.1	31	31	0.0	24	24	0.0
P75	37	37	0.0	31	31	0.0	41	41	0.1	30	30	0.1	41	41	0.1	30	31	0.7
P95	55	55	0.0	41	41	0.0	57	57	0.0	44	44	0.0	57	57	0.0	43	43	0.0
Mean	31	31	0.0	25	25	0.1	34	34	0.0	25	25	0.1	33	33	0.0	25	25	0.1
PUFA(g)																		
P5	7	7	0.0	6	6	0.0	8	8	0.0	6	6	0.0	8	8	0.3	6	6	0.0
P25	11	11	0.0	10	10	0.0	13	13	0.2	9	9	0.1	13	13	0.0	9	9	0.1
P50	15	15	0.0	13	13	0.1	19	19	0.3	13	13	0.1	18	18	0.0	13	13	0.2
P75	21	21	0.0	17	17	0.0	25	25	0.1	17	17	0.0	25	25	0.0	17	17	0.0
P95	34	34	0.0	24	24	0.0	37	37	0.0	27	27	0.0	36	36	0.0	26	26	0.0
Mean	17	17	0.0	14	14	0.1	20	20	0.0	14	14	0.0	19	19	0.0	14	14	0.0
TFA(g)																		
P5	0.4	0.4	**1.6**	0.4	0.4	0.0	0.5	0.5	0.1	0.4	0.4	**1.2**	0.5	0.5	1.0	0.4	0.4	0.0
P25	0.8	0.8	0.2	0.7	0.7	0.4	0.9	0.9	0.0	0.7	0.7	0.0	0.9	0.9	0.1	0.7	0.7	0.1
P50	1.1	1.1	0.0	1.0	1.0	0.2	1.4	1.4	0.0	1.1	1.1	0.0	1.4	1.4	0.1	1.1	1.1	0.2
P75	1.7	1.7	0.1	1.5	1.5	0.4	2.0	2.0	0.0	1.7	1.7	0.0	1.9	1.9	0.1	1.6	1.6	0.0
P95	2.9	2.9	0.4	2.5	2.5	0.9	3.5	3.5	0.0	2.9	2.9	0.0	3.3	3.3	0.0	2.9	2.9	0.7
Mean	1.3	1.3	0.0	1.2	1.2	0.3	1.6	1.6	0.1	1.3	1.3	0.0	1.6	1.6	0.0	1.3	1.3	0.0
ALA(g)																		
P5	642	642	0.0	594	594	0.0	861	836	**2.9**	607	607	0.0	804	802	0.2	607	607	0.0
P25	1020	1020	0.0	914	921	0.7	1378	1378	0.0	1010	1010	0.0	1295	1295	0.0	994	994	0.0
P50	1380	1380	0.0	1205	1205	0.0	1906	1906	0.0	1346	1347	0.1	1806	1802	0.2	1313	1313	0.0
P75	1936	1935	0.0	1600	1601	0.1	2514	2514	0.0	1841	1843	0.1	2437	2432	0.2	1793	1793	0.0
P95	3192	3192	0.0	2286	2286	0.0	3762	3762	0.0	2865	2865	0.0	3615	3612	0.1	2790	2790	0.0

2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
Mean	1569	1569	0.0	1303	1304	0.0	2054	2053	0.0	1494	1495	0.0	1966	1965	0.0	1460	1461	0.0
Marine(g)																		
P5	2	2	0.0	2	2	0.0	3	3	**4.7**	3	3	**2.1**	3	3	**4.7**	3	3	**2.1**
P25	8	8	0.0	7	7	0.4	13	13	0.0	11	11	0.9	12	12	**2.1**	10	10	0.1
P50	16	16	0.4	16	16	0.0	28	28	0.0	25	25	0.0	26	26	0.0	23	23	0.0
P75	40	40	0.8	37	37	0.0	118	118	0.0	105	105	0.0	94	94	0.0	79	79	0.0
P95	383	383	0.0	356	356	0.0	958	958	0.0	960	960	0.0	829	829	0.0	891	891	0.0
Mean	77	77	0.1	70	70	0.0	170	170	0.0	173	173	0.0	153	153	0.0	155	155	0.0
Cholesterol(mg)																		
P5	61	60	**1.5**	57	57	0.0	86	86	0.0	65	65	0.0	77	77	0.5	62	62	0.0
P25	105	105	0.2	90	90	0.0	137	137	0.0	112	112	0.0	130	130	0.3	107	107	0.0
P50	147	148	0.7	127	128	0.1	199	199	0.1	161	161	0.0	188	187	0.3	154	154	0.0
P75	209	210	0.4	182	182	0.4	273	275	0.5	231	231	0.0	260	260	0.0	221	222	0.2
P95	339	337	0.3	298	298	0.0	465	465	0.0	378	378	0.0	447	447	0.0	366	366	0.0
Mean	167	167	0.0	146	146	0.3	222	222	0.0	182	182	0.0	212	212	0.0	176	176	0.1
Carbohydrates(g)																		
P5	187	187	0.0	160	161	0.3	147	147	0.0	117	117	0.0	151	151	0.0	121	121	0.0
P25	238	238	0.0	213	214	0.2	214	214	0.0	171	171	0.0	218	218	0.0	176	175	0.1
P50	287	287	0.0	246	246	0.1	266	266	0.0	208	208	0.0	270	270	0.0	216	216	0.0
P75	344	344	0.0	286	286	0.0	330	330	0.0	251	251	0.0	334	334	0.0	260	260	0.0
P95	454	454	0.0	359	359	0.0	436	434	0.6	338	338	0.0	439	439	0.0	341	341	0.0
Mean	299	299	0.0	252	252	0.0	277	277	0.0	215	215	0.0	281	281	0.0	222	222	0.0
Modisac(g)																		
P5	79	79	0.3	65	65	0.0	42	42	0.0	38	38	0.0	46	46	0.0	41	41	0.6
P25	120	120	0.0	103	103	0.0	82	82	0.0	73	73	0.0	86	86	0.0	76	76	0.1
P50	148	148	0.0	131	131	0.1	114	114	0.0	96	95	0.3	122	122	0.0	103	102	0.0
P75	185	185	0.0	161	161	0.1	159	160	0.2	129	129	0.1	166	167	0.3	137	137	0.0
P95	263	263	0.0	218	218	0.0	243	243	0.0	193	195	**1.3**	246	246	0.0	200	200	0.0

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d
Mean	156	156	0.0	134	134	0.0	125	125	0.0	104	103	0.1	131	131	0.0	109	109	0.1
Polysac(g)																		
P5	82	82	0.0	74	74	0.0	80	80	0.0	61	61	0.0	81	81	0.0	63	63	0.0
P25	113	113	0.0	95	95	0.0	118	119	0.1	88	88	0.0	117	118	0.3	90	90	0.0
P50	136	136	0.0	115	115	0.1	148	148	0.1	108	109	0.1	145	145	0.0	110	110	0.0
P75	166	165	0.0	137	137	0.2	180	180	0.0	130	130	0.1	177	178	0.3	131	131	0.1
P95	221	221	0.0	172	172	0.0	234	234	0.0	175	175	0.2	233	233	0.0	175	175	0.1
Mean	142	142	0.0	118	118	0.0	152	152	0.1	112	112	0.0	150	150	0.0	113	113	0.0
Fibre(g)																		
P5	9.8	9.8	0.0	9.0	9.0	0.0	11.6	11.6	0.0	9.6	9.6	0.0	10.9	10.9	0.0	9.5	9.5	0.0
P25	13.9	13.9	0.0	12.6	12.6	0.0	16.9	16.9	0.1	14.2	14.2	0.3	16.3	16.3	0.3	13.9	13.9	0.0
P50	17.5	17.5	0.1	15.5	15.5	0.1	21.5	21.5	0.0	17.6	17.6	0.0	21.0	21.0	0.3	17.2	17.2	0.0
P75	21.7	21.7	0.0	18.6	18.6	0.0	26.4	26.5	0.2	22.0	22.0	0.0	25.8	25.9	0.1	21.3	21.3	0.0
P95	29.9	30.0	0.1	24.6	24.6	0.0	35.6	35.6	0.1	28.6	28.6	0.0	35.3	35.3	0.0	28.4	28.3	0.3
Mean	18.5	18.5	0.0	16.0	16.0	0.0	22.3	22.3	0.0	18.4	18.4	0.0	21.6	21.6	0.0	18.0	18.0	0.0
Calcium(mg)																		
P5	391	383	**2.0**	332	332	0.0	470	469	0.2	429	429	0.0	448	448	0.0	404	404	0.0
P25	686	685	0.2	616	613	0.4	791	792	0.1	708	708	0.0	769	770	0.0	695	695	0.0
P50	922	922	0.0	828	828	0.0	1092	1092	0.0	944	943	0.1	1049	1050	0.1	924	923	0.1
P75	1232	1232	0.0	1072	1070	0.2	1448	1448	0.0	1195	1193	0.2	1403	1403	0.0	1176	1175	0.1
P95	1832	1832	0.0	1545	1545	0.0	1999	1999	0.0	1702	1702	0.0	1998	1998	0.0	1677	1677	0.0
Mean	995	995	0.1	874	874	0.1	1148	1148	0.0	992	991	0.1	1121	1120	0.0	971	970	0.1
Copper(mg)																		
P5	0.6	0.6	0.0	0.6	0.6	0.0	0.7	0.7	0.0	0.6	0.6	0.0	0.7	0.7	0.0	0.6	0.6	0.0
P25	0.8	0.8	0.0	0.7	0.7	0.0	1.0	1.0	0.0	0.8	0.8	0.3	0.9	0.9	0.0	0.8	0.8	0.0
P50	1.0	1.0	0.0	0.9	0.9	0.0	1.2	1.2	0.0	1.0	1.0	0.2	1.2	1.2	0.0	1.0	1.0	0.0
P75	1.3	1.3	0.0	1.1	1.1	0.0	1.5	1.5	0.2	1.3	1.3	0.2	1.4	1.4	0.2	1.2	1.2	0.0
P95	1.7	1.7	0.5	1.4	1.4	0.1	2.0	2.0	0.0	1.7	1.7	0.0	2.0	2.0	0.0	1.7	1.7	0.2

2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
Mean	1.1	1.1	0.0	0.9	0.9	0.1	1.3	1.3	0.0	1.1	1.1	0.0	1.2	1.2	0.0	1.0	1.0	0.0
Iron(mg)																		
P5	5	5	**1.3**	5	5	0.0	6	6	0.0	6	6	0.0	6	6	0.0	5	5	0.0
P25	7	7	0.0	6	6	0.0	9	9	0.0	8	8	0.0	9	9	0.3	7	7	0.1
P50	9	9	0.1	8	8	0.0	11	11	0.0	9	9	0.1	11	11	0.0	9	9	0.2
P75	11	11	0.0	9	9	0.0	13	13	0.1	11	11	0.1	13	13	0.1	11	11	0.1
P95	15	15	0.6	12	12	0.0	18	18	0.0	15	15	0.0	17	17	0.0	15	15	0.0
Mean	9	9	0.0	8	8	0.0	11	11	0.0	10	10	0.0	11	11	0.0	9	9	0.0
Iodine(µg)																		
P5	122	122	0.0	124	124	0.0	123	123	0.0	120	120	0.0	123	123	0.0	122	122	0.0
P25	196	196	0.0	190	190	0.0	196	196	0.0	185	185	0.0	196	196	0.0	186	186	0.0
P50	240	240	0.0	230	230	0.0	254	254	0.0	238	238	0.0	251	251	0.0	236	236	0.0
P75	297	297	0.0	273	273	0.0	322	322	0.0	292	292	0.0	317	317	0.0	287	287	0.0
P95	389	389	0.0	363	363	0.0	441	441	0.0	413	413	0.0	429	429	0.0	408	408	0.0
Mean	247	247	0.0	236	236	0.0	279	279	0.0	249	249	0.0	273	273	0.0	246	246	0.0
Magnesium(mg)																		
P5	157	157	0.0	143	143	0.1	219	219	0.0	179	178	0.4	190	190	0.0	164	164	0.1
P25	212	212	0.0	194	194	0.0	309	309	0.0	244	244	0.0	284	284	0.1	232	232	0.0
P50	267	267	0.2	236	236	0.0	370	370	0.0	295	294	0.1	355	355	0.1	285	284	0.0
P75	337	337	0.0	280	279	0.2	449	449	0.0	355	355	0.0	436	436	0.0	345	344	0.4
P95	461	461	0.0	366	366	0.0	578	576	0.2	465	473	**1.6**	566	566	0.0	457	457	0.0
Mean	285	285	0.0	241	241	0.1	384	384	0.0	306	306	0.0	366	366	0.0	295	295	0.0
Sodium(mg)																		
P5	1447	1447	0.0	1273	1267	0.5	1625	1622	0.2	1226	1222	0.3	1533	1533	0.0	1234	1224	0.9
P25	1951	1951	0.0	1748	1748	0.0	2357	2357	0.0	1773	1781	0.4	2270	2272	0.1	1767	1772	0.3
P50	2473	2473	0.0	2139	2139	0.0	2970	2963	0.2	2236	2237	0.1	2873	2873	0.0	2225	2225	0.0
P75	3147	3147	0.0	2567	2567	0.0	3630	3630	0.0	2812	2812	0.0	3540	3543	0.1	2782	2782	0.0
P95	4256	4256	0.0	3398	3398	0.0	4792	4792	0.0	3772	3772	0.0	4739	4739	0.0	3668	3668	0.0

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d
Mean	2611	2610	0.0	2195	2195	0.0	3058	3057	0.0	2336	2335	0.0	2977	2976	0.0	2311	2311	0.0
Phosphorus(mg)																		
P5	757	761	0.5	680	680	0.0	982	982	0.0	772	772	0.0	928	928	0.0	750	750	0.0
P25	1038	1038	0.0	940	940	0.0	1350	1351	0.1	1065	1066	0.1	1285	1285	0.0	1029	1029	0.0
P50	1313	1315	0.1	1131	1131	0.0	1637	1637	0.0	1287	1287	0.0	1584	1584	0.0	1264	1264	0.0
P75	1652	1652	0.0	1383	1381	0.1	2023	2023	0.0	1565	1564	0.0	1966	1965	0.1	1545	1544	0.1
P95	2223	2223	0.0	1782	1782	0.0	2641	2641	0.0	2042	2042	0.0	2543	2543	0.0	1981	1981	0.0
Mean	1385	1385	0.0	1176	1176	0.0	1707	1707	0.0	1338	1338	0.0	1649	1649	0.0	1309	1309	0.0
Potassium(mg)																		
P5	2511	2511	0.0	2541	2541	0.0	2728	2728	0.0	2634	2634	0.0	2647	2647	0.0	2630	2630	0.0
P25	3502	3502	0.0	3424	3424	0.0	3898	3898	0.0	3806	3806	0.0	3814	3814	0.0	3713	3713	0.0
P50	4248	4248	0.0	4177	4177	0.0	4981	4981	0.0	4784	4784	0.0	4804	4804	0.0	4671	4671	0.0
P75	4982	4982	0.0	4987	4987	0.0	6386	6386	0.0	6316	6316	0.0	6107	6107	0.0	6046	6046	0.0
P95	6826	6826	0.0	6652	6652	0.0	10584	10584	0.0	9414	9414	0.0	9828	9828	0.0	9058	9058	0.0
Mean	4417	4417	0.0	4304	4304	0.0	5482	5482	0.0	5299	5299	0.0	5288	5288	0.0	5123	5123	0.0
Selenium(µg)																		
P5	19	19	0.5	17	17	0.0	26	26	0.0	21	21	0.0	23	23	0.0	20	20	0.0
P25	27	27	0.2	24	24	0.0	37	37	0.1	29	29	0.0	34	34	0.0	28	28	0.0
P50	34	34	0.0	30	30	0.0	46	46	0.0	37	37	0.0	44	44	0.1	35	35	0.0
P75	45	45	0.0	36	36	0.0	58	58	0.0	47	47	0.0	56	56	0.1	46	46	0.0
P95	64	64	0.3	51	51	0.0	93	93	0.4	69	69	0.0	87	87	0.0	67	66	0.4
Mean	37	37	0.0	32	32	0.0	50	50	0.0	40	40	0.1	48	48	0.0	38	38	0.1
Zinc(mg)																		
P5	5.0	5.0	0.0	4.3	4.3	0.0	6.4	6.4	0.0	5.0	5.0	0.7	5.7	5.7	0.0	4.8	4.8	0.0
P25	6.8	6.8	0.0	6.2	6.2	0.0	9.1	9.1	0.0	7.1	7.1	0.0	8.6	8.6	0.0	6.9	6.9	0.0
P50	8.5	8.5	0.0	7.5	7.5	0.4	10.9	10.9	0.0	8.6	8.6	0.0	10.6	10.6	0.0	8.5	8.5	0.0
P75	10.6	10.6	0.2	8.9	8.9	0.0	13.4	13.4	0.0	10.5	10.5	0.1	13.0	13.0	0.0	10.2	10.2	0.0
P95	14.7	14.7	0.0	11.9	11.9	0.0	17.8	17.8	0.0	13.8	13.8	0.0	17.4	17.4	0.0	13.7	13.7	0.1

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
Mean	9.0	9.0	0.0	7.7	7.7	0.0	11.4	11.4	0.0	9.0	9.0	0.1	11.0	11.0	0.0	8.8	8.8	0.0
RAE(µg)																		
P5	180	179	0.4	179	178	0.4	282	282	0.0	214	214	0.0	259	259	0.0	203	203	0.0
P25	367	368	0.4	325	325	0.0	483	483	0.0	382	382	0.0	460	460	0.0	368	368	0.0
P50	538	537	0.1	460	460	0.0	708	709	0.3	566	567	0.3	668	670	0.3	545	545	0.0
P75	807	819	**1.5**	697	709	**1.7**	1086	1091	0.4	855	853	0.2	1012	1029	**1.7**	828	827	0.1
P95	1703	1703	0.0	1495	1474	**1.4**	2239	2249	0.5	1958	2031	**3.7**	2142	2146	0.2	1879	1879	0.0
Mean	696	702	0.8	611	609	0.3	935	938	0.4	765	766	0.1	891	895	0.4	738	738	0.0
Folate(µg)																		
P5	90	90	0.0	90	91	**1.4**	139	139	0.0	113	113	0.0	124	124	0.0	106	106	0.0
P25	144	144	0.1	134	134	0.0	212	212	0.0	173	172	0.1	193	193	0.2	162	163	0.3
P50	188	188	0.1	170	170	0.1	274	274	0.0	230	230	0.0	260	260	0.0	219	219	0.0
P75	259	259	0.1	223	223	0.0	368	368	0.0	307	307	0.0	352	352	0.0	294	295	0.1
P95	398	398	0.1	335	335	0.0	564	564	0.0	465	465	0.0	550	550	0.0	437	437	0.0
Mean	211	211	0.1	186	186	0.1	304	304	0.0	251	250	0.1	287	287	0.0	239	239	0.1
Vitamin B1(mg)																		
P5	0.4	0.4	0.0	0.4	0.4	0.2	0.6	0.6	0.1	0.4	0.4	0.0	0.5	0.5	0.1	0.4	0.4	0.5
P25	0.7	0.7	0.2	0.6	0.6	0.1	0.8	0.8	0.0	0.7	0.7	0.0	0.8	0.8	0.0	0.7	0.7	0.1
P50	0.9	0.9	0.0	0.8	0.8	0.2	1.1	1.1	0.3	0.9	0.9	0.0	1.1	1.1	0.1	0.9	0.9	0.0
P75	1.3	1.3	0.1	1.1	1.1	0.0	1.5	1.5	0.1	1.2	1.2	0.7	1.5	1.5	0.1	1.2	1.2	0.2
P95	2.0	2.0	0.0	1.6	1.6	0.0	2.4	2.4	0.0	1.9	1.9	0.0	2.3	2.3	0.0	1.9	1.9	0.1
Mean	1.0	1.0	0.0	0.9	0.9	0.0	1.3	1.3	0.0	1.0	1.0	0.2	1.2	1.2	0.0	1.0	1.0	0.2
Vitamin B2(mg)																		
P5	0.6	0.6	0.0	0.5	0.5	0.0	0.8	0.8	0.0	0.6	0.6	0.2	0.7	0.7	0.0	0.6	0.6	0.0
P25	1.0	1.0	0.0	0.9	0.9	0.4	1.2	1.2	0.0	1.0	1.0	1.0	1.2	1.2	0.0	1.0	1.0	1.0
P50	1.4	1.4	0.0	1.2	1.2	0.0	1.6	1.6	0.0	1.3	1.3	0.1	1.6	1.6	0.1	1.3	1.3	0.4
P75	1.9	1.9	0.1	1.6	1.6	0.4	2.2	2.2	0.0	1.7	1.7	**1.2**	2.1	2.1	0.0	1.7	1.7	0.4
P95	2.9	2.9	0.0	2.4	2.4	0.5	3.2	3.2	0.0	2.5	2.5	0.0	3.1	3.1	0.0	2.5	2.5	0.1

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
Mean	1.5	1.5	0.0	1.3	1.3	0.1	1.8	1.8	0.0	1.4	1.4	0.3	1.7	1.7	0.0	1.4	1.4	0.3
Vitamin B6(mg)																		
P5	0.7	0.7	0.7	0.7	0.7	0.0	0.9	0.9	0.0	0.7	0.7	0.2	0.9	0.9	0.2	0.7	0.7	0.1
P25	1.1	1.1	0.1	1.0	1.0	0.0	1.5	1.5	0.6	1.1	1.1	0.6	1.4	1.4	0.0	1.1	1.1	0.3
P50	1.6	1.6	0.0	1.3	1.3	0.5	2.0	2.0	0.1	1.5	1.5	0.0	1.9	1.9	0.0	1.5	1.5	0.3
P75	2.2	2.2	0.0	1.8	1.8	0.0	2.5	2.5	0.0	2.1	2.1	0.0	2.5	2.5	0.1	2.0	2.0	0.4
P95	3.8	3.8	0.0	3.0	3.0	0.0	4.1	4.1	0.0	3.4	3.4	0.0	4.1	4.1	0.0	3.3	3.3	0.4
Mean	1.8	1.8	0.0	1.5	1.5	0.1	2.2	2.2	0.0	1.7	1.7	0.1	2.1	2.1	0.0	1.7	1.7	0.1
Vitamin B12(μg)																		
P5	1.2	1.2	0.6	1.1	1.1	**3.8**	1.8	1.8	0.5	1.3	1.3	0.0	1.7	1.7	0.6	1.3	1.3	0.0
P25	2.3	2.3	0.4	2.0	2.0	0.3	3.2	3.2	0.6	2.4	2.4	0.3	3.0	3.0	0.0	2.3	2.3	0.2
P50	3.3	3.3	0.4	2.8	2.8	0.3	4.5	4.5	0.0	3.4	3.4	0.2	4.3	4.3	0.2	3.3	3.3	0.0
P75	4.8	4.8	0.1	4.0	4.0	0.2	6.3	6.3	0.1	4.9	4.9	0.0	6.1	6.1	0.1	4.7	4.7	0.0
P95	7.6	7.6	0.2	6.2	6.2	0.0	10.7	10.7	0.0	9.3	9.3	0.0	10.4	10.4	0.0	8.6	8.6	0.0
Mean	3.8	3.8	0.0	3.2	3.2	0.2	5.2	5.2	0.1	4.1	4.1	0.1	4.9	4.9	0.1	4.0	4.0	0.0
Vitamin C(mg)																		
P5	24	24	0.0	23	23	0.0	26	27	**2.1**	24	24	0.4	26	26	0.3	24	24	0.4
P25	48	48	0.4	47	47	0.0	50	50	0.2	51	52	**1.9**	50	50	0.0	50	50	0.1
P50	76	76	0.0	73	73	0.0	84	84	0.0	83	83	0.7	82	82	0.0	81	81	0.0
P75	115	115	0.0	110	110	0.1	129	129	0.0	126	126	0.0	127	127	0.0	122	122	0.0
P95	176	176	0.0	169	172	**1.5**	218	222	**1.5**	209	209	0.0	210	210	0.0	204	204	0.0
Mean	86	86	0.1	84	84	0.1	98	98	0.1	95	95	0.1	96	96	0.1	93	93	0.1
Vitamin D(μg)																		
P5	0.7	0.7	0.1	0.6	0.6	0.0	1.1	1.1	**1.0**	0.7	0.7	0.0	1.0	1.0	0.0	0.7	0.7	0.0
P25	1.5	1.5	0.6	1.4	1.4	0.7	2.4	2.4	0.2	1.7	1.7	0.0	2.2	2.2	0.1	1.7	1.6	0.1
P50	2.4	2.4	0.3	2.1	2.1	0.1	3.5	3.5	0.0	2.6	2.6	0.0	3.3	3.3	0.3	2.5	2.5	0.0
P75	3.6	3.6	0.3	2.9	2.9	0.0	4.8	4.8	0.0	3.7	3.7	0.4	4.7	4.7	0.0	3.6	3.6	0.2
P95	5.9	5.9	0.0	4.4	4.4	0.0	7.7	7.7	0.0	6.5	6.5	0.0	7.5	7.5	0.0	6.1	6.0	**2.5**

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
Mean	2.8	2.7	0.2	2.2	2.2	0.1	3.9	3.9	0.0	3.0	3.0	0.1	3.7	3.7	0.0	2.9	2.9	0.1
Vitamin E(mg)																		
P5	5.3	5.2	0.3	4.9	4.9	0.0	6.0	6.0	0.0	4.9	4.9	0.0	5.8	5.7	**1.1**	4.9	4.9	0.0
P25	8.5	8.5	0.4	7.5	7.5	0.0	10.2	10.2	0.0	7.9	7.9	0.0	10.0	10.0	0.1	7.8	7.8	0.0
P50	11.6	11.6	0.1	9.8	9.8	0.0	13.8	13.8	0.4	10.8	10.8	0.0	13.4	13.4	0.0	10.6	10.6	0.0
P75	16.0	16.0	0.0	13.2	13.2	0.0	18.5	18.5	0.0	14.4	14.3	0.3	18.1	18.1	0.0	14.1	14.1	0.5
P95	25.0	25.0	0.0	19.1	18.9	0.9	27.9	27.9	0.0	22.2	21.7	**2.1**	27.1	27.1	0.0	21.1	21.1	0.1
Mean	13.0	13.0	0.0	10.6	10.6	0.1	15.0	15.0	0.0	11.8	11.8	0.2	14.6	14.6	0.0	11.6	11.6	0.2

Percent difference larger than 1% are shown in bold.

Additional File 3-1. The normalized %IncMSEs of the existing facets in each food group calculated by random forests. The normalized percent increase in mean square error (%IncMSEs) (among energy and first 16 nutrients) of the existing facets in each food group calculated by random forests using the data from DNFCS 2007–2010.

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
Potatoes	food_id	0.543	0.545	0.400	0.470	0.486	0.759	0.582	0.533	1.000	0.357	0.640	0.631	0.684	0.630	NA	1.000	0.857
	subgroup A	0.134	0.269	0.054	0.056	0.053	0.087	0.033	0.165	0.013	0.029	0.168	0.367	0.020	0.105	NA	0.361	0.131
	Physical state/form as quantified	0.395	0.736	0.128	0.195	0.142	0.161	0.079	0.150	0.098	0.251	0.513	0.531	0.514	0.488	NA	0.451	0.465
	Cooking method	0.301	0.239	0.258	0.231	0.296	0.491	0.515	0.339	0.283	0.245	0.281	0.188	0.291	0.373	NA	0.483	0.551
	Preservation method	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.115	1.000	1.000	0.594	1.000	1.000	NA	0.649	1.000
	Packing medium	0.014	0.023	0.010	0.015	0.017	-0.006	0.014	-0.003	0.004	0.017	0.012	0.006	0.012	0.012	NA	0.006	0.017
	Brandname(yes/no)	0.278	0.353	0.234	0.237	0.274	0.365	0.256	0.218	0.088	0.213	0.327	0.154	0.342	0.316	NA	0.408	0.368
	Skin consumed	0.218	0.371	0.165	0.163	0.221	0.332	0.320	0.452	0.247	0.148	0.272	0.456	0.282	0.223	NA	0.376	0.277
	Type of fat used	0.170	0.194	0.146	0.163	0.163	0.163	0.123	0.122	0.068	0.249	0.190	0.430	0.191	0.200	NA	0.374	0.236
	Type of milk/liquid used	0.174	0.230	0.151	0.164	0.141	0.143	0.123	0.231	0.056	0.502	0.191	1.000	0.216	0.229	NA	0.804	0.279
Vegetables	food_id	0.927	1.000	0.640	0.805	0.971	0.712	0.433	0.938	0.645	0.495	1.000	1.000	1.000	1.000	NA	0.683	0.484
	subgroup A	0.828	0.765	0.710	0.768	0.736	0.662	0.381	0.892	0.597	0.423	0.778	0.802	0.706	0.913	NA	0.555	1.000
	Physical state/form as quantified	0.831	0.673	0.721	0.815	0.831	0.604	0.487	0.558	0.572	0.444	0.872	0.840	0.682	0.508	NA	0.531	0.233
	Cooking method	0.775	0.607	0.746	0.664	0.877	0.793	0.636	0.858	0.830	0.545	0.775	0.776	0.712	0.765	NA	1.000	0.484
	Preservation method	1.000	0.759	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.758	0.856	0.933	0.684	NA	0.683	0.362
	Packing medium	0.530	0.360	0.329	0.292	0.398	0.442	0.392	0.474	0.537	0.481	0.467	0.510	0.405	0.318	NA	0.472	0.292
	Skin consumed	0.508	0.425	0.272	0.341	0.194	0.198	0.187	0.634	0.165	0.186	0.354	0.385	0.251	0.400	NA	0.373	0.340
Legumes	food_id	0.557	0.397	1.000	1.000	1.000	1.000	NA	1.000	NA	NA	1.000	0.908	0.863	0.528	NA	0.413	1.000
	Physical state/form as quantified	0.154	0.157	0.264	0.258	0.265	0.269	NA	0.146	NA	NA	0.209	0.232	0.229	0.159	NA	0.207	0.280

2

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
	Cooking method	0.204	0.227	-0.015	0.051	0.021	-0.008	NA	0.092	NA	NA	0.114	0.078	0.184	0.064	NA	0.158	0.125
	Preservation method	0.431	0.426	0.060	0.257	0.206	0.017	NA	0.193	NA	NA	0.333	0.319	0.365	0.250	NA	0.207	0.394
	Packing medium	1.000	1.000	0.400	0.444	0.214	0.503	NA	0.880	NA	NA	0.923	1.000	1.000	1.000	NA	1.000	0.372
	Type of fat used	0.149	0.147	0.269	0.251	0.266	0.281	NA	0.104	NA	NA	0.203	0.239	0.172	0.148	NA	0.208	0.267
	Type of milk/liquid used	0.131	0.135	0.259	0.244	0.256	0.268	NA	0.099	NA	NA	0.195	0.220	0.170	0.147	NA	0.201	0.256
Fruits, nuts, olives	food_id	0.643	0.557	0.611	0.600	0.791	0.950	NA	0.887	0.934	1.000	0.554	0.897	1.000	0.807	NA	0.493	0.965
	subgroup A	1.000	1.000	1.000	1.000	1.000	1.000	NA	0.861	0.816	0.275	0.439	0.649	0.903	0.759	NA	0.674	1.000
	Physical state/form as quantified	0.420	0.394	0.355	0.501	0.287	0.513	NA	0.509	0.701	0.085	0.454	0.732	0.436	0.585	NA	0.548	0.426
	Cooking method	0.717	0.502	0.406	0.382	0.738	0.798	NA	0.698	1.000	0.225	0.444	0.680	0.640	0.750	NA	0.392	0.947
	Preservation method	0.773	0.715	0.772	0.771	0.885	0.784	NA	0.605	0.645	0.197	1.000	1.000	0.580	1.000	NA	1.000	0.829
	Packing medium	0.244	0.194	0.174	0.172	0.220	0.181	NA	0.137	0.151	0.039	0.238	0.606	0.284	0.517	NA	0.375	0.385
	Sugar content	0.251	0.217	0.230	0.236	0.293	0.231	NA	0.127	0.090	0.140	0.127	0.241	0.235	0.452	NA	0.318	0.303
	Fat content	0.420	0.403	0.466	0.531	0.567	0.786	NA	1.000	0.556	0.744	0.319	0.524	0.457	0.493	NA	0.280	0.676
	Brandname(yes/no)	0.271	0.271	0.293	0.385	0.371	0.511	NA	0.731	0.665	0.417	0.155	0.427	0.246	0.427	NA	0.361	0.275
	Skin consumed	0.242	0.209	0.206	0.189	0.236	0.222	NA	0.163	0.450	0.089	0.439	0.688	0.295	0.318	NA	0.627	0.408
Dairy (products)	food_id	1.000	0.607	0.629	0.837	0.432	0.248	0.439	0.490	0.530	0.861	1.000	1.000	1.000	0.579	NA	0.727	1.000
	subgroup A	0.930	0.998	0.694	0.838	0.675	0.380	0.565	0.882	0.336	1.000	0.474	0.847	0.458	0.386	NA	0.841	0.390
	subgroup B	0.430	0.276	0.399	0.442	0.370	0.242	0.594	0.529	0.491	0.477	0.097	0.195	0.068	0.056	NA	0.237	0.095
	Source	0.372	0.272	0.313	0.310	0.268	0.176	0.243	0.282	0.013	0.493	0.144	0.328	0.045	0.077	NA	0.435	0.250
	Physical state/form as quantified	0.738	1.000	0.523	0.642	0.454	0.140	0.429	0.661	0.168	0.645	0.297	0.467	0.295	0.229	NA	1.000	0.444
	Cooking method	0.177	0.256	0.101	0.121	0.107	0.110	0.104	0.120	0.023	0.158	0.123	0.128	0.108	0.066	NA	0.201	0.084
	Preservation method	0.298	0.333	0.231	0.236	0.210	0.081	0.292	0.253	0.153	0.290	0.175	0.329	0.153	0.063	NA	0.297	0.166

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
	Packing medium	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	NA	0.000	0.000
	Sugar content	0.358	0.507	0.212	0.258	0.188	0.099	0.267	0.269	0.526	0.419	0.646	0.942	0.354	0.524	NA	0.467	0.160
	Fat content	0.957	0.549	1.000	1.000	1.000	1.000	1.000	1.000	0.678	0.975	0.323	0.913	0.390	1.000	NA	0.636	0.380
	Food production	0.314	0.202	0.169	0.212	0.156	0.216	0.165	0.172	0.598	0.407	0.218	0.421	0.297	0.278	NA	0.191	0.230
	Enriched/fortified	0.334	0.286	0.233	0.259	0.221	0.255	0.197	0.237	0.807	0.336	0.275	0.636	0.156	0.596	NA	0.254	0.277
	Brandname(yes/no)	0.425	0.373	0.288	0.383	0.401	0.429	0.282	0.472	0.229	0.523	0.196	0.449	0.204	0.200	NA	0.357	0.257
	Type of milk/liquid used	0.079	0.041	0.024	0.029	0.022	0.017	0.022	0.013	0.634	0.027	0.191	0.461	0.033	0.206	NA	0.047	0.260
	Flavoured component A	0.199	0.155	0.144	0.147	0.133	0.149	0.158	0.102	0.553	0.241	0.283	0.627	0.186	0.418	NA	0.126	0.302
	Flavoured component B	0.198	0.172	0.097	0.100	0.126	0.119	0.300	0.070	1.000	0.792	0.255	0.465	0.164	0.223	NA	0.061	0.602
	Flavoured component C	0.187	0.196	0.099	0.111	0.089	0.100	0.106	0.087	0.533	0.208	0.308	0.620	0.161	0.350	NA	0.091	0.217
Cereal (products)	food_id	1.000	1.000	0.851	0.551	0.825	1.000	1.000	1.000	1.000	1.000	1.000	0.778	1.000	0.756	NA	0.837	0.781
	subgroup A	0.780	0.679	0.683	0.384	0.472	0.415	0.739	0.397	0.306	0.578	0.737	0.602	0.677	0.576	NA	0.442	0.405
	subgroup B	0.526	0.560	0.313	0.140	0.290	0.346	0.323	0.296	0.160	0.775	0.719	0.339	0.582	0.313	NA	0.197	0.273
	Cooking method	0.421	0.422	0.261	0.203	0.228	0.412	0.432	0.525	0.353	0.729	0.570	1.000	0.432	0.620	NA	0.784	0.654
	Preservation method	0.407	0.296	0.257	0.149	0.204	0.247	0.355	0.275	0.147	0.282	0.363	0.356	0.274	0.339	NA	0.291	0.228
	Packing medium	0.095	0.165	0.242	0.198	0.266	0.019	0.398	0.106	-0.011	0.276	0.215	0.076	0.198	0.023	NA	0.262	0.025
	Sugar content	0.718	0.575	1.000	1.000	1.000	0.668	0.171	0.709	0.930	0.777	0.354	0.525	0.875	0.935	NA	0.352	1.000
	Fat content	0.339	0.375	0.451	0.310	0.654	0.631	0.537	0.373	0.185	0.465	0.611	0.381	0.509	0.915	NA	0.771	0.631
	Food production	0.319	0.261	0.274	0.187	0.325	0.259	0.470	0.333	0.108	0.184	0.297	0.252	0.212	0.387	NA	0.291	0.319
	Enriched/fortified	0.475	0.512	0.509	0.508	0.482	0.473	0.207	0.252	0.845	0.400	0.382	0.493	0.468	0.650	NA	1.000	0.834
	Brandname(yes/no)	0.550	0.461	0.617	0.410	0.442	0.439	0.568	0.875	0.314	0.510	0.601	0.446	0.516	1.000	NA	0.779	0.838
	Type of fat used	-0.039	0.054	-0.045	-0.029	-0.056	-0.023	-0.054	-0.035	0.009	0.010	0.040	0.003	0.024	0.052	NA	0.022	0.020

2

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
	Type of milk/liquid used	0.295	0.143	0.296	0.221	0.234	0.174	0.442	0.259	0.060	0.091	0.213	0.139	0.177	0.129	NA	0.167	0.238
	Flavoured component A	0.288	0.376	0.476	0.327	0.346	0.601	0.196	0.303	0.382	0.219	0.207	0.402	0.473	0.281	NA	0.468	0.681
	Flavoured component B	0.328	0.396	0.394	0.462	0.230	0.406	0.129	0.360	0.257	0.180	0.275	0.450	0.663	0.600	NA	0.730	0.955
	Flavoured component C	0.189	0.270	0.242	0.152	0.151	0.236	0.095	0.193	0.183	0.098	0.226	0.254	0.350	0.409	NA	0.162	0.451
Meat (products)	food_id	1.000	1.000	1.000	1.000	1.000	0.928	1.000	1.000	1.000	1.000	1.000	1.000	0.859	1.000	NA	1.000	1.000
	subgroup B	0.548	0.687	0.579	0.662	0.578	1.000	0.663	0.499	0.508	0.580	0.547	0.771	0.369	0.688	NA	0.683	0.717
	Source	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	NA	0.000	0.000
	Cooking method	0.701	0.855	0.587	0.613	0.585	0.521	0.557	0.536	0.706	0.769	0.950	0.594	1.000	0.547	NA	0.731	0.768
	Preservation method	0.436	0.407	0.417	0.422	0.406	0.465	0.309	0.412	0.469	0.584	-0.007	0.474	0.065	0.422	NA	0.482	0.505
	Packing medium	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	NA	0.000	0.000
	Fat content	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	NA	0.000	0.000
	Brandname(yes/no)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	NA	0.000	0.000
	Skin consumed	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	NA	0.000	0.000
	Visible fat consumed	0.400	0.460	0.476	0.495	0.481	0.410	0.396	0.402	0.406	0.348	0.273	0.464	0.140	0.409	NA	0.372	0.491
Fish and shellfish	food_id	1.000	1.000	1.000	1.000	1.000	1.000	0.835	1.000	1.000	1.000	1.000	0.446	1.000	1.000	0.836	1.000	0.630
	subgroup A	0.388	0.326	0.296	0.316	0.271	0.348	0.370	0.329	0.249	1.000	0.416	0.445	0.436	0.490	0.729	0.941	0.382
	Physical state/form as quantified	0.304	0.394	0.266	0.279	0.216	0.369	0.476	0.374	0.254	0.757	0.346	0.425	0.352	0.387	0.314	0.498	0.321
	Cooking method	0.503	0.542	0.487	0.524	0.454	0.420	0.828	0.432	0.271	0.635	0.840	0.645	0.825	0.986	1.000	0.488	0.517
	Preservation method	0.405	0.390	0.351	0.436	0.424	0.395	0.478	0.410	0.319	0.274	0.374	0.237	0.385	0.355	0.990	0.502	1.000
	Packing medium	0.360	0.256	0.346	0.313	0.325	0.624	1.000	0.694	0.194	0.367	0.263	1.000	0.222	0.237	0.216	0.631	0.640
	Brandname(yes/no)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Skin consumed	0.103	0.135	0.097	0.106	0.109	0.120	0.173	0.109	0.131	0.178	0.354	0.050	0.344	0.211	0.052	0.167	0.079

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
Egg (products)	food_id	1.000	1.000	1.000	1.000	1.000	1.000	NA	1.000	1.000	1.000	1.000	0.969	1.000	NA	NA	1.000	1.000
	Source	0.065	0.043	0.075	0.071	0.055	0.059	NA	0.050	0.064	0.061	-0.088	-0.071	-0.113	NA	NA	-0.002	0.009
	Cooking method	0.115	0.190	0.099	0.142	0.083	0.097	NA	0.149	0.876	0.157	0.892	1.000	0.973	NA	NA	0.218	0.325
Fats and oils	food_id	1.000	1.000	1.000	0.450	1.000	0.924	0.910	1.000	1.000	1.000	0.764	1.000	1.000	0.422	NA	1.000	1.000
	subgroup A	0.659	0.530	0.667	0.409	0.591	0.534	1.000	0.492	0.618	0.819	0.539	0.887	0.466	0.291	NA	0.560	0.547
	Fat content	0.512	0.486	0.500	0.300	0.319	0.291	0.685	0.238	0.257	0.729	0.473	0.419	0.260	1.000	NA	0.271	0.370
	Type of packing	0.817	0.528	0.813	1.000	0.963	1.000	0.885	0.821	0.400	0.680	1.000	0.866	0.899	0.646	NA	0.879	0.309
	Brandname(yes/no)	0.449	0.568	0.450	0.264	0.440	0.376	0.648	0.271	0.511	0.525	0.733	0.503	0.964	0.327	NA	0.975	0.688
Sugar and confectionery	food_id	1.000	1.000	0.671	0.677	0.963	1.000	0.810	0.761	0.885	0.680	0.811	0.856	1.000	1.000	0.195	0.973	0.885
	subgroup A	0.724	0.664	1.000	1.000	1.000	0.818	0.610	0.602	1.000	0.469	0.721	0.735	0.620	0.878	0.229	0.584	1.000
	subgroup B	0.334	0.606	0.641	0.574	0.541	0.337	1.000	0.457	0.116	0.619	0.300	0.316	0.321	0.329	0.189	0.361	0.330
	Sugar content	0.929	0.528	0.748	0.685	0.728	0.641	0.750	1.000	0.412	1.000	1.000	1.000	0.804	0.675	0.072	1.000	0.659
	Enriched/fortified	0.353	0.547	0.217	0.250	0.173	0.217	0.718	0.628	0.340	0.253	0.375	0.368	0.417	0.346	0.067	0.137	0.362
	Brandname(yes/no)	0.470	0.261	0.563	0.737	0.586	0.482	0.535	0.576	0.522	0.457	0.319	0.360	0.514	0.479	0.069	0.964	0.598
	Flavoured component A	0.413	0.443	0.634	0.420	0.782	0.336	0.907	0.259	0.231	0.668	0.346	0.438	0.336	0.390	0.091	0.567	0.588
	Flavoured component B	0.351	0.277	0.484	0.483	0.453	0.362	0.389	0.408	0.165	0.533	0.308	0.325	0.325	0.368	0.030	0.569	0.707
	Flavoured component C	0.324	0.290	0.478	0.463	0.460	0.308	0.774	0.368	0.339	0.607	0.236	0.241	0.299	0.426	1.000	0.446	0.766
Cakes and sweet biscuits	food_id	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	subgroup A	0.507	0.503	0.355	0.357	0.329	0.466	0.438	0.466	0.403	0.418	0.535	0.426	0.228	0.428	0.397	0.483	0.420
	Food production	0.553	0.634	0.507	0.548	0.505	0.786	0.830	0.551	0.664	0.576	0.666	0.812	0.450	0.532	0.328	0.887	0.601
	Enriched/fortified	0.259	0.213	0.150	0.167	0.143	0.225	0.198	0.203	0.160	0.109	0.194	0.175	0.106	0.144	0.053	0.406	0.237
	Brandname(yes/no)	0.505	0.388	0.381	0.379	0.347	0.500	0.443	0.445	0.514	0.335	0.387	0.418	0.336	0.451	0.158	0.615	0.315
	Type of fat used	0.390	0.298	0.256	0.337	0.199	0.415	0.410	0.304	0.334	0.338	0.281	0.215	0.178	0.164	0.138	0.244	0.300

2

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
	Type of milk/liquid used	0.350	0.471	0.217	0.121	0.377	0.406	0.351	0.356	0.221	0.352	0.305	0.315	0.126	0.202	0.344	0.414	0.256
	Flavoured component A	0.479	0.375	0.211	0.236	0.208	0.254	0.222	0.168	0.239	0.238	0.356	0.394	0.193	0.075	0.200	0.323	0.350
	Flavoured component B	0.639	0.288	0.417	0.384	0.443	0.512	0.235	0.338	0.258	0.204	0.421	0.407	0.304	0.214	0.339	0.604	0.609
	Flavoured component C	0.514	0.334	0.262	0.215	0.296	0.443	0.317	0.292	0.296	0.234	0.342	0.475	0.287	0.229	0.138	0.351	0.435
Non-alcoholic beverages	food_id	1.000	0.933	1.000	0.990	1.000	1.000	1.000	0.645	NA	0.963	1.000	0.877	0.618	0.522	0.275	1.000	1.000
	subgroup A	0.516	0.501	0.535	0.578	0.401	0.560	0.598	0.386	NA	0.661	0.462	0.365	0.341	0.348	1.000	0.452	0.568
	subgroup B	0.623	0.782	0.995	1.000	0.704	0.809	0.979	0.505	NA	1.000	0.591	0.521	0.514	0.425	0.548	0.938	0.697
	Physical state/form as quantified	0.347	1.000	0.313	0.418	0.195	0.516	0.169	0.756	NA	0.227	0.705	1.000	0.742	0.315	0.550	0.600	0.402
	Sugar content	0.776	0.649	0.302	0.423	0.185	0.537	0.242	1.000	NA	0.254	0.829	0.761	0.459	0.524	0.077	0.350	0.730
	Food production	0.376	0.969	0.249	0.322	0.105	0.339	0.287	0.887	NA	0.385	0.912	0.777	1.000	1.000	0.266	0.786	0.313
	Enriched/fortified	0.479	0.523	0.177	0.212	0.097	0.217	0.232	0.232	NA	0.197	0.587	0.638	0.402	0.361	0.187	0.249	0.271
	Brandname(yes/no)	0.727	0.589	0.773	0.738	0.484	0.268	0.738	0.520	NA	0.941	0.845	0.674	0.424	0.404	0.203	0.671	0.220
	Flavoured component A	0.106	0.057	0.066	0.058	0.218	0.290	0.000	0.000	NA	0.000	0.150	0.132	0.252	0.048	0.118	0.497	0.000
	Flavoured component B	0.115	0.073	0.080	0.050	0.223	0.267	0.000	0.000	NA	0.000	0.171	0.128	0.252	0.043	0.107	0.491	0.000
	Flavoured component C	0.124	0.075	0.137	0.109	0.263	0.318	0.000	0.000	NA	0.000	0.241	0.225	0.339	0.013	0.200	0.667	0.000
Alcoholic beverages	food_id	0.851	0.943	0.876	1.000	0.941	0.999	1.000	1.000	0.918	0.907	0.625	0.593	0.846	0.619	0.905	0.941	0.872
	subgroup A	1.000	0.931	1.000	0.994	1.000	1.000	0.920	0.950	1.000	1.000	1.000	1.000	1.000	0.554	1.000	1.000	1.000
	Brandname(yes/no)	0.746	1.000	0.655	0.674	0.707	0.777	0.605	0.675	0.606	0.593	0.374	0.468	0.942	1.000	0.992	0.971	0.924
Condiments and sauces	food_id	1.000	0.450	1.000	1.000	1.000	1.000	1.000	0.913	0.764	0.798	0.960	1.000	0.604	1.000	NA	1.000	0.200
	subgroup A	0.251	0.239	0.306	0.240	0.268	0.247	0.154	0.264	0.229	0.288	0.270	0.238	0.240	0.441	NA	0.467	1.000
	subgroup B	0.849	0.262	0.939	0.754	0.924	0.864	0.845	1.000	1.000	1.000	0.536	0.665	0.249	0.601	NA	0.519	0.162

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
	Physical state/form as quantified	0.807	1.000	0.731	0.414	0.778	0.624	0.910	0.797	0.438	0.625	1.000	0.688	1.000	0.797	NA	0.715	0.194
	Food production	0.628	0.520	0.766	0.542	0.722	0.662	0.932	0.675	0.446	0.567	0.689	0.788	0.269	0.440	NA	0.663	0.179
	Brandname(yes/no)	0.555	0.362	0.625	0.594	0.491	0.615	0.872	0.504	0.377	0.370	0.612	0.722	0.304	0.442	NA	0.550	0.093
	Type of fat used	0.335	0.435	0.375	0.428	0.430	0.543	0.896	0.534	0.211	0.488	0.353	0.437	0.250	0.225	NA	0.332	0.773
	Type of milk/liquid used	0.458	0.268	0.490	0.418	0.462	0.440	0.359	0.410	0.175	0.311	0.445	0.436	0.380	0.430	NA	0.657	0.719
Soups	food_id	1.000	1.000	1.000	1.000	1.000	1.000	0.788	1.000	1.000	1.000	1.000	1.000	1.000	1.000	NA	0.795	1.000
	subgroup A	0.448	0.298	0.260	0.211	0.241	0.303	0.339	0.386	0.201	0.295	0.407	0.324	0.386	0.251	NA	0.292	0.279
	Physical state/form as quantified	0.730	0.960	0.970	0.478	0.864	0.773	1.000	0.894	0.989	0.498	0.416	0.892	0.365	0.801	NA	1.000	0.561
	Food production	0.545	0.532	0.702	0.382	0.568	0.469	0.792	0.667	0.597	0.465	0.500	0.798	0.349	0.557	NA	0.528	0.455
	Brandname(yes/no)	0.486	0.312	0.544	0.560	0.495	0.166	0.528	0.308	0.324	0.380	0.474	0.556	0.389	0.388	NA	0.349	0.306
	Type of fat used	0.141	0.176	0.285	0.280	0.154	0.107	0.146	0.287	0.064	0.117	0.069	0.220	0.025	0.068	NA	0.129	0.111
	Type of milk/liquid used	0.449	0.533	0.507	0.491	0.437	0.399	0.345	0.565	0.573	0.347	0.367	0.485	0.184	0.286	NA	0.395	0.328
Miscellaneous	food_id	0.687	0.736	1.000	1.000	1.000	0.976	1.000	0.819	1.000	1.000	0.588	0.549	0.744	1.000	NA	1.000	0.569
	subgroup A	0.385	0.938	0.575	0.576	0.587	0.628	0.590	0.679	0.490	0.614	0.423	0.301	0.550	0.693	NA	0.590	0.398
	subgroup B	0.653	0.924	0.686	0.481	0.386	0.790	0.235	1.000	0.311	0.207	0.588	0.500	0.546	0.973	NA	0.581	0.158
	Physical state/form as quantified	1.000	1.000	0.623	0.448	0.515	0.719	0.456	0.907	0.404	0.356	1.000	0.870	1.000	0.791	NA	0.779	1.000
	Cooking method	0.653	0.417	0.906	0.663	0.727	1.000	0.961	0.944	0.406	0.367	0.743	0.598	0.568	0.563	NA	0.634	0.220
	Preservation method	0.316	0.306	0.281	0.205	0.280	0.316	0.147	0.518	0.249	0.121	0.292	0.248	0.213	0.464	NA	0.336	0.244
	Packing medium	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	NA	0.000	0.000
	Sugar content	0.348	0.269	0.235	0.303	0.371	0.195	0.100	0.248	0.124	0.083	0.215	0.599	0.212	0.233	NA	0.203	0.123
	Fat content	0.242	0.257	0.375	0.223	0.335	0.293	0.167	0.500	0.178	0.146	0.217	0.543	0.221	0.223	NA	0.224	0.145
	Food production	0.162	0.017	0.242	0.190	0.258	0.081	0.194	0.162	-0.012	0.107	0.126	0.205	0.137	0.135	NA	0.363	0.026

2

Groups	Facets	Energy	Protein	Fat	SFA[a]	MUFA[b]	PUFA[c]	TFA[d]	ALA[e]	Marine	Cholesterol	Carbo[f]	Modisac[g]	Polysac[h]	Fibre	Alcohol	Calcium	Copper
	Enriched/fortified	0.373	0.267	0.222	0.302	0.311	0.191	0.110	0.243	0.134	0.074	0.218	1.000	0.264	0.216	NA	0.231	0.122
	Brandname(yes/no)	0.373	0.423	0.738	0.589	0.594	0.447	0.649	0.683	0.470	0.702	0.458	0.273	0.435	0.395	NA	0.462	0.287
	Type of milk/liquid used	0.347	0.425	0.574	0.197	0.293	0.234	0.153	0.306	0.076	0.164	0.405	0.318	0.285	0.357	NA	0.437	0.163
	Flavoured component A	0.242	0.462	0.316	0.196	0.199	0.247	0.104	0.243	0.074	0.098	0.318	0.616	0.246	0.369	NA	0.352	0.188
	Flavoured component B	0.238	0.277	0.286	0.230	0.180	0.207	0.079	0.329	0.102	0.047	0.361	0.631	0.208	0.346	NA	0.201	0.152
	Flavoured component C	0.373	0.561	0.356	0.197	0.344	0.290	0.121	0.483	0.099	0.130	0.433	0.614	0.281	0.354	NA	0.519	0.234

Appendix 3-2. The normalized %IncMSEs of the existing facets in each food group calculated by random forests. The normalized percent increase in mean square error (%IncMSEs) (among energy and last 17 nutrients) of the existing facets in each food group calculated by random forests using the data from DNFCS 2007–2010.

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
Potatoes	food_id	0.558	0.583	0.761	1.000	0.539	0.852	0.842	0.883	1.000	1.000	0.465	0.377	0.330	0.512	1.000	0.492	0.971
	subgroup A	0.052	0.273	0.210	0.040	0.043	0.138	0.253	0.086	0.022	0.318	0.048	0.094	0.424	0.014	0.246	0.010	0.193
	Physical state/form as quantified	0.365	1.000	0.674	0.255	0.673	0.716	0.345	0.714	0.375	0.394	0.228	0.409	1.000	0.511	0.323	0.386	0.365
	Cooking method	0.413	0.148	0.663	0.321	0.231	0.448	0.440	0.667	0.282	0.768	0.321	0.199	0.224	0.309	0.705	0.083	0.503
	Preservation method	0.564	0.685	1.000	0.788	1.000	1.000	1.000	1.000	0.320	0.919	1.000	1.000	0.679	0.360	0.986	0.300	1.000
	Packing medium	0.010	-0.021	0.006	0.010	0.012	0.001	0.014	0.011	0.005	0.013	0.012	0.016	0.015	0.000	0.015	0.000	0.016
	Brandname(yes/no)	0.260	0.690	0.371	0.281	0.361	0.381	0.348	0.396	0.098	0.429	0.218	0.337	0.217	0.179	0.437	0.099	0.404
	Skin consumed	1.000	0.505	0.306	0.159	0.316	0.369	0.272	0.318	0.209	0.324	0.637	0.264	0.606	0.138	0.320	0.157	0.283
	Type of fat used	0.144	0.382	0.232	0.143	0.190	0.244	0.193	0.228	0.832	0.374	0.164	0.254	0.234	0.495	0.223	1.000	0.379
	Type of milk/liquid used	0.172	0.443	0.234	0.168	0.186	0.256	0.343	0.260	0.552	0.564	0.313	0.461	0.220	1.000	0.252	0.299	0.235
Vegetables	food_id	0.441	0.752	0.578	1.000	1.000	1.000	0.922	0.526	0.920	0.321	1.000	0.647	1.000	0.509	1.000	0.481	1.000
	subgroup A	0.388	0.582	0.461	0.616	0.721	0.716	1.000	0.416	1.000	0.380	0.927	1.000	0.696	0.436	0.922	0.413	0.800
	Physical state/form as quantified	0.299	0.354	0.238	0.800	0.628	0.435	0.608	0.477	0.296	0.188	0.581	0.439	0.696	0.480	0.473	0.435	0.512
	Cooking method	0.677	0.485	1.000	0.776	0.476	0.642	0.914	0.669	0.550	1.000	0.693	0.583	0.906	0.565	0.564	0.537	0.803
	Preservation method	0.465	1.000	0.544	0.693	0.598	0.866	0.706	0.460	0.684	0.226	0.844	0.545	0.881	1.000	0.615	1.000	0.939
	Packing medium	0.381	-0.054	0.521	0.521	0.273	0.527	0.469	0.286	0.477	0.168	0.492	0.337	0.656	0.501	0.413	0.477	0.870
	Skin consumed	1.000	0.030	0.466	0.196	0.226	0.479	0.311	1.000	0.348	0.835	0.778	0.345	0.401	0.187	0.264	0.184	0.529
Legumes	food_id	0.965	0.960	1.000	0.940	1.000	1.000	1.000	1.000	0.374	1.000	1.000	1.000	0.957	NA	NA	NA	1.000
	Physical state/form as quantified	0.231	0.000	0.283	0.206	0.339	0.230	0.158	0.239	0.185	0.272	0.294	0.274	0.266	NA	NA	NA	0.269

2

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
	Cooking method	0.191	0.130	-0.035	0.214	0.125	0.196	0.095	0.090	0.147	0.025	0.310	-0.018	0.023	NA	NA	NA	-0.022
	Preservation method	0.406	0.444	0.096	0.589	0.501	0.403	0.087	0.396	0.192	0.460	0.610	0.143	0.317	NA	NA	NA	0.110
	Packing medium	1.000	1.000	0.286	1.000	0.374	0.367	0.224	0.187	1.000	0.336	0.432	0.743	1.000	NA	NA	NA	0.991
	Type of fat used	0.216	0.000	0.272	0.202	0.309	0.178	0.148	0.242	0.193	0.257	0.248	0.281	0.264	NA	NA	NA	0.269
	Type of milk/liquid used	0.178	0.000	0.255	0.182	0.304	0.181	0.139	0.219	0.199	0.250	0.256	0.263	0.249	NA	NA	NA	0.264
Fruits, nuts, olives	food_id	1.000	1.000	0.617	1.000	0.596	1.000	1.000	0.778	1.000	1.000	1.000	0.687	1.000	NA	1.000	NA	0.848
	subgroup A	0.757	0.377	1.000	0.916	1.000	0.602	0.853	1.000	0.477	0.412	0.976	1.000	0.414	NA	0.377	NA	1.000
	Physical state/form as quantified	0.354	0.352	0.385	0.495	0.395	0.453	0.491	0.338	0.473	0.329	0.388	0.495	0.485	NA	0.395	NA	0.246
	Cooking method	0.837	0.498	0.858	0.672	0.787	0.930	0.727	0.855	0.574	0.606	0.859	0.618	0.415	NA	0.672	NA	0.520
	Preservation method	0.937	0.450	0.752	0.732	0.710	0.794	0.750	0.743	0.270	0.403	0.859	0.816	0.431	NA	0.915	NA	0.790
	Packing medium	0.299	0.395	0.264	0.195	0.255	0.522	0.325	0.219	0.264	0.349	0.354	0.594	0.256	NA	0.445	NA	0.247
	Sugar content	0.345	0.289	0.252	0.145	0.235	0.405	0.259	0.252	0.312	0.292	0.337	0.565	0.204	NA	0.201	NA	0.214
	Fat content	0.669	0.144	0.669	0.513	0.746	0.459	0.661	0.737	0.164	0.281	0.797	0.547	0.229	NA	0.101	NA	0.632
	Brandname(yes/no)	0.434	0.176	0.302	0.382	0.315	0.301	0.282	0.317	0.226	0.284	0.644	0.591	0.159	NA	0.152	NA	0.421
	Skin consumed	0.461	0.586	0.309	0.244	0.229	0.615	0.373	0.266	0.322	0.492	0.708	0.244	0.512	NA	0.823	NA	0.245
Dairy (products)	food_id	1.000	1.000	0.732	0.767	1.000	1.000	0.494	0.424	0.622	0.693	0.511	0.656	0.371	0.751	0.729	0.795	0.307
	subgroup A	0.564	0.549	0.775	1.000	0.940	0.700	1.000	1.000	0.912	1.000	0.222	0.547	0.185	0.666	0.389	0.669	0.234
	subgroup B	0.108	0.124	0.356	0.222	0.199	0.139	0.185	0.192	0.645	0.199	0.216	0.286	0.034	0.178	0.176	0.911	0.097
	Source	0.281	0.260	0.324	0.532	0.468	0.160	0.197	0.297	0.475	0.361	0.104	0.207	0.077	0.366	0.189	0.461	0.155
	Physical state/form as quantified	0.543	0.411	0.397	0.625	0.643	0.342	0.537	0.652	0.570	0.463	0.406	0.271	0.165	0.517	0.402	0.589	0.109
	Cooking method	0.119	0.051	0.100	0.134	0.173	0.129	0.144	0.150	0.144	0.180	0.057	0.081	0.042	0.148	0.183	0.189	0.079
	Preservation method	0.127	0.106	0.151	0.333	0.263	0.105	0.305	0.285	0.354	0.255	0.060	0.084	0.052	0.208	0.077	0.281	0.066

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
	Packing medium	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Sugar content	0.369	0.251	0.377	0.421	0.437	0.294	0.353	0.340	0.282	0.402	0.173	0.793	1.000	0.306	0.575	0.417	0.163
	Fat content	0.771	0.687	0.897	0.505	0.798	0.486	0.633	0.315	1.000	0.962	1.000	1.000	0.391	1.000	1.000	1.000	1.000
	Food production	0.339	0.216	0.359	0.173	0.241	0.353	0.238	0.146	0.213	0.335	0.144	0.417	0.154	0.205	0.310	0.354	0.219
	Enriched/fortified	0.413	0.166	0.269	0.207	0.247	0.254	0.261	0.214	0.298	0.401	0.573	0.864	0.807	0.285	0.892	0.359	0.269
	Brandname(yes/no)	0.334	0.283	0.346	0.475	0.566	0.305	0.406	0.224	0.650	0.490	0.116	0.380	0.172	0.415	0.167	0.854	0.107
	Type of milk/liquid used	0.128	0.263	0.507	0.032	0.030	0.194	0.020	0.050	0.172	0.265	0.120	0.574	0.150	0.298	0.157	0.034	0.057
	Flavoured component A	0.652	0.191	0.291	0.134	0.164	0.246	0.169	0.121	0.170	0.366	0.247	0.523	0.230	0.159	0.478	0.244	0.205
	Flavoured component B	0.981	0.362	1.000	0.121	0.094	0.146	0.981	0.190	0.147	0.581	0.081	0.388	0.101	0.395	0.316	0.730	0.417
	Flavoured component C	0.718	0.173	0.265	0.166	0.141	0.211	0.174	0.118	0.123	0.427	0.128	0.380	0.117	0.203	0.393	0.209	0.141
Cereal (products)	food_id	0.643	0.747	0.583	1.000	0.412	1.000	1.000	0.552	0.561	0.514	0.584	0.290	0.568	0.398	0.660	0.680	1.000
	subgroup A	0.466	0.688	0.414	0.937	0.308	0.621	0.650	0.420	0.450	0.321	0.495	0.337	0.355	0.385	0.247	0.213	0.464
	subgroup B	0.187	1.000	0.211	0.560	0.183	0.303	0.475	0.203	0.833	0.130	0.193	0.123	0.166	0.136	0.111	0.102	0.169
	Cooking method	0.489	0.458	0.645	0.436	0.415	0.373	0.474	0.507	0.571	0.379	0.419	0.316	0.229	0.216	0.177	0.114	0.393
	Preservation method	0.288	0.315	0.225	0.340	0.177	0.454	0.349	0.170	0.204	0.232	0.240	0.189	0.150	0.059	0.075	0.016	0.161
	Packing medium	0.040	0.036	0.011	0.288	0.016	0.056	0.188	0.031	0.000	0.071	0.035	0.030	0.022	0.199	0.236	0.270	0.009
	Sugar content	0.545	0.102	0.948	0.699	0.727	0.800	0.869	0.877	-0.038	0.391	0.458	0.330	0.339	0.275	0.515	-0.128	0.427
	Fat content	0.551	0.195	0.783	0.480	0.351	0.899	0.631	0.652	0.989	0.171	0.517	0.243	0.603	0.095	1.000	0.033	0.465
	Food production	0.315	0.219	0.299	0.333	0.255	0.163	0.350	0.394	0.682	0.257	0.312	0.214	0.062	0.027	0.123	0.380	0.098
	Enriched/fortified	1.000	0.365	0.692	0.800	1.000	0.952	0.483	0.636	0.895	1.000	1.000	1.000	1.000	1.000	0.410	0.865	0.452
	Brandname(yes/no)	0.711	0.572	1.000	0.792	0.684	0.627	0.421	1.000	0.977	0.318	0.725	0.310	0.458	0.512	0.518	1.000	0.513
	Type of fat used	0.055	0.045	0.029	-0.047	0.025	0.044	-0.045	0.034	-0.143	-0.022	-0.007	0.000	0.016	0.007	0.041	0.011	0.023

2

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
	Type of milk/liquid used	0.207	0.102	0.172	0.172	0.154	0.089	0.162	0.220	1.000	0.110	0.274	0.207	0.042	0.022	0.060	0.076	0.054
	Flavoured component A	0.256	0.119	0.676	0.382	0.499	0.745	0.401	0.529	0.273	0.206	0.231	0.148	0.132	0.192	0.181	0.247	0.274
	Flavoured component B	0.339	0.117	0.612	0.454	0.601	0.722	0.448	0.529	0.432	0.335	0.357	0.319	0.287	0.356	0.483	0.413	0.294
	Flavoured component C	0.233	0.064	0.449	0.266	0.345	0.535	0.197	0.441	0.238	0.160	0.246	0.195	0.177	0.190	0.225	0.165	0.195
Meat (products)	food_id	1.000	1.000	1.000	0.949	1.000	1.000	1.000	0.870	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	subgroup B	0.982	0.469	0.665	1.000	0.713	0.615	0.573	0.762	0.628	0.643	0.840	0.686	0.545	0.748	0.756	0.910	0.559
	Source	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Cooking method	0.885	0.898	0.816	0.537	0.744	0.812	0.767	1.000	0.614	0.751	0.460	0.867	0.787	0.998	0.872	0.397	0.451
	Preservation method	0.699	0.231	0.484	0.437	0.461	0.464	0.391	0.560	0.358	0.418	0.440	0.385	0.513	0.617	0.528	0.425	0.478
	Packing medium	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Fat content	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Brandname(yes/no)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Skin consumed	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Visible fat consumed	0.517	0.332	0.446	0.284	0.397	0.463	0.552	0.516	0.409	0.447	0.365	0.487	0.457	0.495	0.449	0.301	0.300
Fish and shellfish	food_id	1.000	1.000	0.167	1.000	1.000	1.000	1.000	0.727	1.000	0.973	0.601	1.000	1.000	1.000	0.264	1.000	1.000
	subgroup A	0.395	0.448	0.117	0.742	0.594	0.725	0.169	0.515	0.457	0.459	0.260	0.574	0.266	0.711	0.167	0.339	0.357
	Physical state/form as quantified	0.524	0.449	0.130	0.454	0.685	0.526	0.264	0.449	0.497	0.553	0.279	0.798	0.406	0.753	0.327	0.289	0.581
	Cooking method	0.382	0.692	0.137	0.289	0.732	0.573	0.287	0.729	0.445	0.715	1.000	0.857	0.433	0.601	0.921	0.633	0.747
	Preservation method	0.461	0.507	0.384	0.692	0.801	0.804	0.325	0.589	0.722	1.000	0.331	0.321	0.644	0.364	1.000	0.595	0.579
	Packing medium	0.411	0.398	1.000	0.289	0.632	0.629	0.304	1.000	0.270	0.477	0.143	0.414	0.404	0.362	0.635	0.581	0.365
	Brandname(yes/no)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Skin consumed	0.267	0.254	0.244	0.369	0.186	0.225	0.125	0.054	0.154	0.160	0.393	0.331	0.137	0.324	0.285	0.254	0.201

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
Egg (products)	food_id	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.932	1.000	0.893	1.000	1.000	NA	1.000	1.000
	Source	0.008	-0.135	-0.136	0.097	0.003	0.097	-0.131	-0.031	-0.035	-0.005	0.084	0.120	-0.119	0.077	NA	0.047	-0.122
	Cooking method	0.225	0.294	0.908	-0.073	0.218	0.214	0.886	0.244	0.236	1.000	0.024	1.000	0.899	0.090	NA	0.625	0.958
Fats and oils	food_id	1.000	1.000	0.957	0.891	1.000	1.000	1.000	0.756	0.997	0.560	1.000	1.000	0.876	0.706	NA	0.186	0.441
	subgroup A	0.833	0.908	1.000	0.560	0.853	0.756	0.559	0.693	1.000	0.524	0.521	0.493	0.519	0.495	NA	0.461	0.537
	Fat content	0.625	0.519	0.731	0.558	0.760	0.663	0.326	1.000	0.770	0.235	0.253	0.263	0.240	0.312	NA	1.000	0.305
	Type of packing	0.610	0.403	0.614	1.000	0.565	0.510	0.230	0.406	0.643	0.872	0.871	0.898	0.898	1.000	NA	0.551	1.000
	Brandname(yes/no)	0.404	0.405	0.411	0.531	0.384	0.384	0.513	0.547	0.417	1.000	0.978	0.973	1.000	0.982	NA	0.209	0.284
Sugar and confectionery	food_id	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.706	1.000	1.000	0.843	0.708	1.000	1.000	0.204	0.000
	subgroup A	0.500	0.891	0.516	0.508	0.615	0.626	0.946	0.799	0.942	0.923	0.976	1.000	0.729	0.611	0.896	0.375	0.398
	subgroup B	0.235	0.570	0.461	0.271	0.429	0.354	0.679	0.372	1.000	0.108	0.333	0.687	0.104	0.742	0.000	0.483	0.115
	Sugar content	0.568	0.569	0.781	0.649	0.382	0.532	0.457	0.481	0.312	0.262	0.557	0.798	0.707	0.752	0.392	0.524	0.268
	Enriched/fortified	0.245	0.284	0.380	0.338	0.294	0.280	0.235	0.310	0.259	0.374	0.226	0.321	1.000	0.208	0.558	0.661	0.181
	Brandname(yes/no)	0.257	0.639	0.696	0.308	0.356	0.495	0.953	0.501	0.331	0.446	0.517	0.611	0.813	0.876	0.447	0.778	0.213
	Flavoured component A	0.463	0.341	0.585	0.272	0.621	0.568	0.441	0.528	0.497	0.211	0.439	0.429	0.136	0.364	1.000	0.180	0.173
	Flavoured component B	0.712	0.336	0.621	0.193	0.270	0.671	0.368	0.304	0.527	0.207	0.240	0.490	0.133	0.504	0.485	0.695	0.188
	Flavoured component C	0.569	0.351	0.516	0.164	0.351	0.562	0.510	0.286	0.617	0.170	0.247	0.542	0.127	0.509	0.000	0.277	0.179
Cakes and sweet biscuits	food_id	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.427	0.834	1.000
	subgroup A	0.367	0.490	0.317	0.295	0.377	0.310	0.395	0.385	0.372	0.384	0.549	0.708	0.380	0.339	0.861	1.000	0.600
	Food production	0.672	0.274	0.519	0.621	0.516	0.653	0.763	0.543	0.484	0.597	0.470	0.571	0.754	0.464	0.587	0.401	0.688
	Enriched/fortified	0.303	0.198	0.181	0.120	0.160	0.174	0.113	0.100	0.100	0.133	0.499	0.357	0.445	0.173	0.624	1.000	0.158
	Brandname(yes/no)	0.491	0.183	0.353	0.364	0.325	0.319	0.414	0.470	0.323	0.388	0.842	0.617	0.486	0.293	0.199	0.407	0.486

2

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
	Type of fat used	0.265	0.124	0.267	0.165	0.311	0.276	0.322	0.278	0.331	0.299	0.409	0.448	0.224	0.250	0.650	0.000	0.383
	Type of milk/liquid used	0.112	0.343	0.259	0.234	0.406	0.267	0.363	0.338	0.215	0.433	0.594	0.573	0.115	0.395	1.000	0.462	0.438
	Flavoured component A	0.287	0.355	0.243	0.311	0.348	0.330	0.301	0.281	0.229	0.316	0.673	0.330	0.385	0.211	0.687	0.046	0.246
	Flavoured component B	0.455	0.366	0.249	0.283	0.235	0.314	0.270	0.231	0.189	0.336	0.663	0.650	0.419	0.256	0.181	0.296	0.398
	Flavoured component C	0.254	0.431	0.288	0.342	0.328	0.375	0.287	0.174	0.180	0.388	0.206	0.356	0.303	0.255	0.178	0.278	0.325
Non-alcoholic beverages	food_id	1.000	0.522	0.974	0.624	0.983	0.749	0.506	1.000	1.000	0.774	0.879	0.856	1.000	0.419	1.000	0.984	1.000
	subgroup A	0.554	0.334	0.570	0.509	0.655	0.419	0.259	0.601	0.864	0.378	0.665	0.507	0.313	0.297	0.585	0.684	0.721
	subgroup B	0.925	0.441	0.825	0.569	0.874	0.654	0.497	0.953	0.676	0.448	0.264	0.590	0.103	0.498	0.163	1.000	0.183
	Physical state/form as quantified	0.284	1.000	0.810	0.506	0.717	0.532	0.943	0.388	0.756	0.865	0.976	0.302	0.438	0.521	0.405	0.202	0.972
	Sugar content	0.583	0.263	1.000	1.000	0.758	0.728	0.243	0.511	0.998	1.000	1.000	0.625	0.454	0.266	0.623	0.260	0.649
	Food production	0.364	0.693	0.734	0.302	0.712	1.000	1.000	0.459	0.951	0.562	0.989	1.000	0.861	1.000	0.504	0.368	0.748
	Enriched/fortified	0.290	0.635	0.626	0.392	0.750	0.513	0.429	0.451	0.666	0.578	0.857	0.690	0.776	0.641	0.763	0.188	0.754
	Brandname(yes/no)	0.450	0.540	0.699	0.741	1.000	0.450	0.501	0.825	0.742	0.376	0.835	0.578	0.640	0.532	0.322	0.956	0.986
	Flavoured component A	0.000	0.087	0.118	0.164	0.367	0.087	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Flavoured component B	0.000	0.030	0.071	0.221	0.313	0.079	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Flavoured component C	0.000	0.049	0.109	0.183	0.552	0.102	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Alcoholic beverages	food_id	0.944	0.845	0.969	0.795	0.870	1.000	0.928	0.916	0.983	0.861	1.000	0.857	0.863	0.947	1.000	1.000	1.000
	subgroup A	1.000	1.000	1.000	0.896	0.852	0.979	1.000	0.991	1.000	0.881	0.902	0.835	0.819	0.962	0.616	0.991	0.850
	Brandname(yes/no)	1.000	0.847	0.675	1.000	1.000	0.838	0.825	1.000	0.746	1.000	0.936	1.000	1.000	1.000	0.522	0.716	0.839
Condiments and sauces	food_id	0.457	0.187	0.752	0.991	0.179	0.515	1.000	1.000	1.000	0.163	0.241	0.168	0.735	0.166	1.000	1.000	1.000
	subgroup A	0.209	0.888	0.222	0.439	1.000	0.428	0.306	0.425	0.294	1.000	0.598	1.000	0.913	1.000	0.439	0.242	0.263
	subgroup B	0.500	0.144	0.438	0.393	0.124	0.428	0.638	0.489	0.692	0.101	0.096	0.109	0.737	0.082	0.483	0.468	0.812

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
	Physical state/form as quantified	0.983	0.865	1.000	1.000	0.361	1.000	0.607	0.850	0.600	0.525	1.000	0.378	1.000	0.377	0.588	0.597	0.569
	Food production	1.000	0.268	0.758	0.811	0.237	0.854	0.453	0.625	0.677	0.476	0.390	0.475	0.576	0.176	0.464	0.567	0.784
	Brandname(yes/no)	0.695	0.523	0.603	0.411	0.333	0.619	0.428	0.517	0.568	0.454	0.254	0.330	0.744	0.449	0.467	0.576	0.521
	Type of fat used	0.464	0.331	0.287	0.286	0.321	0.350	0.235	0.310	0.336	0.157	0.410	0.181	0.243	0.106	0.253	0.344	0.474
	Type of milk/liquid used	0.566	1.000	0.506	0.407	0.414	0.499	0.387	0.570	0.416	0.116	0.423	0.159	0.428	0.143	0.234	0.373	0.498
Soups	food_id	1.000	1.000	0.847	0.975	1.000	1.000	0.843	1.000	0.826	0.799	0.833	1.000	0.773	1.000	0.895	1.000	0.964
	subgroup A	0.290	0.449	0.231	0.523	0.251	0.250	0.214	0.256	0.294	0.203	0.235	0.276	0.261	0.187	0.227	0.280	0.326
	Physical state/form as quantified	0.578	0.527	1.000	0.973	0.510	0.557	1.000	0.975	1.000	1.000	1.000	0.891	1.000	0.357	1.000	0.608	1.000
	Food production	0.491	0.502	0.592	0.917	0.555	0.540	0.483	0.604	0.686	0.513	0.633	0.547	0.578	0.309	0.529	0.653	0.594
	Brandname(yes/no)	0.328	0.437	0.391	0.817	0.482	0.497	0.294	0.323	0.319	0.282	0.450	0.252	0.298	0.374	0.329	0.371	0.390
	Type of fat used	0.132	0.226	0.131	0.376	0.118	0.153	0.098	0.131	0.144	0.184	0.101	0.206	0.191	0.028	0.195	0.309	0.068
	Type of milk/liquid used	0.346	0.584	0.388	1.000	0.382	0.416	0.485	0.533	0.555	0.466	0.318	0.532	0.525	0.261	0.475	0.510	0.326
Miscellaneous	food_id	0.868	0.767	0.692	0.471	1.000	0.480	0.593	0.446	0.995	0.689	0.925	1.000	0.998	1.000	0.970	0.599	0.886
	subgroup A	0.633	0.351	0.527	0.373	0.590	0.427	0.392	0.217	0.279	0.498	0.406	0.366	0.397	0.446	0.449	0.207	0.510
	subgroup B	1.000	0.621	0.558	0.706	0.608	0.517	0.160	0.267	0.441	0.127	0.701	0.570	0.517	0.690	0.516	0.201	0.218
	Physical state/form as quantified	0.884	0.429	1.000	1.000	0.720	1.000	1.000	1.000	1.000	1.000	0.825	0.850	1.000	0.697	1.000	1.000	1.000
	Cooking method	0.712	1.000	0.618	0.186	0.639	0.198	0.308	0.361	0.591	0.482	1.000	0.966	0.920	0.671	0.682	0.341	0.599
	Preservation method	0.342	0.137	0.294	0.060	0.363	0.279	0.165	0.177	0.180	0.340	0.468	0.518	0.447	0.332	0.280	0.166	0.293
	Packing medium	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Sugar content	0.503	0.186	0.146	0.074	0.224	0.181	0.102	0.139	0.203	0.159	0.339	0.318	0.366	0.245	0.402	0.156	0.201
	Fat content	0.403	0.231	0.251	0.139	0.390	0.203	0.052	0.203	0.211	0.211	0.536	0.494	0.606	0.563	0.419	0.157	0.299
	Food production	0.154	0.088	0.038	0.211	0.167	0.037	0.131	0.093	0.232	0.015	0.235	0.086	0.090	0.146	0.407	0.091	0.020

2

Groups	Facets	Iron	Iodine	Magnesium	Sodium	Phosphorous	Potassium	Selenium	Zinc	RAE	Folate	Vitamin B1	Vitamin B2	Vitamin B6	Vitamin B12	Vitamin C	Vitamin D	Vitamin E
	Enriched/fortified	0.463	0.184	0.178	0.095	0.245	0.209	0.107	0.132	0.181	0.151	0.329	0.697	0.391	0.505	0.412	0.174	0.198
	Brandname(yes/no)	0.394	0.228	0.321	0.296	0.305	0.323	0.237	0.124	0.203	0.342	0.288	0.210	0.263	0.317	0.188	0.172	0.554
	Type of milk/liquid used	0.448	0.309	0.545	0.071	0.340	0.155	0.173	0.134	0.409	0.173	0.350	0.338	0.346	0.255	0.703	0.178	0.177
	Flavoured component A	0.744	0.570	0.229	0.151	0.316	0.202	0.216	0.205	0.620	0.170	0.374	0.460	0.467	0.345	0.644	0.307	0.272
	Flavoured component B	0.567	0.472	0.188	0.254	0.306	0.277	0.274	0.161	0.477	0.130	0.334	0.367	0.361	0.275	0.572	0.233	0.275
	Flavoured component C	0.653	0.492	0.172	0.182	0.355	0.183	0.196	0.176	0.583	0.197	0.340	0.326	0.426	0.372	0.609	0.302	0.307

Additional File 4. The population two days' average of energy and nutrient intake distributions before and after facets' deletion at cut-off at 1.00. The population means and percentiles of two days' average of energy, macro- and micronutrient intake distributions before and after facets' deletion and percent difference (% D) for the Dutch population aged 7 to 69 years (DNFCS 2007–2010) (n = 3819).

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
Energy(kcal)																		
P5	1457	1456	0.0	1314	1318	-0.3	1482	1481	0.0	1121	1134	-1.1	1481	1475	0.4	1163	1165	-0.2
P25	1859	1857	0.1	1671	1669	0.1	2048	2042	0.3	1539	1539	0.0	2010	2009	0.1	1565	1567	-0.1
P50	2226	2226	0.0	1915	1920	-0.2	2483	2487	-0.2	1837	1834	0.2	2421	2427	-0.2	1853	1851	0.1
P75	2653	2657	-0.1	2200	2199	0.0	2964	2965	0.0	2223	2225	-0.1	2907	2906	0.0	2220	2225	-0.2
P95	3676	3654	0.6	2694	2694	0.0	3749	3764	-0.4	2834	2830	0.2	3742	3745	-0.1	2813	2806	0.2
Mean	2326	2328	-0.1	1951	1953	-0.1	2534	2539	-0.2	1902	1905	-0.1	2496	2500	-0.2	1911	1914	-0.1
Protein(g)																		
P5	41	41	-0.4	37	37	0.0	55	55	0.0	42	42	0.0	49	49	0.0	41	41	-0.2
P25	56	56	-0.3	50	50	-0.1	74	74	-0.1	57	57	-0.1	70	70	-0.4	55	55	-0.3
P50	69	70	-0.8	59	59	0.2	87	88	-0.2	68	68	0.0	84	84	0.3	66	66	-0.1
P75	86	86	0.2	70	70	-0.2	105	105	0.4	82	82	0.1	102	102	-0.1	80	80	0.0
P95	114	114	0.0	91	91	0.4	133	132	0.5	102	102	0.0	131	130	0.3	101	101	0.2
Mean	72	72	0.0	61	61	0.0	90	90	0.0	70	70	0.1	87	87	0.0	69	68	0.1
Fat(g)																		
P5	45	45	0.1	40	41	-0.9	47	47	-0.1	34	34	0.2	47	47	-0.4	35	35	-0.8
P25	63	63	0.1	56	56	0.1	73	73	-0.3	52	53	-0.4	71	71	0.4	53	53	-0.4
P50	83	83	-0.9	70	70	-0.8	95	95	-0.3	68	69	-0.1	92	93	-0.3	69	69	-0.1
P75	102	103	-0.9	86	87	-1.2	118	119	-1.0	90	90	-0.5	116	117	-1.3	89	89	-0.4

2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	151	151	0.0	112	112	0.0	158	158	-0.3	122	123	-0.7	158	158	-0.1	120	121	-0.9
Mean	87	87	-0.3	72	73	-0.4	98	98	-0.4	73	73	-0.4	96	96	-0.4	73	73	-0.4
SFA(g)																		
P5	16	16	-0.9	14	14	0.2	16	16	-0.1	12	12	-0.1	16	16	-0.1	12	13	-4.1
P25	23	23	0.6	21	21	0.5	27	27	-0.4	20	20	-1.7	26	26	-0.3	20	20	-1.8
P50	30	30	-0.7	27	27	-1.3	34	35	-0.3	26	27	-1.4	33	34	-0.5	26	27	-1.1
P75	39	39	-0.4	33	33	-0.3	44	44	-0.5	35	35	-0.1	43	43	-0.5	34	34	-0.5
P95	54	54	-0.7	44	44	-0.2	60	61	-0.6	49	49	0.6	58	59	-0.8	48	48	-0.1
Mean	32	32	-0.3	27	28	-0.4	36	36	-0.7	28	28	-0.7	35	35	-0.6	28	28	-0.6
MUFA(g)																		
P5	14	14	-0.2	13	13	-1.1	15	15	-0.7	10	10	-4.0	15	15	0.2	11	11	-2.5
P25	22	22	0.0	19	19	0.2	24	24	0.0	17	17	0.1	24	24	0.2	17	18	-0.7
P50	29	28	1.0	25	25	0.0	32	32	-0.9	23	23	-0.2	31	31	-0.3	24	24	-0.5
P75	37	37	-0.4	31	31	0.2	41	41	-0.3	30	30	-0.7	41	41	-0.6	30	31	-0.7
P95	55	55	-0.2	41	41	0.0	57	57	0.2	44	44	0.0	57	57	0.2	43	43	-0.4
Mean	31	31	-0.2	25	25	-0.4	34	34	-0.4	25	25	-0.3	33	33	-0.3	25	25	-0.3
PUFA(g)																		
P5	7	7	0.8	6	6	-2.6	8	8	0.4	6	6	-1.1	8	8	0.2	6	6	-1.4
P25	11	11	0.0	10	10	-1.0	13	13	1.0	9	9	-0.5	13	13	0.2	9	9	-1.0
P50	15	15	-0.2	13	13	0.0	19	19	0.3	13	13	-0.1	18	18	0.0	13	13	0.0
P75	21	21	-0.1	17	17	0.8	25	25	0.0	17	17	-0.3	25	25	-0.9	17	17	0.3

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	34	34	-0.3	24	24	-0.5	37	37	-1.4	27	27	-0.2	36	37	-1.1	26	26	-0.6
Mean	17	17	-0.4	14	14	-0.3	20	20	0.0	14	14	-0.3	19	19	0.0	14	14	-0.3
TFA(g)																		
P5	0.4	0.4	1.7	0.4	0.4	0.0	0.5	0.5	0.1	0.4	0.4	-1.3	0.5	0.5	1.4	0.4	0.4	-0.9
P25	0.8	0.8	0.2	0.7	0.7	0.0	0.9	0.9	-0.1	0.7	0.7	0.1	0.9	0.9	-0.5	0.7	0.7	-0.1
P50	1.1	1.1	-1.4	1.0	1.1	-1.3	1.4	1.4	-0.2	1.1	1.1	-0.2	1.4	1.4	-0.1	1.1	1.1	-0.4
P75	1.7	1.7	-1.0	1.5	1.5	-1.5	2.0	2.1	-3.0	1.7	1.7	-1.2	1.9	2.0	-3.5	1.6	1.6	-0.4
P95	2.9	2.9	-0.2	2.5	2.6	-3.0	3.5	3.5	0.2	2.9	3.0	-0.6	3.3	3.4	-3.8	2.9	2.9	-0.7
Mean	1.3	1.3	-0.4	1.2	1.2	-0.8	1.6	1.6	-0.8	1.3	1.3	-0.3	1.6	1.6	-0.7	1.3	1.3	-0.4
ALA(mg)																		
P5	642	640	0.3	594	607	-2.1	861	831	3.5	607	606	0.2	804	802	0.2	607	606	0.1
P25	1020	1020	0.0	914	923	-1.0	1378	1385	-0.5	1010	1013	-0.3	1295	1299	-0.3	994	1005	-1.1
P50	1380	1380	0.0	1205	1209	-0.3	1906	1899	0.4	1346	1352	-0.5	1806	1811	-0.2	1313	1317	-0.3
P75	1936	1936	0.0	1600	1611	-0.7	2514	2553	-1.6	1841	1865	-1.3	2437	2427	0.4	1793	1801	-0.4
P95	3192	3199	-0.2	2286	2369	-3.7	3762	3714	1.3	2865	2856	0.3	3615	3614	0.0	2790	2773	0.6
Mean	1569	1574	-0.3	1303	1312	-0.7	2054	2057	-0.1	1494	1504	-0.7	1966	1969	-0.2	1460	1470	-0.7
Marine(mg)																		
P5	2	2	-4.2	2	2	0.0	3	3	-4.7	3	3	**10.5**	3	3	-4.7	3	3	4.3
P25	8	8	3.6	7	7	4.2	13	13	-1.0	11	11	-4.1	12	12	-3.4	10	10	-2.4
P50	16	16	0.1	16	16	1.4	28	28	-0.8	25	25	0.9	26	26	0.0	23	23	-0.4
P75	40	40	0.1	37	38	-4.4	118	118	0.0	105	106	-0.6	94	92	2.9	79	78	0.9

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	383	383	0.0	356	322	**9.5**	958	843	**12.0**	960	949	1.1	829	828	0.2	891	844	**5.3**
Mean	77	77	-0.2	70	70	0.7	170	167	1.6	173	172	0.2	153	151	1.4	155	154	0.2
Cholesterol(mg)																		
P5	61	60	1.6	57	58	-1.8	86	86	0.0	65	66	-1.8	77	77	-0.6	62	63	-2.6
P25	105	106	-0.8	90	92	-1.3	137	139	-1.8	112	112	-0.1	130	133	-1.9	107	107	0.0
P50	147	147	0.0	127	129	-1.1	199	200	0.0	161	161	0.0	188	189	-0.3	154	155	-1.0
P75	209	209	0.3	182	183	-0.6	273	276	-1.2	231	232	-0.7	260	263	-1.2	221	223	-1.0
P95	339	339	0.0	298	295	1.1	465	462	0.7	378	381	-0.6	447	447	0.0	366	366	0.0
Mean	167	167	-0.4	146	147	-0.8	222	224	-0.5	182	184	-0.8	212	213	-0.5	176	177	-0.8
Carbohydrates(g)																		
P5	187	188	-0.1	160	161	-0.3	147	147	0.0	117	117	0.4	151	150	0.1	121	121	-0.1
P25	238	237	0.4	213	212	0.5	214	213	0.4	171	170	0.1	218	219	-0.1	176	177	-0.5
P50	287	288	0.0	246	246	0.0	266	265	0.3	208	209	-0.3	270	270	-0.2	216	216	0.0
P75	344	344	0.0	286	286	0.2	330	330	0.0	251	251	0.1	334	334	-0.1	260	260	0.1
P95	454	454	0.0	359	361	-0.4	436	436	0.1	338	337	0.2	439	438	0.2	341	340	0.2
Mean	299	299	0.0	252	251	0.1	277	277	-0.1	215	215	0.0	281	281	-0.1	222	222	0.0
Modisac(g)																		
P5	79	79	0.9	65	65	-0.2	42	43	-0.4	38	38	-1.1	46	45	0.4	41	41	0.3
P25	120	120	0.1	103	102	0.8	82	81	0.8	73	72	0.4	86	86	0.0	76	76	0.3
P50	148	148	0.3	131	130	0.2	114	114	-0.1	96	96	0.0	122	122	0.3	103	102	0.1
P75	185	185	0.3	161	160	1.0	159	159	0.2	129	129	0.2	166	165	0.3	137	137	0.0

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	263	263	0.0	218	218	-0.1	243	243	0.2	193	194	-1.0	246	246	0.0	200	201	-0.6
Mean	156	156	0.2	134	133	0.4	125	125	0.1	104	104	-0.1	131	130	0.1	109	109	0.0
Polysac(g)																		
P5	82	82	-0.4	74	74	0.0	80	80	1.0	61	61	0.0	81	80	0.7	63	63	0.3
P25	113	113	-0.3	95	95	-0.8	118	119	-0.2	88	88	-0.1	117	118	-0.5	90	90	0.3
P50	136	135	0.1	115	115	-0.2	148	148	-0.1	108	109	-0.5	145	145	0.0	110	110	-0.3
P75	166	166	-0.2	137	136	0.4	180	180	-0.1	130	130	-0.1	177	178	-0.3	131	131	0.0
P95	221	222	-0.3	172	173	-0.4	234	237	-1.3	175	175	0.1	233	234	-0.5	175	174	0.7
Mean	142	143	-0.2	118	118	-0.1	152	152	-0.2	112	112	0.0	150	150	-0.2	113	113	0.0
Fibre(g)																		
P5	9.8	9.79	0.0	9.0	9.1	-0.3	11.6	11.6	0.1	9.6	9.6	0.0	10.9	10.9	0.0	9.5	9.5	0.0
P25	13.9	14.2	-2.2	12.6	12.7	-0.8	16.9	16.9	-0.1	14.2	14.1	0.5	16.3	16.3	-0.3	13.9	13.9	-0.2
P50	17.5	17.8	-1.9	15.5	15.4	0.6	21.5	21.6	-0.2	17.6	17.5	0.1	21.0	21.0	-0.1	17.2	17.2	0.0
P75	21.7	21.8	-0.4	18.6	18.8	-0.7	26.4	26.6	-0.8	22.0	22.0	0.1	25.8	26.0	-0.7	21.3	21.3	-0.1
P95	29.9	29.9	0.0	24.6	25.0	-1.5	35.6	35.3	0.8	28.6	28.6	-0.1	35.3	35.1	0.7	28.4	28.3	0.3
Mean	18.5	18.5	-0.4	16.0	16.1	-0.5	22.3	22.3	0.0	18.4	18.4	0.2	21.6	21.6	0.0	18.0	18.0	0.0
Calcium(mg)																		
P5	391	391	-0.1	332	332	0.0	470	469	0.2	429	429	0.0	448	449	-0.2	404	404	0.0
P25	686	690	-0.5	616	627	-1.9	791	792	-0.1	708	710	-0.3	769	769	0.0	695	696	-0.2
P50	922	917	0.5	828	834	-0.7	1092	1091	0.1	944	941	0.3	1049	1055	-0.6	924	924	0.0
P75	1232	1238	-0.5	1072	1080	-0.7	1448	1441	0.4	1195	1199	-0.3	1403	1401	0.1	1176	1179	-0.2

2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d
P95	1832	1865	-1.8	1545	1545	0.0	1999	1999	0.0	1702	1702	0.0	1998	1998	0.0	1677	1681	-0.2
Mean	995	1000	-0.4	874	880	-0.6	1148	1150	-0.1	992	992	0.0	1121	1122	-0.1	971	973	-0.1
Copper(mg)																		
P5	0.6	0.6	0.0	0.6	0.6	-0.2	0.7	0.7	-0.3	0.6	0.6	0.5	0.7	0.7	-1.4	0.6	0.6	-0.4
P25	0.8	0.8	0.3	0.7	0.7	-0.1	1.0	1.0	-1.1	0.8	0.8	0.3	0.9	0.9	-0.2	0.8	0.8	0.1
P50	1.0	1.0	0.5	0.9	0.9	-0.4	1.2	1.2	0.2	1.0	1.0	0.2	1.2	1.2	-0.5	1.0	1.0	0.0
P75	1.3	1.3	-0.2	1.1	1.1	0.0	1.5	1.5	-0.3	1.3	1.3	0.0	1.4	1.4	-0.3	1.2	1.2	-0.3
P95	1.7	1.7	-0.9	1.4	1.4	0.3	2.0	2.0	0.1	1.7	1.7	0.9	2.0	2.0	-0.5	1.7	1.7	0.3
Mean	1.1	1.1	-0.3	0.9	0.9	-0.1	1.3	1.3	-0.3	1.1	1.1	0.4	1.2	1.2	-0.3	1.0	1.0	0.3
Iron(mg)																		
P5	5.1	4.9	3.6	4.7	4.7	0.1	6.5	6.4	0.4	5.6	5.6	0.8	5.9	5.9	-0.8	5.4	5.4	-0.4
P25	7.1	7.1	0.0	6.4	6.4	0.0	9.2	9.2	0.0	7.6	7.6	0.2	8.7	8.7	0.0	7.3	7.3	0.0
P50	8.8	8.8	0.4	7.7	7.7	0.0	11.1	11.1	-0.2	9.2	9.2	0.4	10.8	10.7	0.4	8.9	8.9	0.1
P75	10.9	10.9	0.3	9.3	9.3	-0.4	13.4	13.3	0.3	11.4	11.3	0.2	13.1	13.1	0.1	11.0	11.0	0.0
P95	14.9	14.9	-0.2	12.2	12.2	-0.3	17.6	17.5	0.5	15.2	15.1	1.0	17.4	17.2	0.8	14.9	14.9	0.1
Mean	9.2	9.2	0.1	8.0	8.0	-0.3	11.5	11.5	0.0	9.7	9.7	0.3	11.1	11.1	0.1	9.4	9.4	0.2
Iodine(µg)																		
P5	78	77	1.2	75	74	0.6	99	98	0.6	84	84	0.0	93	94	-0.9	82	82	-0.3
P25	125	125	0.1	109	109	-0.2	152	153	0.0	120	120	0.1	147	146	0.5	118	118	0.3
P50	164	164	0.0	138	138	0.2	192	192	0.0	151	151	-0.1	187	187	0.0	149	149	-0.1
P75	211	211	0.0	170	169	0.8	241	241	0.1	188	188	0.0	235	235	0.0	186	185	0.2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d	Original	Adapted	% d
P95	287	287	0.1	230	230	0.2	322	323	-0.3	249	247	0.6	321	319	0.5	247	245	0.7
Mean	171	171	0.2	142	142	0.3	199	198	0.2	157	157	0.1	194	193	0.2	154	154	0.1
Magnesium(mg)																		
P5	157	157	-0.3	143	141	1.6	219	217	0.5	179	179	0.3	190	191	-0.5	164	164	0.0
P25	212	215	-1.5	194	194	0.3	309	309	0.0	244	244	-0.3	284	284	-0.2	232	232	0.0
P50	267	269	-0.6	236	236	-0.3	370	372	-0.4	295	294	0.1	355	355	-0.1	285	284	0.1
P75	337	338	-0.1	280	280	-0.3	449	449	-0.2	355	354	0.3	436	436	0.0	345	345	0.1
P95	461	461	0.0	366	366	0.0	578	578	0.0	465	464	0.3	566	569	-0.4	457	457	-0.2
Mean	285	286	-0.2	241	242	-0.1	384	384	-0.1	306	306	0.0	366	366	-0.1	295	295	0.0
Sodium(mg)																		
P5	1447	1444	0.2	1273	1270	0.2	1625	1622	0.2	1226	1219	0.6	1533	1559	-1.7	1234	1228	0.5
P25	1951	1953	-0.1	1748	1754	-0.3	2357	2364	-0.3	1773	1789	-0.9	2270	2272	-0.1	1767	1785	-1.0
P50	2473	2492	-0.8	2139	2134	0.2	2970	2986	-0.5	2236	2239	-0.1	2873	2890	-0.6	2225	2224	0.1
P75	3147	3152	-0.2	2567	2568	0.0	3630	3644	-0.4	2812	2825	-0.5	3540	3561	-0.6	2782	2782	0.0
P95	4256	4260	-0.1	3398	3409	-0.3	4792	4830	-0.8	3772	3780	-0.2	4739	4745	-0.1	3668	3674	-0.2
Mean	2611	2619	-0.3	2195	2198	-0.1	3058	3065	-0.2	2336	2346	-0.4	2977	2984	-0.2	2311	2320	-0.4
Phosphorus(mg)																		
P5	757	766	-1.2	680	677	0.5	982	981	0.0	772	770	0.3	928	921	0.7	750	751	-0.1
P25	1038	1033	0.5	940	942	-0.2	1350	1352	-0.2	1065	1063	0.2	1285	1290	-0.4	1029	1030	-0.1
P50	1313	1319	-0.5	1131	1129	0.2	1637	1640	-0.2	1287	1292	-0.4	1584	1590	-0.4	1264	1264	0.0
P75	1652	1656	-0.2	1383	1389	-0.5	2023	2023	0.0	1565	1563	0.1	1966	1966	0.0	1545	1539	0.4

2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	2223	2217	0.3	1782	1782	0.0	2641	2641	0.0	2042	2055	-0.7	2543	2545	-0.1	1981	1981	0.0
Mean	1385	1388	-0.2	1176	1179	-0.3	1707	1711	-0.2	1338	1338	0.0	1649	1652	-0.2	1309	1310	0.0
Potassium(mg)																		
P5	1502	1506	-0.2	1371	1356	1.1	2167	2140	1.3	1789	1798	-0.5	1895	1896	-0.1	1652	1652	0.0
P25	2170	2170	0.0	1939	1933	0.3	3012	3022	-0.3	2362	2372	-0.4	2799	2799	0.0	2279	2273	0.3
P50	2648	2641	0.3	2362	2361	0.0	3594	3610	-0.4	2888	2885	0.1	3421	3426	-0.2	2808	2808	0.0
P75	3271	3272	0.0	2787	2783	0.2	4348	4348	0.0	3541	3538	0.1	4253	4245	0.2	3423	3403	0.6
P95	4268	4288	-0.5	3574	3582	-0.2	5672	5633	0.7	4555	4527	0.6	5483	5489	-0.1	4395	4389	0.1
Mean	2770	2774	-0.1	2400	2401	-0.1	3735	3739	-0.1	2995	2994	0.0	3559	3563	-0.1	2890	2889.444	0.0
Selenium(µg)																		
P5	19	19	-0.2	17	17	0.0	26	26	0.3	21	20	0.6	23	23	0.0	20	20	0.0
P25	27	27	-0.2	24	24	-0.1	37	37	-0.3	29	29	0.1	34	34	0.1	28	28	0.2
P50	34	35	-0.4	30	30	0.1	46	46	0.1	37	37	0.2	44	44	-0.2	35	35	0.0
P75	45	44	0.6	36	36	0.0	58	58	-0.1	47	47	0.9	56	56	-0.1	46	45	0.6
P95	64	64	0.3	51	52	-1.3	93	94	-1.3	69	70	-1.0	87	90	-3.0	67	67	-0.8
Mean	37	37	0.0	32	32	0.0	50	51	-0.4	40	40	0.0	48	48	-0.3	38	38	0.0
Zinc(mg)																		
P5	5.0	5.0	-0.1	4.3	4.2	0.7	6.4	6.4	0.0	5.0	5.1	-1.4	5.7	5.7	0.3	4.8	4.8	0.0
P25	6.8	6.8	0.1	6.2	6.2	-0.2	9.1	9.1	-0.3	7.1	7.0	0.3	8.6	8.6	-0.1	6.9	6.9	-0.2
P50	8.5	8.5	0.4	7.5	7.5	0.1	10.9	10.9	0.0	8.6	8.6	0.0	10.6	10.6	0.1	8.5	8.4	0.5
P75	10.6	10.6	0.0	8.9	8.9	0.3	13.4	13.4	0.1	10.5	10.5	0.2	13.0	13.0	0.0	10.2	10.2	0.0

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	14.7	14.7	-0.1	11.9	11.9	0.0	17.8	17.7	0.5	13.8	14.0	-1.1	17.4	17.4	0.1	13.7	13.8	-0.8
Mean	9.0	9.0	0.0	7.7	7.7	0.1	11.4	11.4	0.1	9.0	9.0	0.3	11.0	11.0	0.1	8.8	8.8	0.2
RAE(μg)																		
P5	180	176	2.1	179	181	-1.3	282	286	-1.4	214	213	0.6	259	259	0.0	203	203	0.0
P25	367	371	-1.3	325	320	1.5	483	485	-0.4	382	384	-0.5	460	463	-0.5	368	371	-0.9
P50	538	544	-1.2	460	459	0.1	708	708	0.0	566	568	-0.5	668	677	-1.5	545	545	0.0
P75	807	843	-4.5	697	699	-0.3	1086	1094	-0.7	855	858	-0.4	1012	1026	-1.4	828	834	-0.8
P95	1703	1703	0.0	1495	1491	0.3	2239	2240	0.0	1958	2010	-2.7	2142	2153	-0.5	1879	1874	0.3
Mean	696	701	-0.8	611	610	0.2	935	938	-0.4	765	753	1.5	891	895	-0.4	738	728	1.3
Folate(μg)																		
P5	90	89	0.8	90	89	0.8	139	143	-2.6	113	115	-1.1	124	124	0.3	106	106	-0.4
P25	144	144	0.1	134	134	0.0	212	213	-0.4	173	173	-0.5	193	196	-1.4	162	162	0.1
P50	188	189	-0.4	170	169	0.3	274	274	0.0	230	229	0.6	260	260	0.0	219	219	0.4
P75	259	257	0.7	223	224	-0.4	368	367	0.4	307	307	0.0	352	350	0.7	294	292	0.7
P95	398	397	0.4	335	339	-1.1	564	563	0.1	465	465	0.0	550	540	1.8	437	437	0.0
Mean	211	211	0.1	186	186	0.0	304	304	0.0	251	250	0.3	287	287	0.0	239	239	0.2
Vitamin B1(mg)																		
P5	0.4	0.4	3.5	0.4	0.4	1.2	0.6	0.6	-0.1	0.4	0.4	-0.2	0.5	0.5	0.0	0.4	0.4	0.1
P25	0.7	0.7	-1.7	0.6	0.6	-0.6	0.8	0.8	0.0	0.7	0.7	-0.7	0.8	0.8	-0.5	0.7	0.7	-1.1
P50	0.9	0.9	-2.8	0.8	0.8	-2.3	1.1	1.1	-2.8	0.9	0.9	-2.1	1.1	1.1	-2.0	0.9	0.9	-2.0
P75	1.3	1.3	-3.0	1.1	1.1	-4.2	1.5	1.5	-1.3	1.2	1.2	-0.8	1.5	1.5	-1.8	1.2	1.2	-1.5

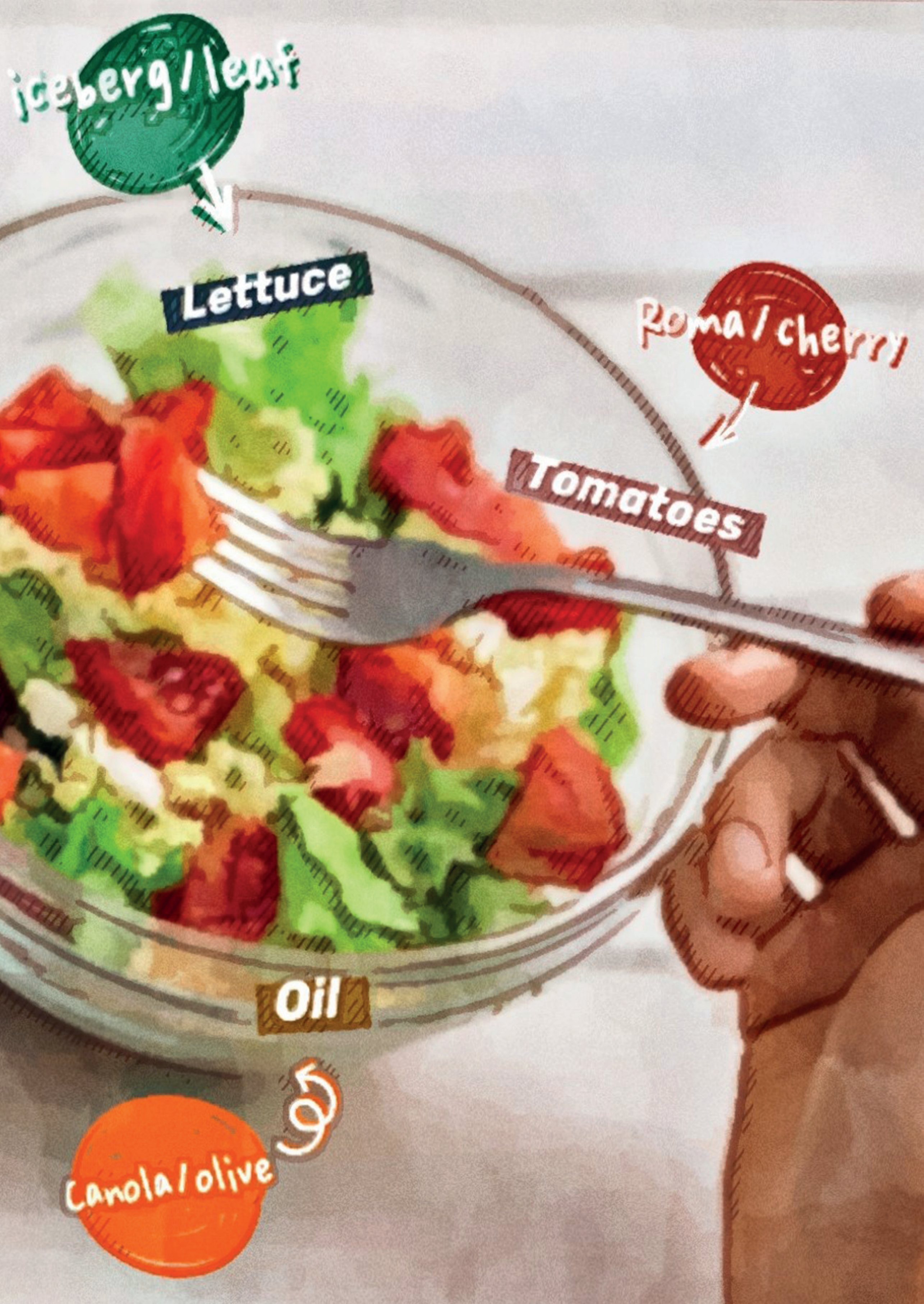
2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	2.0	2.0	-2.1	1.6	1.7	-4.7	2.4	2.4	1.2	1.9	1.9	-1.8	2.3	2.3	-1.3	1.9	1.9	-1.5
Mean	1.0	1.1	-2.9	0.9	0.9	-3.8	1.3	1.3	-0.9	1.0	1.0	-1.2	1.2	1.2	-1.2	1.0	1.0	-1.6
Vitamin B2(mg)																		
P5	0.6	0.6	-0.6	0.5	0.5	-0.6	0.8	0.8	-1.0	0.6	0.6	-2.0	0.7	0.7	-1.0	0.6	0.6	-1.0
P25	1.0	1.1	-4.4	0.9	0.9	-2.6	1.2	1.3	-1.2	1.0	1.0	-0.1	1.2	1.2	-1.5	1.0	1.0	-0.5
P50	1.4	1.5	**-6.9**	1.2	1.3	-4.6	1.6	1.7	-2.1	1.3	1.4	-0.4	1.6	1.6	-1.6	1.3	1.3	-1.6
P75	1.9	2.1	**-9.6**	1.6	1.7	**-6.7**	2.2	2.2	-2.5	1.7	1.7	-1.4	2.1	2.2	-4.5	1.7	1.7	-2.1
P95	2.9	3.3	**-12.8**	2.4	2.8	**-14.8**	3.2	3.4	**-5.8**	2.5	2.6	**-7.2**	3.1	3.4	**-7.4**	2.5	2.7	**-9.5**
Mean	1.5	1.7	**-9.7**	1.3	1.4	**-7.4**	1.8	1.8	-2.6	1.4	1.5	-2.4	1.7	1.8	-3.7	1.4	1.5	-3.2
Vitamin B6(mg)																		
P5	0.7	0.7	-4.6	0.7	0.7	-3.0	0.9	0.9	-0.2	0.7	0.7	0.2	0.9	0.9	-1.5	0.7	0.7	-2.1
P25	1.1	1.2	-3.7	1.0	1.0	-4.1	1.5	1.5	-1.9	1.1	1.1	-1.0	1.4	1.4	-2.2	1.1	1.1	-2.0
P50	1.6	1.6	-4.2	1.3	1.4	**-6.2**	2.0	2.0	-1.5	1.5	1.6	-1.2	1.9	1.9	-2.6	1.5	1.5	-2.8
P75	2.2	2.4	**-10.4**	1.8	1.9	**-5.1**	2.5	2.6	-3.0	2.1	2.1	-1.0	2.5	2.6	-4.4	2.0	2.1	-2.2
P95	3.8	4.3	**-14.1**	3.0	3.4	**-13.5**	4.1	4.3	-2.6	3.4	3.5	-3.1	4.1	4.3	-4.1	3.3	3.5	**-6.8**
Mean	1.8	1.9	**-7.5**	1.5	1.6	**-7.6**	2.2	2.2	-2.5	1.7	1.8	-1.9	2.1	2.2	-3.3	1.7	1.7	-2.8
Vitamin B12(μg)																		
P5	1.2	1.2	-0.6	1.1	1.1	**-5.5**	1.8	1.8	0.7	1.3	1.3	0.3	1.7	1.7	0.4	1.3	1.3	0.5
P25	2.3	2.3	-1.0	2.0	2.1	-1.1	3.2	3.2	-0.5	2.4	2.4	0.2	3.0	3.0	-0.7	2.3	2.3	-0.6
P50	3.3	3.4	-3.2	2.8	2.8	-1.0	4.5	4.6	-0.3	3.4	3.4	0.0	4.3	4.3	-0.9	3.3	3.3	0.2
P75	4.8	4.9	-2.0	4.0	4.0	-0.2	6.3	6.3	0.0	4.9	4.9	-0.4	6.1	6.1	0.1	4.7	4.7	0.2

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	7.6	7.6	0.0	6.2	6.3	-0.8	10.7	10.9	-1.6	9.3	8.9	4.2	10.4	10.6	-1.6	8.6	8.6	0.0
Mean	3.8	3.8	-1.6	3.2	3.2	-1.6	5.2	5.2	-0.3	4.1	4.0	2.0	4.9	4.9	-0.5	4.0	3.9	1.5
Vitamin C(mg)																		
P5	24	24	1.3	23	21	**8.3**	26	27	-1.2	24	24	0.3	26	26	0.0	24	23	2.0
P25	48	46	3.6	47	44	**5.2**	50	50	-0.1	51	51	0.0	50	49	1.2	50	48	2.8
P50	76	72	**5.4**	73	67	**8.4**	84	83	1.1	83	81	2.1	82	80	1.7	81	79	2.2
P75	115	109	4.9	110	106	4.1	129	127	2.0	126	124	1.6	127	123	3.3	122	119	2.3
P95	176	167	**5.2**	169	165	2.6	218	218	0.0	209	206	1.3	210	208	1.0	204	201	1.9
Mean	86	82	4.8	84	79	**5.3**	98	96	1.5	95	94	1.6	96	94	2.1	93	91	2.2
Vitamin D(μg)																		
P5	0.7	0.7	1.8	0.6	0.6	-1.3	1.1	1.1	0.3	0.7	0.7	-3.8	1.0	1.1	-1.3	0.7	0.7	-2.2
P25	1.5	1.5	0.1	1.4	1.4	0.9	2.4	2.4	-0.6	1.7	1.7	0.0	2.2	2.2	0.1	1.7	1.7	-0.1
P50	2.4	2.4	0.4	2.1	2.1	0.1	3.5	3.5	0.5	2.6	2.6	-1.1	3.3	3.3	0.4	2.5	2.5	-0.3
P75	3.6	3.6	0.1	2.9	2.9	-0.7	4.8	4.8	0.6	3.7	3.7	1.2	4.7	4.7	0.0	3.6	3.6	0.3
P95	5.9	5.9	0.0	4.4	4.5	-0.7	7.7	7.6	1.3	6.5	6.4	1.5	7.5	7.5	0.6	6.1	6.0	2.5
Mean	2.8	2.7	0.3	2.2	2.2	-0.4	3.9	3.8	1.0	3.0	3.0	0.3	3.7	3.6	0.9	2.9	2.9	0.2
Vitamin E(mg)																		
P5	5.3	5.4	-2.8	4.9	4.9	-0.5	6.0	5.8	3.0	4.9	4.8	1.5	5.8	5.7	0.9	4.9	4.9	0.7
P25	8.5	8.7	-2.3	7.5	7.7	-2.8	10.2	10.2	0.5	7.9	8.0	-0.1	10.0	9.9	0.5	7.8	7.9	-1.3
P50	11.6	12.0	-3.6	9.8	10.0	-2.1	13.8	13.7	0.4	10.8	10.8	0.1	13.4	13.4	-0.3	10.6	10.6	0.0
P75	16.0	16.1	-0.6	13.2	12.9	2.2	18.5	18.4	0.4	14.4	14.3	0.3	18.1	18.1	0.1	14.1	14.1	0.4

	7-18 years						19-69 years						All ages					
	Male			Female			Male			Female			Male			Female		
Nutrients	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**	Original	Adapted	**% d**
P95	25.0	24.9	0.4	19.1	19.5	-2.0	27.9	27.5	1.3	22.2	21.5	3.0	27.1	27.1	0.1	21.1	20.9	1.0
Mean	13.0	13.2	-2.0	10.6	10.8	-1.7	15.0	15.0	0.3	11.8	11.7	0.9	14.6	14.7	-0.1	11.6	11.5	0.5

Percent difference larger than 5% are shown in bold.


iceberg/leaf
Lettuce
Roma/cherry
Tomatoes
Oil
Canola/olive

Chapter 3

How Does a Simplified Recipe Collection Procedure in Dietary Assessment Tools Affect the Food Group and Nutrient Intake Distributions of the Population

Liangzi Zhang, Hendriek Boshuizen, Marga Ocké

British Journal of Nutrition 2020, 124(2), 189-198

Abstract

Technology advancements have driven the use of self-administered dietary assessment methods in large-scale dietary surveys. Interviewer-assisted methods generally have a complicated recipe recording procedure enabling the adjustment from a standard recipe. In order to decide if this functionality can be omitted for self-administered dietary assessment, this study aimed to assess the extent of standard recipe modifications in the Dutch National Food Consumption Survey, and measure the impact on the food group and nutrient intake distributions of the population when the modifications were disregarded. A two-scenario simulation analysis was conducted. Firstly, the individual recipe scenario omitted the full modifications to the standard recipes made by people who knew their recipes. Secondly, the modified recipe scenario omitted the modifications made by those who partially modified the standard recipe due to their limited knowledge. The weighted percentage differences for the nutrient and food group intake distributions between the scenarios and the original dataset were calculated. The highest percentage of energy consumed through mixed dishes was 10% for females aged 19 to 79. Comparing the combined scenario and the original dataset, the average of the absolute percentage difference for the population mean intakes was 1.6% across all food groups and 0.6% for nutrients. The soup group (-6.6%) and docosahexaenoic acid (DHA) (-2.3%) showed the largest percentage difference. The recipe simplification caused a slight underestimation of the consumed amount of both foods (-0.2%) and nutrients (-0.4%). These results are promising for developing self-administered 24hR or food diary applications without complex recipe function.

Introduction

Inappropriate dietary intakes have been recognized as major risk factors for developing chronic diseases[1,2]. Many countries, therefore, carry out national food consumption surveys to monitor food consumption and nutrient intakes of their populations[3]. The most frequently used dietary assessment methods in Europe for collecting national food consumption data are 24-hour-recalls (24hRs) and food records[4], both open methods aim to assess the intake of all foods and drinks on a specific day(s). 24hRs require low literacy levels of participants and are less likely to alter eating behaviours than food records[5,6], whereas food records have less recalling bias[7]. To collect harmonised data among the EU Member states, the European Food Safety Authority recommended collecting two non-consecutive 24hRs for adults and two non-consecutive food records for children. Moreover, the use of validated and standardized software was advised, for example, GloboDiet (formerly known as Epic-Soft)[8,9,10]. The EFSA guidelines were based on the experiences and recommendations from various European projects, such as the EFCOSUM-project[11], the EFCOVAL project[12], the PANCAKE project[13] and the PAN-EU project[14].

Although detailed food consumption information can be captured, the current interviewer-administered dietary assessment method induces high costs and logistic complications for data collection and handling[15,16]. This limitation encourages efforts to explore solutions that could enhance the cost-efficiency of implementing large-scale nutrition monitoring surveys[17]. The increased access to the Internet has fostered the development of many self-administered dietary assessment methods, including web-based and smartphone-based tools[18]. The overall quality of collected data from these tools is comparable with the interviewer-administered method[19]. Participants have greater flexibility and fewer time constraints to complete the survey[17]. Costs could be greatly reduced with automated coding and less interviewer involvement. Moreover, the incorporation of more objective food recognition features (e.g., photographs, barcodes) could enhance efficiency and reducing unintentional under-reporting in recording real-time food intake[20,21,22,23]. Review studies have indicated great potential for mobile dietary assessment applications to be used in large-scale studies[20,24,25]. Hence, moving towards self-administered tools from interviewer-administered tools seems a promising effort to explore for future national food consumption surveys[26]. However, the complexity of self-reporting tools is a real concern for certain people to participate and complete the survey[17]. A simplification of certain comprehensive features might be a crucial step in facilitating migrations from interviewer-administered tool to a self-administered tool.

The feature of recording mixed meal intake comprises complicated procedures in GloboDiet. Mixed recipes are collected through a specific recipe pathway[27] , which starts by automatically searching entered recipes within a pre-existing standard recipe list[9,28]. The standard recipe is entered into the system unless the participants know that the actual recipe they consumed has different ingredient than the standard recipe. In this case, ingredients in standard recipes can be replaced, and the amounts of ingredients can be adjusted[15,29]. Different from portion size estimation of reported single food items which are always estimated "as consumed", for mixed recipes, more steps are needed to estimate the amount of each ingredient. After the portion size of the consumed mixed dish has been estimated, the ingredient amounts in the whole prepared recipe can be reported as raw or as consumed. With only raw amounts known, a consumed amount is calculated using pre-defined algorithms and standard food-specific coefficients (e.g., raw-to-cooked yield factors, density, or edible part coefficients)[9,10]. This additional ingredient adjustment is complicated to implement and requires much work and knowledge from the participants. Besides, estimating ingredient amounts in a mixed meal is without question a difficult task, given that people already find it hard to estimate portions in a single food item[14]. The common practice for current self-administered tools is to choose standard mixed dishes directly or to create new recipes from scratch[6,30]. Although omitting modifications to the standard recipes can save much effort, it could potentially bias the actual ingredient intake. Hence, the impact of using standard recipes without modifications on the nutrient and food group intake at the population level should be investigated.

This study aims to provide evidence to support the decision on whether a standard recipe modification feature in self-administered 24hRs or food diary apps is needed for large-scale dietary surveys. Firstly, we evaluated how often a home-prepared mixed meal is consumed in the Dutch diet and how often alterations were being made to standard recipes. Subsequently, we did a simulation analysis using national survey data in which standard recipes were adjusted by the interviewers and assessed the impact of ignoring these changes but using the standard ingredients. We then compared the observed food group and nutrient intake distributions of the population between the original and simulated data.

Methods

Data Collection

In this study, the importance of recipes in the Dutch diet was analysed and a simulation study was conducted using the data of the Dutch National Food Consumption Survey 2012-2016[31]. This survey was conducted among 4313 Dutch men and women aged 1-79 years old. Subjects were excluded if they were pregnant, lactating, or institutionalized. Participants completed a questionnaire covering various background factors, such as educational level, working status, native country, family composition, various lifestyle factors, such as patterns of physical activity, smoking, use of alcoholic beverages and various general characteristics of the diet. Dietary intake of participants was collected through two 24hRs on non-consecutive days with 2-6 weeks in between. The 24hRs for children between 1 to 15 years old and older adults between 70 to 79 years old were collected by face-to-face interviews by trained dieticians with a food diary completed one day before the interview as an aid. For children aged 1 to 8 years, their parents or caretakers were interviewed. The 24hRs for 16 to 70 years olds were conducted through two telephone interviews. In both the face-to-face and the telephone-based 24hR interviews, a computer-assisted software called GloboDiet developed by the International Agency for Research on Cancer (IARC) was used[8].

Current Recipe Collection

The feature within GloboDiet that could record mixed meal intakes was called the recipe pathway. As a starting point, a standard recipe list with 378 pre-defined recipes embedded in the recipe pathway was used if a pre-defined recipe resembled the mixed dish reported by the participants. Then, participants were asked whether the recipe was commercial or homemade. Commercial recipes were those with brand names from commercial sources such as supermarkets and restaurants. For home-prepared dishes, different procedures were followed depending on the participant's knowledge of their dishes. For those who were aware of the detailed information, an individual recipe was created by going through several steps to modify the standard ingredients according to their situations. For people not knowing much about their dishes, standard recipes were applied instead. For situations that ingredients were visually recognized in the mixed dish, ingredients in standard recipes were substituted, this type of recipes was regarded as a modified recipe. For ingredients that were reported as raw, raw-to-cooked yield factors and edible part coefficients were multiplied with the raw amount to calculate the consumed amount. A complete flow chart explaining the recipe pathway can be found in Figure 1. All reported food items, including the recipe ingredients, were linked to the most appropriate food code in the Dutch National Food Composition Database (NEVO table 2016/5.0)[32] by trained dieticians. Each food item/ingredient were categorized according to the GloboDiet food group classification system[33].

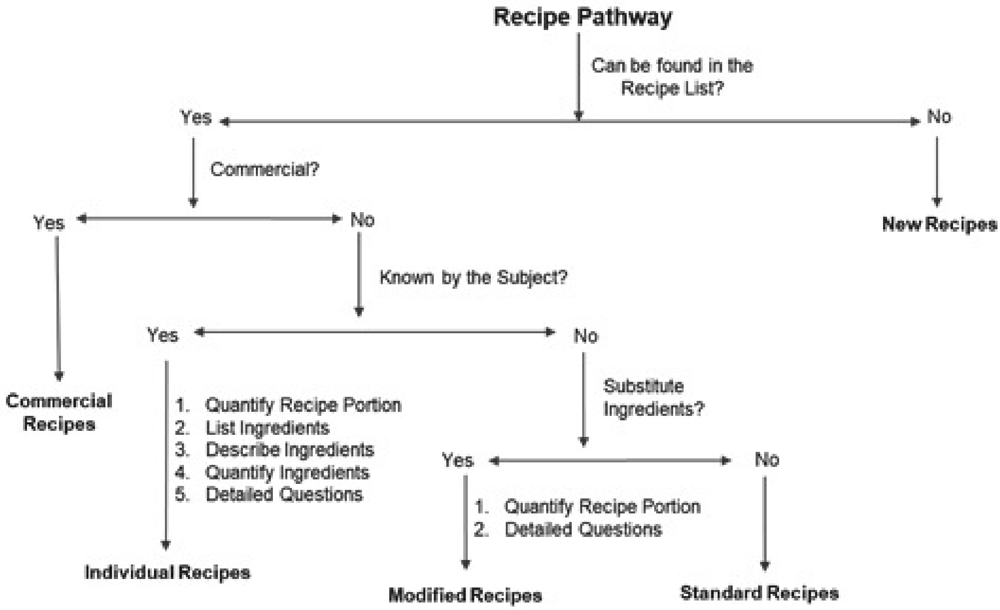


Figure 1. The flow chart of the mixed meal pathway in GloboDiet. Dishes were defined as homemade dishes if they could be found in the pre-defined recipe list and were not derived from commercial sources. Individual recipes were defined when people knew the information, they could substitute the predefined ingredients or adjust the amount of the ingredients of a standard recipe. For those who did not know the recipe, standard recipes would be used instead. For situations where the participants partly knew the recipe, adjustments of the ingredients were possible. These were regarded as modified recipes. New recipes were created if the name of the dish could not be found in the pre-defined recipe list.

Simulation Procedure

A two-scenario simulation study was conducted to evaluate whether the distributions of population nutrient and food group intake changed significantly when only standard recipes were used. The individual recipe scenario only ignored modifications to standard recipes for people who knew the recipes. In other words, the ingredients of individual recipes were switched to ingredients of standard recipes. The modified recipe scenario only ignored modifications to standard recipes during or after the interview for people who did not know all details of the recipe (but they could see some ingredients or had some insight in the used ingredients but not amounts). In both scenarios, the portion consumed for each recipe was kept the same with the original individual or modified recipe. The amount of ingredients were calculated according to the predefined percentage of the recipe total weight. All the ingredients were linked to the food code in the NEVO automatically if the same food item was linked already in the original database, otherwise they were linked by dieticians. The individual recipe scenario and the modified recipe scenario were also taken together in a combined scenario. Scenario analyses were run with all participants including those that did not use recipes, and in the subset of participants that did consume either mixed recipes that were reported as individual recipes or modified recipes. The details of preparing commercial recipes were not known by the participants, and newly created recipes were created from scratch without having a corresponding standard recipe to compare with. Hence, the ingredients were kept unchanged for recipes that were originally commercial, for unmodified standard recipes and for new recipes.

Data Analysis

The following study population characteristics were summarized. The highest educational level of the participants or the parents/carers of participants under the age of 19, who is the main earner of the family. Educational level was categorized into low (primary education, lower vocational education, advanced elementary education), middle (intermediate vocational education, higher secondary education), and high (higher vocational education and university). Percentages of energy and macronutrient intake consumed through recipes from the individual's total intake were calculated for the total population and per age and sex category. Percentage of energy intake consumed through recipes per eating occasion, recipe types, recipe groups was calculated. All population means were weighted for socio-demographic characteristics, day of the week and season of data collection, to give results that are representative for the Dutch population and representative for all days of the week and all seasons.

The nutrient level and quantities of food groups consumed were summarized per person by day and averaged over two days in both the dataset with original ingredients and the one with ingredients from standard recipes. The weighted mean, median, 5th, 25th, 75th, 95th percentile and the percentage differences of consumption per nutrient and food group between the original and the new dataset were calculated for the total population and within people who used individual and modified recipes in each scenario. The nutrient intake estimation was conducted for two scenarios, both separately and combined. The number of food items in each food group was also compared between the original state and the combined scenario. The descriptive summary and population nutrient intake distributions were conducted using the SAS 9.4, the replacement of ingredients from standard recipes to original dataset were conducted using R x64 3.5.0. The percentage differences between the original and newly linked dataset were calculated using Excel 2016 software.

Results

The general characteristics of the survey participants are shown in Table 1. The study included equal percentages for each age-gender group. The average BMIs for boys (18.0 kg/m^2) and males (26.0 kg/m^2) were similar with those for girls (18.1 kg/m^2) and females (26.6 kg/m^2), respectively. More than half of the boys and girls had a highly educated head of the household (54%). More adult males (38%) had a higher education level than females

(28%). The mean intake of energy per day was generally higher in boys (1988 kcal) and males (2543 kcal) than in girls (1685 kcal) and females (1860 kcal). The percentages of energy consumed through mixed dishes were lower or equal to 10% for the four age-gender groups; adult female (10%) consumed more energy through mixed dishes than other age-gender groups.

Figure 2 illustrates the percentage of energy consumed through mixed dishes differentiated by eating occasions, by recipe types (new, individual, modified, standard) and by recipe groups based on the food group of the main ingredients. Dinner was the main occasion for consuming mixed dishes (73.2%). More than half of the people who consumed mixed dishes knew the content of the recipe and reported individual recipes (62.9%). The modified recipes (15.1%) were reported as the second most frequent recipe type. Among all the recipe groups, energy from cereal- (52.5%) and vegetable- (22.6) based mixed dishes were higher than other recipe groups.

Table 1. General characteristics of the population aged 1-79 years old from the Dutch National Food Consumption Survey 2012-2016, weighted for socio-demographic characteristics and season, and day of the week.

		Total	**1-18 years old**		**19-79 years old**	
Gender (n)		4313	Boys (1122)	Girls (1113)	Males (1043)	Females (1035)
Education n(%)	Low	815 (19)	108 (9)	105 (9)	242 (23)	360 (35)
	Middle	1628 (38)	413 (37)	408 (37)	406 (39)	383 (37)
	High	1888 (44)	601 (54)	600 (54)	395 (38)	292 (28)
Mean BMI kg/m² (SD)			18.0 (3.1)	18.1 (3.4)	26.0 (4.6)	26.6 (5.6)
Mean Energy intake in kcal per day (SD)			1988 (21)	1685 (16)	2543 (27)	1860 (19)
Mean % kcal from home-made recipes (SD)			8 (0.32)	8 (0.34)	9 (0.38)	10 (0.53)

SD, standard deviation.

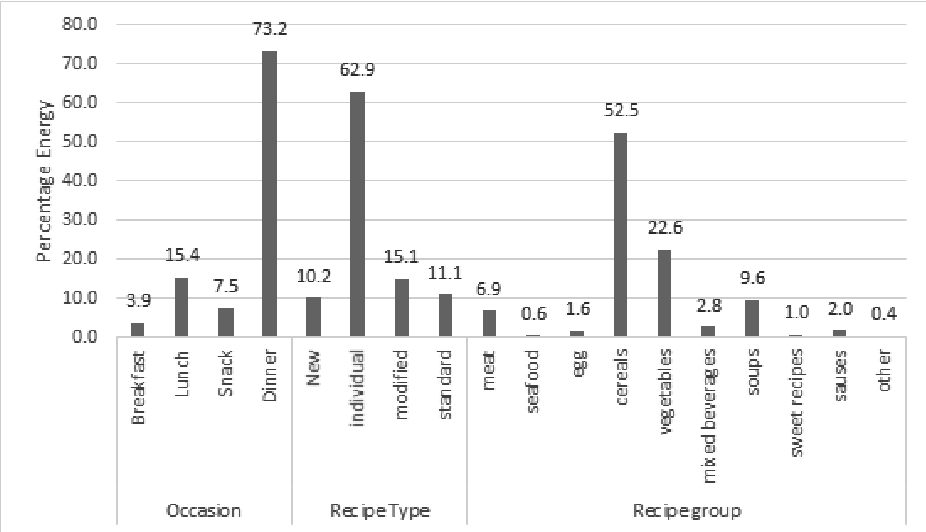


Figure 2. Energy consumed through mixed dishes partitioned (%) by different occasions, recipe types and recipe groups from the Dutch National Food Consumption Survey 2012-2016.

Stratified by food groups, the impact of the combined scenario on the consumed amount of ingredients at a population level are shown in Table 2. In the individual recipe scenario, we disregarded modifications made by people who knew their recipes, while in the modified recipe scenario, the substitutions made by people who did not know the exact recipes were disregarded. Detailed results for sub-food groups can be found in Appendix 1. From Table 2, the average of the percentage difference in mean intakes over all food groups was -0.2%, while the average of the absolute percentage difference was 1.6%. For eight out of 17 food groups, the percentage difference in mean consumed amount was larger than 1% or lower than -1% between the combined scenario and the original dataset. Among the food groups that were overestimated by the standard recipes, meat has the highest percentage difference (3.6%). Specifically, ingredients from the meat group were overestimated the most by the standard recipes of hamburgers and meat wraps. Potatoes (1.2%) and legumes (0.7%) also showed an overestimation of the consumed amount but an underestimation in the count of the food ingredients by the standard recipes. Another observation was that the standard recipes tended to be less specific for certain food groups. For example, there were more unclassified meat products in standard recipes than in individual recipes (Appendix 1). A similar finding was also observed in the fats group.

For the food groups with an underestimated consumed amount by the standard recipes, soups and stocks had been underestimated to the greatest extent in average intake (-6.6%). The underestimation was mainly due to the existence of water in standard recipes of soups that were made from soup powders, whereas stock from the soup group was reported in individual and modified recipes. Similarly, the total amount of vegetables was underestimated by the standard recipes, especially in spaghetti bolognese, greek salad, chicken-related dishes (e.g., wrap, curry, siam) and in different kinds of soups. On the contrary, there was a higher occurrence of different vegetables in standard ingredients. When we looked at the detailed results of food subgroups (Appendix 1), fruiting vegetables, cabbages, mushrooms, and stalk vegetables were the main contributors to the contradictory result. In other words, these subgroups were used more often in standard recipes but in small amounts.

As for the results of the nutrient analysis, Table 3 shows the percentage difference and the difference of the actual amount of 26 nutrients between the combined scenario and the original dataset within the total population. The average of the percentage difference was 0.6% for the absolute mean intakes across all nutrients. The averages for the other five percentiles of the intake distributions were slightly higher; the 25th percentile has the highest average of 1.0%. The percentage difference in mean of five nutrients was larger than 1.0% or lower than -1.0%. Most nutrient intakes (73%) were underestimated by using standard recipes, with an average percentage difference of -0.4% for the population mean intakes. The largest negative mean percentage difference was in DHA (-2.3%) with an actual amount difference of -2.6mg, while the largest positive mean percentage difference was in vitamin B1 (1.8%) with an actual amount difference of 0.02mg. A relatively larger percentage difference with a low actual amount difference was also observed in trans fatty acids (-1.1%, -0.01g). To compare the impact to the total population with only those who consumed mixed dishes, seven nutrients that have higher percentage differences than the other 19 nutrients from the combined scenario are included in Figure 3a. The impact within people who consumed mixed dishes was larger than the impact on the total population for every nutrient. When we looked at Appendix 2b that has the percentage, and actual amount difference for all nutrients, the effect within people consumed mixed dishes has more nutrients with a mean percentage difference larger than 1.0% or lower than -1.0% than within total population.

The separate effects of each scenario on the nutrient intake of the total population is shown in Figure 3b. Either scenario has a smaller impact than the combined effect as shown in Figure 3a. The individual recipe scenario has a larger impact on the nutrient intake distribution than the modified recipe scenario. The results with all nutrients for each scenario separately is shown in Appendix 3a & 3b. The individual recipe scenario has an average of the absolute mean percentage difference of 0.5% with five nutrients larger than 1.0% or lower than -1.0%. While the modified recipe scenario has an average of the absolute mean percentage difference of 0.2% with all nutrients fell within –1.0% to 1.0%. About 63% of the nutrients were underestimated in scenario 1, while 88% of the nutrients were underestimated in scenario 2. Figure 3a and 3b also illustrate that the intake of most nutrients was underestimated by using standard recipes. Exceptions were vitamin B1 and ALA.

Vitamin B1 was overestimated in all scenarios. ALA showed contradictory results between the two scenarios and was higher in combined scenarios than the original dataset.

Table 2. The percentage and amount difference of the food group intake distribution of the population between the combined scenario and the original data.

Food Groups	Percentage Difference (%)				Amount Difference (g)				Difference in the number of ingredient occurrence
	Mean	Median	P75	P95	Mean	Median	P75	P95	
Potatoes and other tubers	1.2	2.5	0.0	0.9	0.8	1.5	0.0	1.9	-31
Vegetables	-4.0	-6.4	-4.1	-3.8	-5.3	-7.2	-7.3	-12.0	1454
Legumes	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-18
Fruits, nuts and seeds, olives	-0.6	-0.8	-1.1	0.0	-0.7	-0.8	-2.1	-0.1	50
Dairy products and substitutes	0.1	0.6	0.0	0.4	0.4	1.9	0.0	3.4	254
Cereals and cereal products	1.6	1.4	1.4	2.3	3.1	2.5	3.5	8.5	163
Meat, meat products and substitutes	3.6	3.8	2.8	1.7	3.5	3.3	3.7	3.9	49
Fish, shellfish and amphibians	-3.0	0.0	0.0	0.0	-0.5	0.0	0.0	0.0	-42
Eggs and egg products	2.6	0.0	0.0	6.2	0.3	0.0	0.0	3.1	88
Fats and oils	2.4	3.1	1.9	-0.2	0.5	0.6	0.6	-0.1	662
Sugar and confectionery	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	-68
Cakes and sweet biscuits	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-4
Non-alcoholic beverages	0.2	0.2	0.0	-0.5	3.6	2.9	0.5	-14.9	416
Condiments, spices, sauces and yeast	-0.5	-1.7	0.0	0.2	-0.2	-0.4	0.0	0.2	32
Soups and stocks	-6.6	0.0	-10.9	-4.3	-2.8	0.0	-6.8	-9.9	-460
Miscellaneous	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-9
Savoury snacks	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-11
Average (|Percentage Difference|)	1.6	1.2	1.3	1.2	-	-	-	-	-
Average (Percentage Difference)	-0.2	0.2	-0.6	0.2	-	-	-	-	-

P75, 75th percentile. P90, 95th percentile. | Percentage Difference |: the absolute value of percentage difference.

3

Table 3. The percentage and amount difference of the nutrient intake distribution of the population between the combined scenario and the original data.

	Percentage Difference (%)						Amount Difference		
Nutrients	**Mean**	**P5**	**P25**	**Median**	**P75**	**P95**	**Mean**	**P5**	**P95**
Energy (kcal)	0.2	1.4	0.5	-0.3	-0.4	-0.1	4	16	-4
Protein (g)	0.0	0.1	0.3	0.2	0.6	1.4	0.0	0.0	1.8
Carbohydrates (g)	0.6	0.6	0.4	0.7	0.3	0.6	1.3	0.7	2.4
Mono- and disaccharides (g)	-0.1	0.1	-0.1	-0.2	-0.2	-0.3	-0.1	0.0	-0.6
Fibre (g)	-0.8	-0.2	-1.2	-0.6	-0.5	0.1	-0.2	0.0	0.0
Fat (g)	-0.2	0.8	0.0	-0.4	-1.1	0.0	-0.1	0.3	0.0
SFA (g)	-0.5	-0.4	0.5	-0.7	0.2	-0.8	-0.1	0.0	-0.4
ALA (g)	0.2	4.2	1.0	-0.1	1.0	-1.6	0.00	0.02	-0.06
TFA (g)	-1.1	-2.5	-0.6	-1.2	-1.2	0.3	0.0	0.0	0.0
DHA (mg)	-2.3	0.0	-9.4	-10.1	-2.6	-2.1	-2.63	0.00	-14.51
Calcium (mg)	-0.1	1.1	0.5	0.0	0.3	-1.3	-1	4	-23
Iron (mg)	-0.8	0.4	-0.6	-0.9	-1.1	-0.5	-0.1	0.0	-0.1
Sodium (mg)	0.4	-1.2	0.3	-0.1	-0.6	1.3	9	-13	54
Potassium (mg)	-0.5	0.3	0.0	-1.1	0.0	0.2	-16	4	8
Zinc (mg)	-0.2	-0.3	-0.2	0.1	-1.0	-1.2	-0.02	-0.01	-0.21
Beta-carotene (µg)	-1.3	2.4	-0.3	0.0	1.6	-2.8	-27	5	-207
Retinol (µg)	0.2	2.1	0.5	0.1	0.3	-0.2	1	3	-4
Folate equivalents(µg)	-0.9	-0.1	-1.0	-1.2	-1.0	-0.2	-2.1	-0.1	-0.8
Vitamin B1 (mg)	1.8	1.0	2.1	0.8	1.7	3.6	0.02	0.00	0.07
Vitamin B2 (mg)	-0.2	0.6	-1.1	0.0	-0.3	-0.6	0.00	0.00	-0.02
Vitamin B3 (mg)	-0.4	-0.8	-0.9	-1.3	-1.2	0.2	-0.1	-0.1	0.1
Vitamin B6 (mg)	-0.5	-0.2	-0.3	-1.4	-0.5	0.7	-0.008	-0.002	0.021
Vitamin B12 (µg)	-0.5	-1.9	-0.5	-0.4	-1.0	0.0	-0.02	-0.03	0.00
Vitamin C (mg)	-1.8	-0.1	-1.8	-1.8	-1.8	-2.6	-2	0	-5
Vitamin D (µg)	-0.2	0.0	0.1	0.0	-0.4	0.6	0.0	0.0	0.0
Vitamin E (µg)	-0.6	-1.1	-0.6	-0.7	-0.6	-0.1	-0.1	-0.1	0.0
Average (|Percentage Difference|)	0.6	0.9	1.0	0.9	0.8	0.9	-	-	-
Average (Percentage Difference)	-0.4	0.2	-0.5	-0.8	-0.4	-0.2	-	-	-

P5, 5th percentile. P25, 25th percentile. P75, 75th percentile. P90, 95th percentile. SFA, saturated fatty acids. ALA, alpha-Linolenic acids. TFA, trans-fatty acids. DHA, docosahexaenoic acids. | Percentage Difference |: the absolute value of percentage difference.

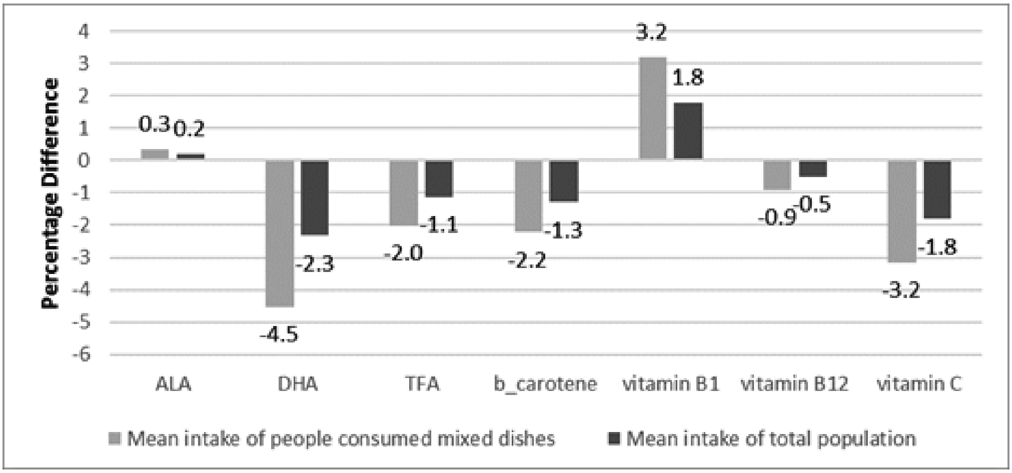


Figure 3a. The percentage difference of the mean intake of 7 nutrients of the total population and within people who consumed mixed dishes between the combined scenario and the original dataset.

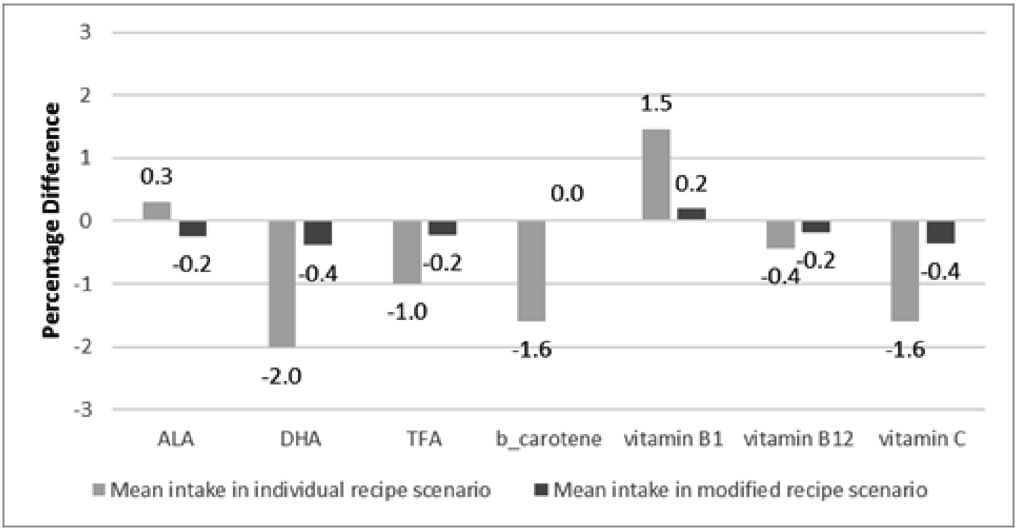


Figure 3b. The percentage difference of the mean intake of 7 nutrients of the total population between each scenario and the original dataset.

Discussion

A replacement of complete recipe recording steps with a simplified recipe recording procedure would help improve the cost-effectiveness of recording mixed meal intake and was explored to be used in the Dutch National Food Consumption Surveys (DNFCS). Therefore the impact of replacing individual with standard recipes was investigated using data collected in DNFCS 2012-2016. With a few exceptions, this study found that using only pre-defined standard recipes caused less than one percent differences in mean nutrient intakes and food consumption compared to standard recipes being modified according to participant declaration. The main contributing factor for the insignificant impact was the small portion of the energy consumed (approximately 10%) from home-made mixed meals, according to DNFCS 2012-2016. This observation is in line with the trend of preparing less mixed dishes at home due to peoples' tendency to eating quick and ready meals(34). Also, compared to countries where mixed dishes were dominant(35), the western diet includes relatively few dishes that mix all ingredients(36). An additional explaining factor was that 20% of the home-made mixed meals were entered as new recipes or unmodified standard recipes, both of which could not be simplified in this study.

Despite the small overall difference in main food groups, a larger difference was found in some subgroups of the main food group. The reason is that the standard recipes contained more ingredients from undefined food subgroups while individual recipes contained more ingredients from specific food subgroups. A seemingly contradictory outcome was found in several food groups where the average consumed amount was lower, while the number of food items was higher in standard recipes, the vegetable group is a notable example of this. One possible explanation might be that the participants deemed vegetables as healthy foods hence overestimated the consumed amount in individual recipes (37). Another reason is that the standard recipes in our study were purposely created with more varieties of vegetables in smaller portion size of each type in order to make them representative for different versions of a recipe (lasagne with mushrooms, or with leek, or with carrots).

The change in the ingredients would inevitably cause a change in nutrient intake(38,39). The overall difference was small across nutrients with only a few exceptions. DHA has the largest average percentage difference and was underestimated when replacing individual recipes with standard recipes (-2.3%), which was mainly due to the fact that people put fish in dishes that do not have fish in the corresponding standard recipes (e.g., oven dishes, salads, foreign dishes). On the contrary, vitamin B1 has the largest positive average percentage difference of 1.8%, which was probably due to the higher average amount of dairy, cereals, and meat in standard recipes. These differences seem unsubstantial for dietary monitoring purposes with a large sample size. However, to better accommodate real-life variations, the development of future standard recipes should consider the fact that people tend to take fewer varieties from certain food groups (e.g., vegetables) but higher amounts of available varieties in certain dishes. The specificity of food subgroups should be defined in standard recipes with ingredients from, for example, the meat group. Also, acknowledge that people might exclude or replace the main ingredients of certain dishes with ingredients from other food groups. Without the modification functionality, identical standard recipes with different main ingredient options should be listed individually, with key ingredients shown in the recipe title for easier identification. A study comparing nutrition results from more varieties of unmodifiable standard recipes with results from original modifiable standard recipes could provide more relevant insight.

As far as we know, this is the first study investigating the impact of replacing individual with standard recipes. The study contained a large sample size (n=4313), the population was representative of the Dutch population, and the survey results were representative for all days of the week and all seasons. The study results are transferable to surveys which use Globodiet as their main instrument of collecting dietary data; however, it may not apply to countries where mixed dishes are dominant in the diet. Unlike many other large food consumption surveys that allocate a composite dish into one food group(40,41), surveys that use Globodiet disaggregate ingredients of recipes and distinguish the food group of every ingredient(42). The disaggregation simplifies the procedure of replacing old ingredients with standard ingredients and calculating nutrient and food group difference between the original and new scenarios. Another advantage of the study is that the between-person variation did not impact the results since the manipulated dataset was derived from the original dataset, and thus on data from the same participants(37).

There are also some limitations to the study. Firstly, some of the complex foods were not considered as recipes in Globodiet(9), such as cakes, biscuits, desserts, sauces, and some snacks. As a result, the percentage of the home-prepared mixed meal might have been underestimated as well as the impact on intake. However, the influence is estimated to be small due to a high proportion of eating industrially prepared food and out-of-home eating for sweets, especially for northern European countries such as the Netherlands(39,43,44). Secondly, only the impact on food groups and nutrients were considered, while other aspects related to food can also be important. For example, since standard recipes contain mostly generic food items, this would underestimate the consumption of branded or specific food items, and hence their environmental impact as well as exposure to potentially harmful substances of the population. Lastly, the quality, completeness, and specificity of the standard recipe database is also an essential aspect in estimating the actual intake of the population. In our study, the standard recipe list was derived from a widely-used cookbook in the Netherlands, the deviation of standard recipes from the real-life intake is unknown.

As opposed to creating a new individual recipe from scratch, good quality standard recipes could save time, supplement commonly forgotten ingredients such as seasonings(7,35), and correct misreporting out of embarrassment and inconvenience[45]. Hence, standard recipes were embedded in most of the dietary apps and software, as well as dietary assessment surveys in many countries[39,46]. While numerous commercial and research-based apps have the option of creating new individual recipes[47], there are no self-administered methods incorporated modifiable standard recipes as far as we know[48]. The reason for the less popularity of modifiable standard recipes in self-administered software is that incorporating recipe modification would increase the time and effort for the participants and part of the respondents might not provide valuable answers due to their limited knowledge about the recipe. Also, when applying technologies like photo recognition and analysis in smartphones[45,49,50], challenges exist especially for mixed dishes where not all ingredients are visible[51].

According to the study results and current limitation on technology, a recipe function that could balance the workload of participants and capture deviation with real-life intakes is proposed. In self-reported food diaries or 24hRs, participants could choose well-described unchangeable standard recipes if they are representative for the real preparation habits of the population. For participants that have consumed a mixed dish that cannot be classified as one of the available recipes, an individual recipe could be created. In this way, the number of participants that are requested to provide recipe details is limited. Such an approach needs to be evaluated in terms of usability for the users, and in terms of the validity of the consumption data.

Conclusion: Disregarding modification steps of a recipe functionality in 24hR software has a small impact on the distribution of food group consumption and nutrient intake of the Dutch population. Therefore, there seems to be minor loss in validity for food group and nutrient intake if no recipe function is available and mixed dishes are treated as food (with standard ingredients). Using good quality standard recipes without modification is a promising solution for reducing participant burden on self-administered 24hR or food diary.

References:

1. Arens-Volland AG, Spassova L, Bohn T (2015) Promising approaches of computer-supported dietary assessment and management-Current research status and available applications. Int J Med Inform 84, 997-1008.
2. Mozaffarian D, Wilson PWF, Kannel WB (2008) Beyond established and novel risk factors - Lifestyle risk factors for cardiovascular disease. Circulation 117, 3031-3038.
3. Huybrechts I, Aglago EK, Mullee A et al. (2017) Global comparison of national individual food consumption surveys as a basis for health research and integration in national health surveillance programmes. Proc Nutr Soc 76, 549-567.
4. Rippin H, Hutchinson J, Evans C et al. (2018) National nutrition surveys in Europe: a review on the current status in the 53 countries of the WHO European region. Food Nutr Res 62.
5. Ahluwalia N, Dwyer J, Terry A et al. (2016) Update on NHANES Dietary Data: Focus on Collection, Release, Analytical Considerations, and Uses to Inform Public Policy. Adv Nutr 7, 121-134.
6. Jacques S, Lemieux S, Lamarche B et al. (2016) Development of a Web-Based 24-h Dietary Recall for a French-Canadian Population. Nutrients 8.
7. Shim J-S, Oh K, Kim HC (2014) Dietary assessment methods in epidemiologic studies. Epidemiol Health 36, e2014009-e2014009.
8. Slimani N, Casagrande C, Nicolas G et al. (2011) The standardized computerized 24-h dietary recall method EPIC-Soft adapted for pan-European dietary monitoring. Eur J Clin Nutr 65 Suppl 1, S5-15.
9. Slimani N, Deharveng G, Charrondiere RU et al. (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. Comput Methods Programs Biomed 58, 251-266.
10. EFSA (2014) Guidance on the EU Menu methodology. In EFSA Journal, vol. 12, pp. 3944: European Food Safety Authority.
11. Brussaard JH, Lowik MR, Steingrimsdottir L et al. (2002) A European food consumption survey method--conclusions and recommendations. Eur J Clin Nutr 56 Suppl 2, S89-94.
12. de Boer EJ, Slimani N, van 't Veer P et al. (2011) The European Food Consumption Validation Project: conclusions and recommendations. Eur J Clin Nutr 65, S102-S107.
13. Ocke M, Brants H, Dofkova M et al. (2015) Feasibility of dietary assessment methods, other tools and procedures for a pan-European food consumption survey among infants, toddlers and children. Eur J Nutr 54, 721-732.
14. Ambrus Á, Horváth Z, Farkas Z et al. (2013) Pilot study in the view of a Pan-European dietary survey – adolescents, adults and elderly. EFSA Supporting Publications 10, 508E-n/a.
15. Thompson FE, Subar AF (2017) Dietary assessment methodology. In Nutrition in the Prevention and Treatment of Disease, pp. 5-48: Elsevier.

16. Ocké MC, Slimani N, Brants H et al. (2011) Potential and requirements for a standardized pan-European food consumption survey using the EPIC-Soft software. Eur J Clin Nutr 65, S48.
17. Touvier M, Kesse-Guyot E, Mejean C et al. (2011) Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. Br J Nutr 105, 1055-1064.
18. Taren D, Dwyer J, Freedman L et al. (2002) Dietary assessment methods: where do we go from here? Public Health Nutr 5.
19. Kirkpatrick SI, Subar AF, Douglass D et al. (2014) Performance of the Automated Self-Administered 24-hour Recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. Am J Clin Nutr 100.
20. Ambrosini GL, Hurworth M, Giglia R et al. (2018) Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr J 17, 5.
21. Illner AK, Freisling H, Boeing H et al. (2012) Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 41, 1187-1203.
22. Maringer M, van't Veer P, Klepacz N et al. (2018) User-documented food consumption data from publicly available apps: an analysis of opportunities and challenges for nutrition research. Nutr J 17, 59.
23. Cade JE (2017) Measuring Diet in the 21st Century: Use of New Technologies. Proc Nutr Soc 76, 276-282.
24. Eldridge AL, Piernas C, Illner AK et al. (2018) Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients 11.
25. Sharp DB, Allman-Farinelli M (2014) Feasibility and validity of mobile phones to assess dietary intake. Nutrition 30, 1257-1266.
26. Subar AF, Kirkpatrick SI, Mittl B et al. (2012) The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. J Acad Nutr Diet 112, 1134-1137.
27. Cosgrove M, Flynn A, Kiely M (2005) Impact of disaggregation of composite foods on estimates of intakes of meat and meat products in Irish adults. Public Health Nutr 8, 327-337.
28. Raper N, Perloff B, Ingwersen L et al. (2004) An overview of USDA's Dietary Intake Data System. J Food Compos Anal 17, 545-555.
29. Slimani N, Ferrari P, Ocke M et al. (2000) Standardization of the 24-hour diet recall calibration method used in the european prospective investigation into cancer and nutrition (EPIC): general concepts and preliminary results. Eur J Clin Nutr 54, 900-917.
30. Leclercq C, Arcella D, Piccinelli R et al. (2009) The Italian National Food Consumption Survey INRAN-SCAI 2005-06: main results in terms of food consumption. Public Health Nutr 12, 2504-2532.

31. Van Rossum CT, Nelis K, Wilson C et al. (2018) National dietary survey in 2012-2016 on the general population aged 1-79 years in the Netherlands. EFSA Supporting Publications 15.
32. RIVM (2011) NEVO-table. Series NEVO-table. https://nevo-online.rivm.nl/ (accessed July 2018).
33. Van Rossum C, Buurma-Rethans E, Vennemann F et al. (2016) The diet of the Dutch: Results of the first two years of the Dutch National Food Consumption Survey 2012-2016.
34. Geurts M, Bakel AMv, Van Rossum CT et al. (2016) Food consumption in the Netherlands and its determinants. Bilthoven, The Netherlands: National Institute for Public Health and the Environment.
35. Yun SH, Choi B-Y, Kim M-K (2009) The effect of seasoning on the distribution of nutrient intakes by a food-frequency questionnaire in a rural area. Korean Journal of Nutrition 42, 246-255.
36. Wang O, Gellynck X, Verbeke W (2016) Perceptions of Chinese traditional food and European food among Chinese consumers. British Food Journal 118, 2855-2872.
37. Macdiarmid J, Blundell J (1998) Assessing dietary intake: Who, what and why of under-reporting. Nutr Res Rev 11, 231-253.
38. Fitt E, Mak TN, Stephen AM et al. (2010) Disaggregating composite food codes in the UK National Diet and Nutrition Survey food composition databank. Eur J Clin Nutr 64, S32.
39. Fitt E, Prynne CJ, Teucher B et al. (2009) National Diet and Nutrition Survey: Assigning mixed dishes to food groups in the nutrient databank. J Food Compos Anal 22, S52-S56.
40. Sui Z, Raubenheimer D, Rangan A (2017) Consumption patterns of meat, poultry, and fish after disaggregation of mixed dishes: secondary analysis of the Australian National Nutrition and Physical Activity Survey 2011–12. BMC Nutrition 3, 52.
41. Agudo A (2005) Measuring intake of fruit and vegetables. Kobe, Japan: World Health Organization.
42. Roe M, Bell S, Oseredczuk M et al. (2013) Updated food composition database for nutrient intake. EFSA Supporting Publications 10, 355E.
43. Slimani N, Deharveng G, Southgate DA et al. (2009) Contribution of highly industrially processed foods to the nutrient intakes and patterns of middle-aged populations in the European Prospective Investigation into Cancer and Nutrition study. Eur J Clin Nutr 63 Suppl 4, S206-225.
44. Orfanos P, Naska A, Trichopoulos D et al. (2007) Eating out of home and its correlates in 10 European countries. The European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr 10, 1515-1525.
45. Macdiarmid JI, Blundell JE (1997) Dietary under-reporting: what people say about recording their food intake. Eur J Clin Nutr 51, 199-200.

46. Shinozaki N, Murakami K, Masayasu S et al. (2019) Development and simulated validation of a dish composition database for estimating food group and nutrient intakes in Japan. Public Health Nutr, 1-14.
47. Zhang L, Nawijn E, Boshuizen H et al. (2019) Evaluation of the Recipe Function in Popular Dietary Smartphone Applications, with Emphasize on Features Relevant for Nutrition Assessment in Large-Scale Studies. Nutrients 11.
48. Institute TNC (2018) Instructions for the ASA24 Researcher Website: ASA24-2018 (U.S. version only), The National Cancer Institute.
49. Bucher T, Weltert M, Rollo ME et al. (2017) The international food unit: a new measurement aid that can improve portion size estimation. Int J Behav Nutr Phys Act 14, 124.
50. Yang Y, Jia W, Bucher T et al. (2019) Image-based food portion size estimation using a smartphone without a fiducial marker. Public Health Nutr 22, 1180-1192.
51. Naska A, Lagiou A, Lagiou P (2017) Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Res 6, 926.

Appendix 1: Stratified by food sub-groups, the impact of the combined scenario on the consumed amount of ingredients.

		Percentage Difference		Amount Difference		
Food Groups	Food sub-groups	Mean	P95	Mean	P95	Occurrence
Potatoes and other tubers	Unclassified, mixed and other tubers	-24.2	0.0	-0.1	0.0	-13
	Potatoes	1.3	0.9	1.0	1.9	-18
Vegetables	Unclassified, mixed salad/vegetables	-18.4	-13.7	-2.2	-11.0	-266
	Leafy vegetables	-2.9	-2.5	-0.6	-2.4	-169
	Fruiting vegetables	-5.7	-1.9	-2.8	-3.2	71
	Root vegetables	4.7	-1.9	0.6	-1.6	635
	Cabbages	-2.5	-4.7	-0.5	-5.3	209
	Mushrooms	-12.3	-14.5	-0.4	-2.8	137
	Grain and pod vegetables	1.1	0.8	0.0	0.1	-21
	Leek, onion, garlic	5.6	-1.5	0.7	-0.7	668
	Stalk vegetables, sprouts	-7.3	-26.3	-0.2	-2.2	190
Legumes	Legumes	0.7	0.0	0.0	0.0	-18
Fruits, nuts and seeds, olives	Fruits	-0.5	0.4	-0.5	1.4	136
	Fruit compote	-1.7	0.0	-0.1	0.0	-2
	Nuts, peanuts, seeds	-2.1	0.0	-0.1	0.0	-67
	Peanut butter, nut/seeds spread	0.7	0.0	0.0	0.0	-4
	Olives	-3.7	-100.0	0.0	-1.4	-13
Dairy products and substitutes	Non fermented milk and milk beverages	0.4	0.2	0.5	0.8	187
	Fermented milk, milk beverages and yogurt	-0.2	0.0	-0.1	0.0	-3
	Milk substitutes	-3.0	0.0	-0.3	0.0	-19
	Yoghurt	1.7	0.0	0.9	0.0	20
	Fromage blanc, petits suisses	-0.2	0.0	0.0	0.0	-11
	Cheeses (including spread cheeses)	-0.9	-1.0	-0.3	-0.9	70
	Unclassified creams	0.0	0.0	0.0	0.0	-16
	Dairy creams and creamers	-1.0	-2.3	0.0	-0.4	60
	Non dairy creams and creamers	-67.9	0.0	-0.4	0.0	-32
	Ice cream	-0.1	0.0	0.0	0.0	-2
Cereals and cereal products	Flours, starches, flakes, semolina	5.0	4.9	0.1	0.7	69
	Pasta, rice, other grain	4.0	6.0	1.9	10.5	-7
	Bread	0.8	1.2	1.0	3.1	118
	Crispbread, rusks	-0.2	0.0	0.0	0.0	1
	Breakfast cereals	-0.2	0.0	0.0	0.0	-10
	Dough and pastry	1.8	1.6	0.1	0.9	-8
Meat, meat products and substitutes	Unclassified and combined meat	-6.9	0.0	-0.2	0.0	-23
	Unclassified, mixed and other mammals	58.2	26.1	2.2	8.0	355
	Beef	-9.2	-2.7	-1.1	-1.9	-307
	Veal	-8.9	0.0	0.0	0.0	-5
	Pork	12.1	0.0	1.6	0.0	177
	Mutton/lamb	36.7	0.0	0.2	0.0	27
	Chicken, hen	5.7	12.0	0.9	9.2	-208
	Turkey, young turkey	-17.2	0.0	-0.1	0.0	-8
	Game	-10.2	0.0	0.0	0.0	-1

		Percentage Difference		Amount Difference		
Food Groups	Food sub-groups	Mean	P95	Mean	P95	Occurrence
	Hot processed meat	1.0	3.2	0.3	3.2	36
	Cold processed meat	-0.1	-1.5	0.0	-1.0	42
	Hot meat substitutes	-16.5	0.0	-0.2	0.0	-35
	Cold meat substitutes	-0.6	0.0	0.0	0.0	-1
Fish, shellfish and amphibians	Unclassified and combined fish products	-30.4	0.0	0.0	0.0	-1
	Fish	-3.4	0.0	-0.4	0.0	-40
	Crustaceans, molluscs	-4.1	0.0	-0.1	0.0	2
	Fish products	-0.9	- 100.0	0.0	-0.4	-3
Eggs and egg products	Eggs	2.6	6.2	0.3	3.1	88
Fats and oils	Unclassified and combined fats	78.1	38.7	1.1	3.2	1758
	Vegetable oils	-3.7	-2.8	-0.1	-0.4	-403
	Butter	-2.9	-2.3	-0.1	-0.3	-101
	Margarines and cooking fats	-2.8	-2.9	-0.4	-1.2	-589
	Other animal fats (including fish oils)	-2.2	0.0	0.0	0.0	-3
Sugar and confectionery	Sugar	-0.5	0.0	0.0	0.0	-54
	Jam, jelly, marmelade	0.0	0.0	0.0	0.0	-2
	Honey	-1.2	-11.1	0.0	-0.7	-14
	Other sweet spread	0.1	0.0	0.0	0.0	-4
	Syrup	1.1	0.0	0.0	0.0	-1
	Unclassified and other chocolate	-0.1	0.0	0.0	0.0	-2
	Chocolate spread and chocolate powder	-0.1	0.0	0.0	0.0	-2
	Confectionery non chocolate	-0.1	0.0	0.0	0.0	11
Cakes and sweet biscuits	Cakes, pies, pastries, puddings	0.0	0.0	0.0	0.0	-4
Non alcoholic beverages	Unclassified and combined non alc. Drinks	-1.1	0.0	-0.1	0.0	-31
	Fruit and vegetable juices	2.2	0.0	1.2	0.0	160
	Carbonated/soft/isotonic drinks	0.0	0.0	0.0	0.0	-17
	Waters	0.4	0.0	0.0	0.0	304
Condiments, spices, sauces and yeast	Unclassified or combined condiments	-2.8	0.0	0.0	0.0	-1
	Other and mixed sauces	-6.7	-4.9	-0.9	-2.9	-133
	Tomato sauces	3.8	0.6	0.3	0.2	29
	Dressing sauces, mayonnaises and similar	-2.1	-0.5	-0.2	-0.2	-139
	Mayonnaise based spreads	-0.6	0.0	0.0	0.0	-8
	Spices, herbs and flavourings	-47.1	0.0	0.0	0.0	-21
	Unclassified and combined condiments	36.7	14.9	0.7	1.6	339
	Vinegar	-21.1	-48.4	0.0	-0.4	-34
Soups and stocks	Soups	-42.9	-25.1	-8.0	-35.7	-481
	Stocks	21.6	11.4	5.2	16.2	21
Miscellaneous	Artificial sweeteners	0.0	0.0	0.0	0.0	-7
	Meal substitutes	-0.1	0.0	0.0	0.0	-2
Savoury snacks	Savoury snacks, biscuits and crisps	-0.2	0.0	0.0	0.0	-10
	Savoury filled buns, croissants	0.0	0.0	0.0	0.0	-1

Appendix 2a: The percentage and actual difference of the nutrient intake distribution of the total population between the combined scenario and the original data.

	Percentage Difference				Amount Difference			
Nutrients	**Mean**	**P5**	**P50**	**P95**	**Mean**	**P5**	**P50**	**P95**
Energy (kcal)	0.2	1.4	-0.3	-0.1	3.82	16.17	-5.44	-4.11
Protein (g)	0.0	0.1	0.2	1.4	0.00	0.05	0.15	1.75
Carbonhydrates (g)	0.6	0.6	0.7	0.6	1.31	0.74	1.55	2.45
Mono- and disaccharides (g)	-0.1	0.1	-0.2	-0.3	-0.13	0.03	-0.17	-0.65
Fibre (g)	-0.8	-0.2	-0.6	0.1	-0.15	-0.02	-0.12	0.03
Fat (g)	-0.1	-0.6	0.3	-0.8	-0.01	-0.02	0.04	-0.20
SFA (g)	-0.5	-0.4	-0.7	-0.8	-0.14	-0.04	-0.20	-0.41
TFA (g)	-1.1	-2.5	-1.2	0.3	-0.01	-0.01	-0.01	0.01
ALA (g)	0.2	4.2	-0.1	-1.6	0.00	0.02	0.00	-0.06
DHA (mg)	-2.3	0.0	-10.1	-2.1	-2.63	0.00	-0.77	-14.51
Calcium (mg)	-0.1	1.1	0.0	-1.3	-0.60	4.35	0.00	-23.18
Iron (mg)	-0.8	0.4	-0.9	-0.5	-0.08	0.02	-0.09	-0.08
Sodium (mg)	0.4	-1.2	-0.1	1.3	9.11	-12.82	-2.99	54.34
Potassium (mg)	-0.5	0.3	-1.1	0.2	-16.48	4.17	-33.81	7.90
Zinc (mg)	-0.2	-0.3	0.1	-1.2	-0.02	-0.01	0.01	-0.21
Beta carotene (µg)	-1.3	2.4	0.0	-2.8	-26.67	4.94	-0.43	-207.32
Retinol (µg)	0.2	2.1	0.1	-0.2	1.15	2.75	0.60	-3.62
Folate equivalents(µg)	-0.9	-0.1	-1.2	-0.2	-2.09	-0.15	-2.84	-0.80
Vitamin B1 (mg)	1.8	1.0	0.8	3.6	0.02	0.00	0.01	0.07
Vitamin B2 (mg)	-0.2	0.6	0.0	-0.6	0.00	0.00	0.00	-0.02
Vitamin B3 (mg)	-0.4	-0.8	-1.3	0.2	-0.08	-0.06	-0.21	0.06
Vitamin B6 (mg)	-0.5	-0.2	-1.4	0.7	-0.01	0.00	-0.02	0.02
Vitamin B12 (µg)	-0.5	-1.9	-0.4	0.0	-0.02	-0.03	-0.01	0.00
Vitamin C (mg)	-1.8	-0.1	-1.8	-2.6	-1.66	-0.01	-1.37	-5.47
Vitamin D (µg)	-0.2	0.0	0.0	0.6	-0.01	0.00	0.00	0.04
Vitamin E (µg)	-0.6	-1.1	-0.7	-0.1	-0.07	-0.05	-0.08	-0.02

Appendix 2b: The percentage and actual difference of the nutrient intake distribution of the recipe population between the combined scenario and the original data.

	Percentage Difference				Amount Difference			
Nutrients	Mean	P5	P50	P95	Mean	P5	P50	P95
Energy (kcal)	0.3	2.1	-0.2	0.7	6.76	24.14	3.77	24.61
Protein (g)	0.0	0.3	0.0	1.2	0.00	0.13	0.03	1.59
Carbonhydrates (g)	1.0	0.6	1.2	3.1	2.32	0.72	2.72	11.78
Mono- and disaccharides (g)	-0.2	0.1	-0.1	0.0	0.22	0.03	0.09	0.09
Fibre (g)	-1.3	-1.1	-0.9	-0.9	0.27	0.11	0.17	0.30
Fat (g)	-0.3	0.6	-0.2	1.1	0.23	0.22	0.17	1.60
SFA (g)	-0.8	0.1	-1.2	-2.4	0.24	0.01	0.34	1.32
TFA (g)	-2.0	-5.2	-2.4	0.6	0.02	0.01	0.02	0.01
ALA (g)	0.3	4.7	-0.6	-2.1	0.01	0.03	0.01	0.07
DHA (mg)	-4.5	0.0	-12.8	-8.4	4.66	0.00	0.93	52.71
Calcium (mg)	-0.1	1.9	0.1	-1.5	1.05	7.77	1.34	26.35
Iron (mg)	-1.3	0.7	-1.5	-0.7	0.13	0.03	0.14	0.12
Sodium (mg)	0.7	-1.6	0.1	3.6	16.13	18.19	2.68	150.66
Potassium (mg)	-0.9	0.2	-2.0	-0.7	29.18	3.30	61.16	35.97
Zinc (mg)	-0.3	-0.7	0.0	-1.6	0.03	0.04	0.00	0.27
Beta carotene (µg)	-2.2	2.3	0.6	-5.2	47.23	5.48	7.60	391.76
Retinol (µg)	0.3	2.7	0.2	-0.6	2.03	3.64	0.81	10.60
Folate equivalents(µg)	-1.5	0.6	-2.2	-1.1	3.70	0.62	5.00	4.59
Vitamin B1 (mg)	3.2	1.3	1.6	10.0	0.03	0.01	0.01	0.19
Vitamin B2 (mg)	-0.4	0.4	0.5	-0.9	0.01	0.00	0.01	0.02
Vitamin B3 (mg)	-0.8	-1.1	-2.3	0.6	0.15	0.08	0.37	0.22
Vitamin B6 (mg)	-0.9	-0.3	-1.6	1.5	0.01	0.00	0.02	0.04
Vitamin B12 (µg)	-0.9	-2.7	-0.6	-1.7	0.04	0.04	0.02	0.15
Vitamin C (mg)	-3.2	-3.9	-2.9	-3.2	2.93	0.96	2.21	6.68
Vitamin D (µg)	-0.3	3.1	0.3	0.4	0.01	0.02	0.01	0.03
Vitamin E (µg)	-1.1	-1.0	-2.5	0.1	0.13	0.05	0.27	0.03

Appendix 3a: The percentage and actual difference of the nutrient intake distribution of the total population between the individual scenario and the original data.

	Percentage Difference				Amount Difference			
Nutrients	**Mean**	**P5**	**P50**	**P95**	**Mean**	**P5**	**P50**	**P95**
Energy (kcal)	0.2	1.2	-0.4	0.6	4.62	14.51	7.24	22.04
Protein (g)	0.1	-0.1	0.2	1.5	0.08	0.03	0.15	1.91
Carbonhydrates (g)	0.5	0.9	0.7	0.6	1.24	1.11	1.55	2.45
Mono- and disaccharides (g)	-0.1	0.0	-0.1	-0.3	0.11	0.00	0.15	0.65
Fibre (g)	-0.7	-0.2	-0.5	0.1	0.14	0.02	0.10	0.03
Fat (g)	-0.1	0.4	-0.2	-0.1	0.05	0.16	0.14	0.09
SFA (g)	-0.4	-0.7	-0.5	-1.2	0.12	0.08	0.15	0.66
TFA (g)	-1.0	-1.8	-1.2	0.1	0.01	0.00	0.01	0.00
ALA (g)	0.3	5.2	-0.1	-0.4	0.01	0.03	0.00	0.01
DHA (mg)	-2.0	0.0	-7.4	-2.5	2.28	0.00	0.56	16.56
Calcium (mg)	-0.1	1.1	0.3	-1.2	0.56	4.30	2.86	20.44
Iron (mg)	-0.6	0.4	-0.9	-0.3	0.06	0.02	0.09	0.06
Sodium (mg)	0.3	-1.0	0.0	1.2	7.76	10.72	0.25	47.90
Potassium (mg)	-0.4	0.3	-0.8	0.0	13.99	4.20	25.64	0.00
Zinc (mg)	-0.1	-0.3	0.0	-1.2	0.01	0.01	0.00	0.21
Beta carotene (µg)	-1.6	1.8	-0.4	-5.2	32.98	3.74	5.09	382.04
Retinol (µg)	0.2	0.0	0.2	-0.2	1.08	0.00	0.63	3.62
Folate equivalents(µg)	-0.7	0.0	-1.2	-0.5	1.75	0.05	2.84	2.06
Vitamin B1 (mg)	1.5	0.7	0.7	2.7	0.02	0.00	0.01	0.06
Vitamin B2 (mg)	-0.2	0.6	0.0	-0.6	0.00	0.00	0.00	0.02
Vitamin B3 (mg)	-0.2	-0.7	-0.7	0.1	0.04	0.05	0.12	0.03
Vitamin B6 (mg)	-0.3	0.0	-0.9	1.1	0.01	0.00	0.01	0.03
Vitamin B12 (µg)	-0.4	-1.2	-0.6	-0.4	0.02	0.02	0.02	0.03
Vitamin C (mg)	-1.6	1.0	-1.9	-3.1	1.46	0.23	1.46	6.31
Vitamin D (µg)	-0.2	-0.1	0.1	0.0	0.00	0.00	0.00	0.00
Vitamin E (µg)	-0.4	-0.7	-0.6	-0.1	0.05	0.03	0.07	0.02

3

Appendix 3b: The percentage and actual difference of the nutrient intake distribution of the total population between the modified scenario and the original data.

	Percentage Difference				Amount Difference			
Nutrients	**Mean**	**P5**	**P50**	**P95**	**Mean**	**P5**	**P50**	**P95**
Energy (kcal)	-0.1	-0.5	0.0	-0.3	3.06	5.30	0.00	11.02
Protein (g)	-0.2	0.0	-0.4	-0.5	0.19	0.02	0.28	0.69
Carbonhydrates (g)	-0.1	0.0	0.0	0.0	0.14	0.00	0.00	0.00
Mono- and disaccharides (g)	0.0	0.1	-0.1	-0.3	0.05	0.03	0.15	0.65
Fibre (g)	-0.2	-0.2	-0.2	0.0	0.04	0.02	0.03	0.00
Fat (g)	-0.2	-0.1	-0.1	-0.1	0.19	0.05	0.04	0.09
SFA (g)	-0.1	0.6	-0.1	0.0	0.05	0.08	0.03	0.00
TFA (g)	-0.2	-1.1	-0.4	0.1	0.00	0.00	0.00	0.00
ALA (g)	-0.2	0.1	-0.1	-1.2	0.00	0.00	0.00	0.04
DHA (mg)	-0.4	0.0	-2.7	0.0	0.44	0.00	0.21	0.00
Calcium (mg)	-0.1	-0.3	-0.2	-0.6	0.84	1.13	1.71	9.92
Iron (mg)	-0.3	-0.3	-0.1	-0.3	0.03	0.02	0.01	0.06
Sodium (mg)	-0.1	-0.6	-0.2	0.0	3.01	7.10	4.36	1.06
Potassium (mg)	-0.2	0.0	-0.4	0.0	6.04	0.63	11.25	0.00
Zinc (mg)	-0.2	0.0	-0.2	-1.0	0.02	0.00	0.02	0.16
Beta carotene (µg)	0.0	-1.0	0.2	0.5	0.51	2.07	2.29	35.07
Retinol (µg)	0.0	1.0	-0.1	-0.2	0.32	1.28	0.50	3.62
Folate equivalents(µg)	-0.3	-0.7	-0.5	0.4	0.61	0.77	1.06	1.90
Vitamin B1 (mg)	0.2	-0.2	0.0	0.1	0.00	0.00	0.00	0.00
Vitamin B2 (mg)	-0.1	-0.1	-0.1	-0.1	0.00	0.00	0.00	0.00
Vitamin B3 (mg)	-0.3	0.0	-0.9	0.0	0.06	0.00	0.14	0.01
Vitamin B6 (mg)	-0.3	0.4	-0.8	0.0	0.00	0.00	0.01	0.00
Vitamin B12 (µg)	-0.2	-1.0	-0.3	0.0	0.01	0.01	0.01	0.00
Vitamin C (mg)	-0.4	-2.5	-0.1	-0.1	0.33	0.58	0.09	0.25
Vitamin D (µg)	-0.1	0.0	-0.1	0.3	0.00	0.00	0.00	0.02
Vitamin E (µg)	-0.3	-0.4	-0.2	-0.1	0.04	0.02	0.02	0.03


22:55
Food Record Apps
MyFitnessPal
FatSecret
YAZIO
Lifesum
Lose It!
MyNetDiary
MyPlate
Nutracheck
Food

Chapter 4

Evaluation of the Recipe Function in Popular Dietary Smartphone Applications, with Emphasize on Features Relevant for Nutrition Assessment in Large-Scale Studies

Liangzi Zhang, Eline Nawijn, Hendriek Boshuizen, Marga Ocké

Nutrients 2019, 11(1), 200

Abstract

Nutrient estimations from mixed dishes require detailed information collection and should account for nutrient loss during cooking. This study aims to make an inventory of recipe creating features in popular food diary apps from a research perspective and to evaluate their nutrient calculation. A total of 12 out of 57 screened popular dietary assessment apps included a recipe function and were scored based on a pre-defined criteria list. Energy and nutrient content of three recipes calculated by the apps were compared with a reference procedure, which takes nutrient retention due to cooking into account. The quality of the recipe function varies across selected apps with a mean score of 3.0 (out of 5). More relevant differences (larger than 5% of the Daily Reference Intake) between apps and the reference were observed in micronutrients (49%) than in energy and macronutrients (20%). The primary source of these differences lies in the variation in food composition databases underlying each app. Applying retention factors decreased the micronutrient contents from 0% for calcium in all recipes to more than 45% for vitamins B6, B12, and folate in one recipe. Overall, recipe features and their ability to capture true nutrient intake are limited in current apps.

1. Introduction

When assessing the dietary intake of a large population, an accurate dietary assessment plays a fundamental role [1]. Self-report dietary assessment methods, such as 24-hour dietary recall (24HDR), dietary record (DR), and food frequency questionnaire (FFQ), are commonly used to assess food consumption at both individual and population level [2]. Since underreporting, overreporting, misreporting, and interviewer bias can occur in those methods **[3-5]**, assessing dietary intake with a high level of accuracy continues to be a major challenge in nutritional epidemiology and monitoring [6,7]. Moreover, cumbersome procedures of collecting details of foods are time-consuming and are associated with a high burden for both the respondent and the researcher [8]. This is especially the case for 24HDR and DR, which are open methods, and for which repeated measurements are needed to estimate usual dietary intake [9]. The burden laid on respondents can also lead to a low response rate, which may lead to bias in the survey results and diminish the representativeness of the sample [10].

Progress in Information and Communication Technology (ICT) in the past few decades has led to investigations into innovative strategies to overcome drawbacks of traditional pen-and-paper and interviewer-based dietary assessment methods [11,12]. One such innovative strategy is the use of mobile applications (apps) on smartphones for a dietary record. In the last decade, an increase in the number of smartphone users has led to a proliferation of mobile applications (apps) [13]. A popular category within all these apps are the health and fitness-related apps [14], mostly aimed at supporting dietary change and weight management [15,16]. Those apps usually include a food diary function, in which users can record the foods consumed and the consumed quantities. Apart from searching in a pre-defined food and beverage list and selecting pre-defined portion sizes [17], various features are available to help identify consumed foods, estimate portion size, and decrease the burden of food entering. Examples of those features are image-based food recognition and barcode scanner. Their potential on reducing the respondents' burden, decreasing the effort of multiple self-administrations and on improving food recording accuracy have been investigated in both experimental and observational epidemiological studies, and have shown some promising results [6,18]. However, the knowledge on the performance of other specific features is still limited [19].

One feature of food diary apps is the recipe function for entering mixed dishes prepared at home. These are dishes consisting of multiple foods, with specific food preparation and often with cooking involved. For user-friendliness, the recipe function should be structured in a way that could easily guide the users in recording necessary information of a recipe. It should be able to assess the recipe intake of an individual, while mixed dishes are often prepared for more than one person [19]. Furthermore, for a better estimation of nutrient intake, an accurate recipe calculation should take nutrient loss of ingredients during cooking and food processing into account [20].

Some food diary apps have introduced a recipe function through the recent years [21,22]. The effectiveness of these recipe functions in capturing the food consumption and nutrient intake has not been fully evaluated. Moreover, the question whether the features of available recipe functions are also appropriate for dietary assessment as part of large-scale studies remains unanswered. Therefore, the aim of this study was to make an inventory of recipe function features in apps that could facilitate the estimation of nutrient intake of a large population. Furthermore, another aim was to evaluate the accuracy of the recipe function in capturing nutrient intake of popular dietary assessment apps by comparing their nutrient calculation with a standard calculation procedure.

2. Materials and Methods

The starting point for app selection was an identification of dietary assessment smartphone apps in the Health & Fitness category of iTunes App Store and Google Play Store in the Netherlands between 15th and 23rd of October 2016. This selection was performed by Maringer et al. [20] and resulted in the identification of 176 dietary assessment apps. Further screening was performed in August 2017. Inclusion of a subselection of apps for this study required the app to meet the following criteria: (1) user rating >3 in iTunes App Store and Google Play Store, (2) user rating count >500 in iTunes App Store and Google Play Store, (3) >10,000 downloads in the both stores, (4) a recipe function which was freely available, actually present and functional. A recipe function was defined as “a functionality in which the user can create a mixed dish by entering and specifying the amount of each ingredient within the dish” [23,24]. Each app underwent initial screening based on descriptions and associated images in the app stores to check for the presence of a recipe function. Apps were downloaded onto a OnePlus 3T smartphone running Android 7.1.1 and a Huawei Mate 8 running EMUI 5.0.1 for analysis. The apps were checked manually to confirm whether a recipe function was freely available, actually present, and functional. Basic descriptive information about the apps was identified, such as app name, version number, operating platforms, number of installs, ratings, whether they can synchronize with their website, and country of origin. Subsequently, the recipe function of the selected apps was evaluated.

To our knowledge, no widely accepted standard evaluation of the quality of the recipe function of apps exists. Therefore, a criteria list was made for evaluating features in the individual recipe function of apps. For each feature on the criteria list a rubric of assessment was created with a 1 (low)–5 (high) scoring scale. The criteria list and assessment rubric were modified upon findings from a pilot scoring and feedback from two nutritionists and three dietitians with different specializations. The criteria list and assessment include the following aspects of creating an individual recipe: options in searching ingredients, ways to record relevant information of the recipe, whether raw or cooked ingredients could be selected, consumed amount for both ingredients and the whole recipe, energy and nutrient expression, and whether the recipe could be saved and edited later (Table 1). Two researchers scored all the selected apps according to the criteria list independently. Inconsistent scores among the two researchers were discussed to reach agreed final scores. For scoring the criterion whether both raw and cooked foods are available in the food list, nine foods from the three most frequently used Dutch recipes (explained in next paragraph) were entered in each app (kale, potato, milk, mushroom, onion, salami, beef, pepper, and tomato).

To be able to evaluate the accuracy of energy and nutrient content estimations, three recipes were entered into the individual recipe function of each app. The selection of recipes was performed by exploring the most frequent reported recipes in the Dutch diet using the data of the Dutch National Food Consumption Survey (DNFCS) 2007–2010 [25]. Three recipes with different preparation methods, like stewing, baking, and frying, were chosen from the twenty most frequently consumed recipes. The chosen recipes were boerenkool stamppot (mashed potato with kale), pizza with salami, tomato, and mushrooms, and hachee (a traditional Dutch stew based on beef and onions). Raw ingredients of the recipes were entered in the selected apps and a set of rules for entering ingredients were followed, in case the exact match of food items or amount indications could not be found across apps. If available, energy, macro- and micronutrient values of the recipe were obtained based on the displayed nutrient content in the app. For those apps where the nutrient contents were not shown at the recipe level, values from ingredients of a recipe were added up by researchers. Then, nutrient contents from the apps were compared with nutrient contents derived from the Dutch food composition database (NEVO) [26]. To account for nutrient loss due to cooking, retention factors suggested by the European Food Information Resource [27] were applied to the nutrients derived by NEVO, see complete calculation in Supplementary Material (Table S1–S11). A retention factor larger than 0 and lower than 1 implied nutrient loss due to cooking. A retention factor of 1 was used for energy and macronutrients for all ingredients in all recipes since they were not easily affected by cooking. Next to energy and macronutrient, micronutrients such as sodium, potassium, vitamin A represented as retinol equivalent (RE), vitamin C, calcium, vitamin E, vitamin B1, vitamin B2, vitamin B6, vitamin B12, and folate were selected for comparison between apps and the reference measure. Of these, sodium, potassium, and vitamin E had a retention factor of 1 for all ingredients in the three recipes mentioned above, hence, were deleted from analysis. Calcium also had a retention factor of 1, but was maintained in the analysis as an example.

General characteristics of the 12 evaluated dietary assessment apps with recipe function were summarized. For each app, the mean score and standard deviation over all nine criteria was calculated (see Table 1). The mean and standard deviation of scores across apps were calculated for each criterion. Energy and nutrient content estimations of the three recipes for each app were analyzed using descriptive statistics. For nutrients with retention factor of 1, a direct comparison could be made with the nutrient contents derived from NEVO combining nutrient contents of raw ingredients in the appropriate amounts. For the micronutrients with retention factors below 1, the reference was the NEVO nutrient contents of the raw ingredients after applying the relevant retention factors. For showing the effect of the retention factors, a comparison with NEVO nutrient contents of raw ingredients without applying retention factors was also made. A difference in values between apps and the reference of more than 5% from the Daily Reference Intake (DRI) for adults was considered out of range [28] .

To visualize the correlation between apps and nutrients, a principal component analysis (PCA) was conducted for each recipe separately with energy and macronutrients divided by their DRIs being set as variables. The first two principal components represent the most variation. This was done for energy and macronutrients only, since only 3 apps showed information on absolute amounts of micronutrients. The descriptive statistics were calculated using Excel 2016 software and the PCA was conducted in R version 3.5.0 (The R Foundation for Statistical Computing, Vienna, Austria).

Table 1. Rubric for assessment of the individual recipe function in dietary assessment apps, giving a score between 1 (low) and 5 (high) per feature.

Feature	Mark for feature 1	2	3	4	5
Recipe creation options (name, photo, ingredients, servings)	The user can only create a recipe by adding ingredients and amounts	The user can create a recipe by giving it a name and adding ingredients and amounts	The user can create a recipe by giving it a name, adding ingredients and amounts, and number of servings.	The user can create a recipe by giving it a name, add ingredients and amounts, number of servings, and explanation of preparation.	The user can create a recipe by giving it a name, add ingredients and amounts, number of servings, explanation of preparation, and a photo.
Ingredients search options within a recipe	Can only search in one way	Can search in 2 or 3 ways	Can search in 4 or 5 ways	Can search in 6 or 7 ways	Can search in 8 or more ways
Reminders for frequently forgotten ingredients (e.g., olive oil, butter, salt)	App does not give reminders for frequently forgotten ingredients				App gives reminders for frequently forgotten ingredients
Preparation indication of ingredients	It is unclear whether the entered ingredients are prepared or not	The user can only select the prepared or the unprepared ingredient from the food list	The user can select both the prepared and the unprepared ingredients from the food list for some foods	The user can select both the prepared and the unprepared ingredients from the food list for all foods	The user can select an ingredient and indicate the preparation (unprepared, prepared (cooked, grilled, etc.))
Entering consumed amount at recipe level	User cannot indicate consumed amount	User can indicate the consumed amount, but the type of indication of the amount is inappropriate	User can indicate the consumed amount, and the appropriate type(s) of indication is given. However, inappropriate amounts are also given	User can indicate consumed amount and the appropriate type(s) of indication are given	User can indicate the consumed amount and the user can choose from a lot of appropriate types of indications (grams, portion in grams, portion as photo, fraction of recipe) OR can manually add amount indications
Entering prepared amount at ingredient level	User cannot indicate prepared amount	User can indicate the prepared amount, but the type of indication of the amount is limited (1 or 2 options) OR other types of indications (portion in grams, portion as photo, fraction of recipe)	User can indicate prepared amount from more than 2 options.	User can indicate prepared amount from more than 2 options. AND other types of indications (portion in grams, portion as photo, fraction of recipe)	User can indicate prepared amount from more than 2 options, and other types of indications (grams, portion in grams, portion as photo, fraction of recipe), and can manually add amount indications

Feature	Mark for feature				
	1	2	3	4	5
Save and edit function for recipe	The user can create recipe, but cannot save it to use it later	The user can create a recipe and save it to use it later	The user can create a recipe and save it in a categorized way OR the user can create a recipe and edit it; premium only	The user can create a recipe and edit it later	The user can save the created recipe to use it later, edit it later on, and can save it in a categorized way
Energy and macronutrient information at recipe level	Energy and macronutrient content are not shown	Energy content is shown in kcal (KJ), macronutrient content is not shown	Energy content is shown in kcal (KJ), macronutrient content is shown in grams OR energy is shown in % of Reference Daily Allowance (RDA)*; premium only	Energy content is shown in kcal (KJ) and % of RDA, macronutrient content is shown in grams OR macronutrient content is shown in grams and % of RDA; premium only	Energy content is shown in kcal (KJ) and % of RDA, macronutrient content is shown in grams and % of RDA
Micronutrient information at recipe level	No micronutrient information available	Micronutrient information exists for only premium account	Information on less than 3 micronutrients	Information on 3–6 micronutrients	Information on more than 6 micronutrients

*Reference Daily Allowance (RDA): The average daily dietary intake level sufficient to meet the nutrient requirement (for the specified indicator of adequacy) of nearly all (97% to 98%) healthy individuals in a particular life stage and gender group .

3. Results

3.1. App Selection

The starting point was a selection of 176 popular dietary assessment apps with food recording and available in English identified by Maringer et al. [21]. Then, apps were further narrowed down, with inclusion criteria of a user rating >3 in the iTunes App Store and Google Play Store, a user rating count > 500 in iTunes App Store and Google Play Store, >10,000 downloads in the Google Play Store, and a claimed recipe function in the app description. After manually checking for the presence of an individual recipe function in 30 included apps, 17 apps were excluded from further evaluation because of dysfunction of the app, the absence or dysfunctionality of a recipe function, or the inability to use the app due to requirements of a membership. After final exclusion of one app with a non-functioning individual recipe function, a total of 12 apps (21% of 57) were selected for evaluation in detail (Figure 1).

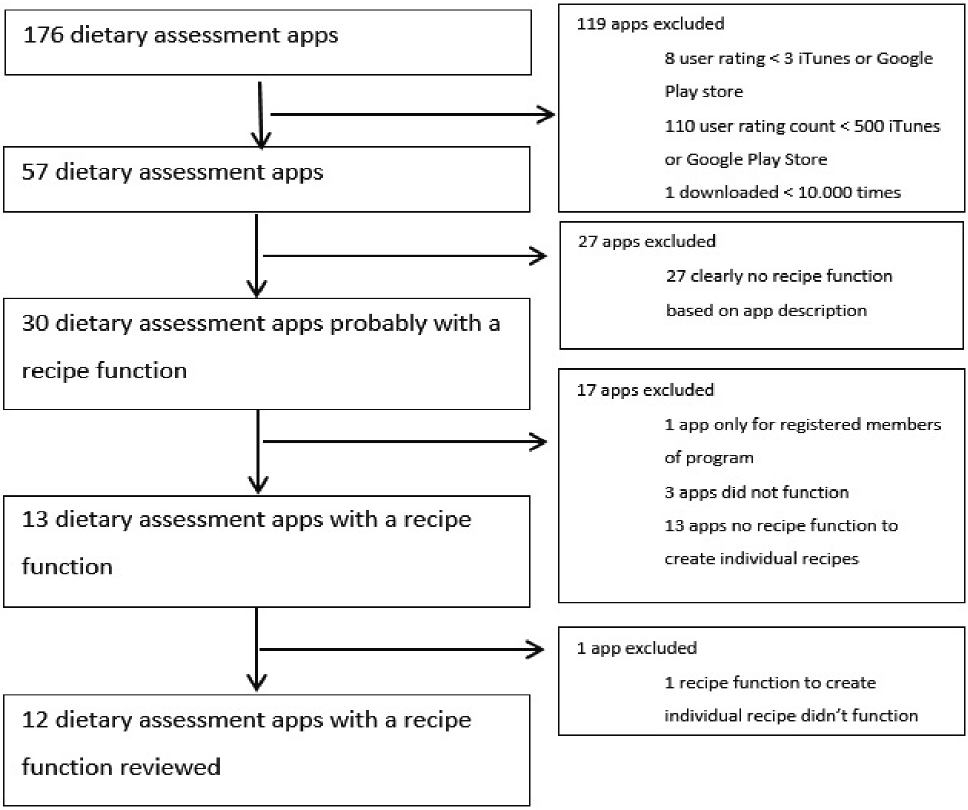


Figure 1. Flow diagram of selection procedure of dietary assessment apps with recipe function showing the number of apps included or excluded.

General characteristics of the remaining 12 apps can be found in Table 2. All apps operated on an Android platform, whereas IOS ranked as the second most-prevalent platform (10 apps). The highest number of installs was 50 million with 1844 thousand ratings for MyFitnessPal, the lowest was 100 thousand installs and 2000 ratings for Nutracheck. The rating scores among the apps ranged from 4.2 to 4.6 with the maximum score of 5.0. Four apps were made by US companies, two apps were made in Germany, and the rest of apps were made in other countries, mostly northwest Europe.

3.2. Qualitative Recipe Function Assessment

Agreed scores given to recipe functions of each app are shown in Table 3. Mean overall score of both apps and criteria was 3.0 (out of 5.0). The app Calories! had the highest score for its recipe function with an average score of 3.9 however, in contrast, Calories! had a rating score and number of installations at the lower range compared to other apps (Table 2). MyPlate and Health Infinity, on average, had the lowest scores of 2.2 and 2.3, respectively.

The apps that had relative higher popularity, such as MyFitnessPal, Lose It!, Lifesum, and MyPlate, did not have any criterion that scored 5, while Calories! was achieved a score of 5 three times. Health Infinity scored 1 most often (three times) compared to other apps.

Specifically, most of the evaluated apps could save a self-created recipe and edit it later, hence, this criterion ranked the highest (mean = 4.3) compared to other criteria. None of the apps included reminders for frequently forgotten ingredients, therefore, all apps scored 1 for that criterion. The available options that existed for searching ingredients for recipes included text search, barcode scanning, voice record, recent/frequent/saved food, create new food, choose from categories, and choose from a list of all food in alphabetic order. The number of options ranged from 2 to 6, where half of the apps had only 2 to 3 options, while only Nutracheck had all 6 options. The most frequently adopted options were search in a textbox and barcode scanning. FatSecret and Virtuagym Food had four searching options for food entering, but only two options for adding ingredients to recipes. In terms of options in searching raw or cooked foods, nearly all apps had both raw and cooked options for all or at least some foods in their dataset (mean = 3.3). An exception was The Secret of Weight, where, for the most foods, the text indicated raw while the picture showed cooked foods. In terms of indicating consumed amount in both ingredients and recipes, in Calories!, one could manually add a new serving unit to ingredients but not in recipes whereas, in Virtuagym Food, this was the other way around. Health Infinity had no options to chooe the amount of recipe consumed (scored as 1), and had only one built-in option when choosing the amount of ingredients. In terms of macronutrient information, Calories! was the only app that had energy and macronutrients expressed as both absolute amounts (mg, µg, etc.) and % of Recommended Daily Allowance (RDA). Most apps had energy and macronutrients shown only in absolute amounts. Since only four apps showed micronutrient for recipes, the average score for micronutrient availability ranked the second lowest with a score of 2.7. Among the apps with micronutrients, Calories! and MyNetDiary had both absolute amounts and % RDA for more than six micronutrients, while Virtuagym Food had only actual amounts. MyFitnessPal had only % RDA of less than six micronutrients.

3.3. Accuracy of Energy and Macronutrient Content Estimations

The differences in energy and macronutrient content estimations of the three recipes between the 12 popular dietary assessment apps and the value derived from NEVO are presented in Table 4. Macronutrient contents for both recipes and ingredients were not available in The Secret of Weight. Heterogeneity in differences was observed between recipes and between nutrients. Pizza had fewer differences >5% (n = 7) in the DRI as compared to boerenkool stamppot (n = 10) and hachee (n = 12). Carbohydrates (n = 2) and energy (n = 3) contents had fewer differences >5% in the DRI than protein (*n* = 13) and fat (*n* = 11). In total, around 20% of the differences were >5% DRI. Most apps underestimated the macronutrient content in boerenkool stamppot and pizza, while this was not observed in hachee.

With 7 out of 12, Nutracheck had the most discrepancies >5% in the DRI compared to the reference, mainly caused by a discrepancy in fat and protein contents. YAZIO and Lifesum only had one difference of more than 5%. Health Infinity had lower protein contents in all three recipes, whereas Lose It! had lower fat in all three recipes. Virtuagym Food and YAZIO had similar patterns in all recipes, and both had lower fat in hachee as outliers. MyNetDiary had all macronutrients being out of range once, including a lower carbohydrate, lower protein, and higher fat in three recipes, respectively. In Figure 2, apps are plotted against the first and second principal component of all differences in macronutrient contents. Macronutrients plotted further from the center indicate a larger variance. Apps situated in the same direction with a certain nutrient indicate an overestimation of the nutrient and vice versa. Nutracheck laid outside compared to other apps for all three recipes. MyFitnessPal was the only app without discrepancies of more than 5%. Therefore, it was located around the center of the graph in all three recipes.

Table 2. General characteristics, such as platforms available, number of installs on Google Play Store, user rating on Google Play Store and country of twelve popular dietary assessment apps with a recipe function (n = 12).

	App Name (Version)	**Platforms**	**Installs Google Play Store (Million)**	**Rating Google Play Store (# Ratings/1000)**	**Country**
1	**MyFitnessPal (18.6.0)**	Android, IOS, Windows Phone	50–100	4.6 (1844)	USA
2	**FatSecret (7.8.27)**	Android, IOS, Windows Phone, Watch OS, Blackberry OS	10–50	4.4 (223)	Australia
3	**YAZIO (4.0.1)**	Android, IOS	5–10	4.6 (109)	Germany
4	**Lose It! (9.4.5)**	Android, IOS	5–10	4.4 (68)	USA
5	**Lifesum (6.2.4)**	Android, IOS, Watch OS, Android Wear,	5–10	4.4 (165)	Sweden
6	**MyPlate (3.2.2)**	Android, IOS, Watch OS	1–5	4.6 (22)	USA
7	**MyNetDiary (6.4.7)**	Android, IOS, Watch OS	1–5	4.5 (26)	USA
8	**Calories! (8.1.6)**	Android	1–5	4.3 (10)	Germany
9	**The Secret of Weight (2.4.24)**	Android, IOS	1–5	4.3 (14)	France
10	**Virtuagym Food (2.4.0)**	Android, IOS	1–5	4.5 (28)	Netherlands
11	**Health Infinity (HI) (2.0.58)**	Android	0.1–0.5	4.2 (9)	India
12	**Nutracheck (5.0.12)**	Android, IOS	0.1–0.5	4.3 (2)	UK

4

Table 3. Agreed scores for the recipe function of 12 popular dietary assessment apps using the criteria list based on a 1(low)–5 (high) scale.

App Name (Version) / **Criteria List**	MyFitnessPal (18.6.0)	FatSecret (7.8.27)	YAZIO (4.0.1)	Lose It! (9.4.5)	Lifesum (6.2.4)	MyPlate (3.2.2)	MyNetDiary (6.4.7)	Calories! (8.1.6)	The Secret of Weight (2.4.24)	Virtuagym Food (2.4.0)	Health Infinity (HI) (2.0.58)	Nutracheck (5.0.12)	Mean	SD
Options (name, photo, ingredients, servings)	3	5	5	3	4	2	5	4	5	2	2	3	3.6	1.2
Options to search ingredients	2	2	3	3	3	2	3	3	2	2	2	4	2.6	0.6
Reminders for frequently forgotten ingredients (e.g., oil, spices, salt)	1	1	1	1	1	1	1	1	1	1	1	1	1.0	0.0
Entering ingredients—preparation indication	4	3	3	4	3	3	3	4	2	4	4	3	3.3	0.6
Consumed amount recipe level	4	4	4	4	4	2	4	4	4	5	1	4	3.7	1.0
Consumed amount ingredient level	3	3	3	3	3	3	3	5	2	3	2	3	3.0	0.7
Save and edit	4	5	5	4	4	4	3	4	4	4	5	5	4.3	0.6
Energy and macronutrient expression at recipe level	4	4	3	3	3	2	3	5	2	3	3	3	3.2	0.8
Micronutrient availability at recipe level	4	3	3	1	2	1	5	5	1	5	1	1	2.7	1.6
Mean	3.2	3.3	3.3	2.9	3.0	2.2	3.3	3.9	2.6	3.2	2.3	3.0	3.0	0.5
SD	1.0	1.2	1.2	1.1	0.9	0.9	1.2	1.2	1.3	1.3	1.3	1.2	0.9	-

Table 4. Difference in energy (kcal) and macronutrient content (gram) estimations for one portion of each of three recipes between 12 dietary assessment apps and reference values using NEVO.

Recipes	Macronutrients	NEVO [a]	MyFitness Pal	FatSecret	YAZIO	Lose It!	Lifesum	MyPlate	MyNet Diary	Calories!	The Secret of Weight	Virtuagym Food	Nutra check	HI	Mean	SD
Boerenkool stamppot	Energy (kcal)	472	4	−42	10	−16	−69	−28	−53	−93	−116*	−62	59	−44	−38	46
	Fat (g)	10.9	−0.2	−5.1*	−0.4	−3.7*	−1.0	−0.6	−0.2	−0.9	-	0.9	6.6*	−2.9	−0.7	2.9
	Protein (g)	17.0	0.1	0.4	0.8	0.9	−5.2*	−1.7	−0.2	−5.3*	-	−1.9	−11.1*	−17.0*	−3.6*	5.4
	Carbohydrate (g)	70.4	−0.1	0.3	11.8	1.2	−2.9	10.2	−15.1*	−14.1*	-	−11.4	−9.0	−6.1	−3.2	8.6
Pizza with salami, tomato, and mushroom	Energy (kcal)	483	−36	−5	−2	−42	−5	−35	0	−24	−7	−8	−47	−41	−21	17
	Fat (g)	25.9	−2.6	−0.3	0.3	−2.9	−0.3	−4.4*	−0.7	−1.6	-	−0.1	−5.4*	−2.9	−1.9	1.8
	Protein (g)	22.1	−2.3	−1.2	−0.2	−2.7*	−0.8	−2.6*	−5.1*	−1.0	-	−0.9	−2.6*	−3.8*	−2.1	1.4
	Carbohydrate (g)	38.8	0.1	0.6	0.3	1.9	1.9	11.8	−0.8	−2.8	-	−0.4	4.2	−2.8	1.3	3.9
Hachee	Energy (kcal)	316	15	−43	−47	−119*	7	12	75	32	58	−46	142*	19	9	65
	Fat (g)	17.9	2.2	−4.3*	−4.5*	−8.8*	2.5	1.7	8.4*	−0.3	-	−5.1*	10.8*	2.4	0.4	5.6
	Protein (g)	23.3	−0.9	−0.6	−1.0	−11.2*	−0.8	−21.3*	−1.3	12.5*	-	1.3	9.0*	−0.1	−1.3	8.5
	Carbohydrate (g)	13.7	1.7	3.8	3.7	3.8	−0.9	2.3	−4.7	−4.1	-	−0.5	3.1	−1.7	0.6	3.0

[a] Energy and macronutrient contents of one recipe portion by adding nutrient contents of raw ingredients derived from Dutch food composition database (NEVO); retention factors were all 1. * Discrepancy with reference >5% of the Dietary Reference Intakes (DRI), which is 100 kcal out of 2000 kcal for energy, 3.5 g out of 70g for fat, 2.5 g out of 50g for protein, and 13 g out of 260g for carbohydrate.

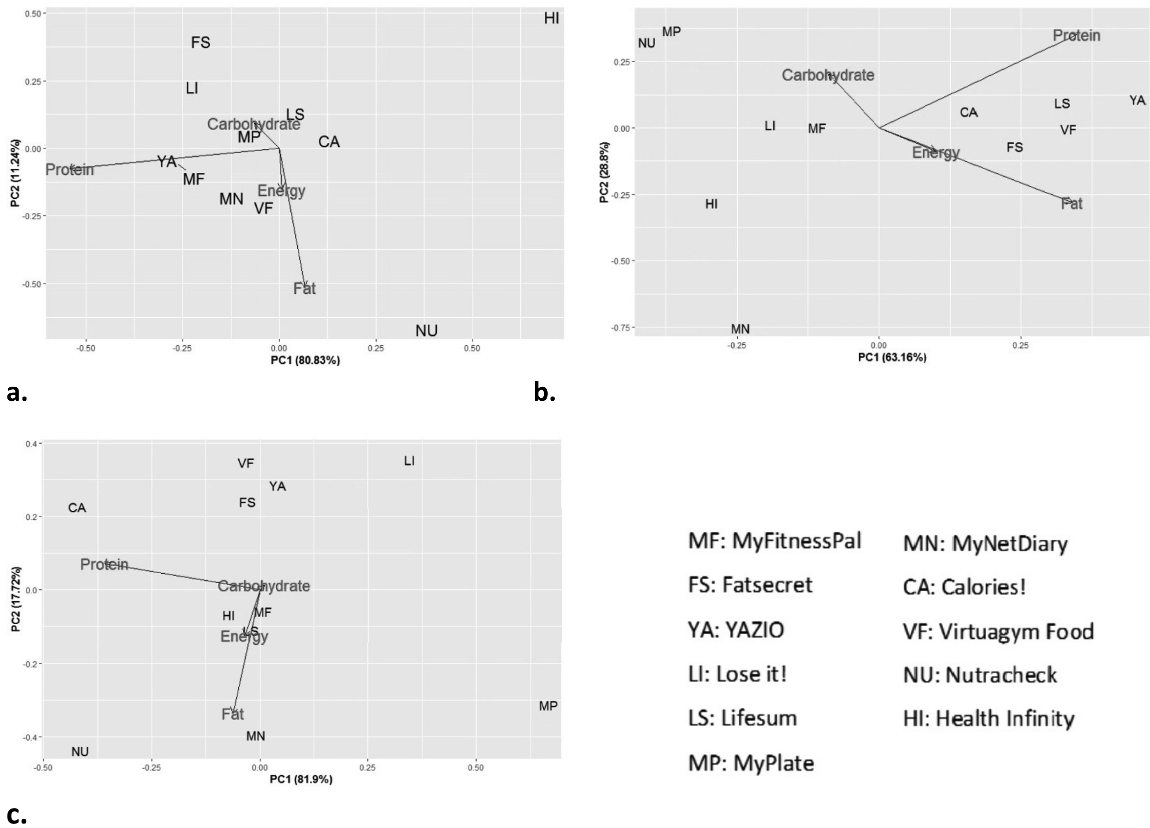


Figure 2. Principal Component Analysis (PCA) showing the variation strength and trend of macronutrient difference compared with the reference contents from different apps in (**a**) Boerenkool stamppot, (**b**) Pizza with salami, tomato, and mushrooms, (**c**) Hachee.

3.4. Accuracy of micronutrient content estimations

The micronutrient contents were analyzed for MyNetDiary, Calories! and Virtuagym in which it was available. The differences in micronutrient content estimations of the three recipes between the three popular dietary assessment apps, the micronutrient calculated from NEVO values in raw foods and the reference where retention factors was applied to NEVO are presented in Table 5. For most micronutrients except calcium, applying retention factors resulted in lower micronutrient levels than micronutrient levels in raw ingredients. The relative differences between the reference and using NEVO without applying retention factors ranged from 0% for calcium in all recipes, vitB12 in stamppot and vitB2 in hachee to more than 45% for vitamins B6, B12 and folate in hachee. Over the 3 recipes, 8 out of 24 differences (33%) were relevant (>5% of DRI) in case of a high content and high vulnerability of these nutrients of raw ingredients in a certain recipe. The relatively large difference in vitamin B6 and B12 in Hachee can be explained by the sensitivity to heat and the two cooking procedures in this recipe, i.e. frying and stewing. Whereas, boerenkool stamppot (*n* = 5) had more relevant differences than the other two recipes (*n* = 1 and 2 respectively), due to its high contents of vitamin C, vitamin A, vitamin B1, vitamin B6 and folate even if the retention factor was not so different from 1 (for example, vitamin A with a retention factor of 0.9).

A larger proportion of difference >5% DRI was found in micronutrients (49%) than in energy and macronutrients (20%) when compared with the reference values. Among the three apps, MyNetDiary showed more differences > 5% DRI (*n* = 14 out of 24) than the other two apps (Virtuagym *n* =10, Calories! *n* =11) when comparing micronutrient values with the reference. In contrast to macronutrient comparisons, apps more often overestimated the contents of micronutrient in the recipes. The number and extent of overestimations were slightly larger when comparing with the reference than comparing with NEVO without applying retention factors, since the retention factors resulted in lower micronutrient contents in the reference values. The proportions of relevant differences found after comparing the apps to NEVO with or without applying retention factors were rather similar (49% vs 51%), illustrating that in many cases the effects of differences in nutrient databases were much larger than differences due to applying retention factors.

Table 5. Comparison of micronutrient contents between recipes added by raw ingredients from three apps with recipes added by raw ingredients from the Dutch food composition database (NEVO), with NEVO multiplied by retention factors.

Recipes	Micronutrients	NEVO[a]	R[b]	NEVO-R	MyNetDiary App	MyNetDiary App-NEVO	MyNetDiary App-R	Calories! App	Calories! App-NEVO	Calories! App-R	Virtuagym App	Virtuagym App-NEVO	Virtuagym App-R
Boerenkool stamppot	Calcium(mg)	494	494	0	431	–63*	–63*	573	80*	80*	391	–102*	–102*
	Vitamin C(mg)	294	187	107*	327	33*	140*	327	33*	140*	362	68*	174*
	Vitamin A(µg)	1774	1606	168 *	2557	783*	951*	2320	546*	714*	74	–1701*	–1532*
	Vitamin B1(mg)	0.66	0.60	0.06*	0.32	–0.34*	–0.28*	0.57	–0.09*	–0.03	0.49	–0.17*	–0.11*
	Vitamin B2(mg)	0.43	0.41	0.02	0.56	0.13*	0.15*	0.79	0.36*	0.38*	0.48	0.05	0.07
	Vitamin B6(mg)	1.49	1.34	0.15*	1.38	–0.11*	0.04	1.70	0.21*	0.36*	1.30	–0.19*	–0.04
	Vitamin B12(µg)	0.11	0.11	0.00	0.43	0.32*	0.32*	-	-	-	0.19	0.08	0.08
	Folate(µg)	198	142	56*	407	208*	265*	94	–104*	–48*	-	-	-
Pizza	Calcium(mg)	339	339	0	293	–46	–46	290	–48	–48	293	–46	–46
	Vitamin C(mg)	6	5	1	5	–1	3	8	2	3	5	–1	0
	Vitamin A(µg)	188	183	5	205	17	22	204	17	22	97	–91 *	–86*
	Vitamin B1(mg)	0.21	0.18	0.03	0.75	0.54*	0.57*	0.29	0.08*	0.11*	0.77	0.56*	0.59*
	Vitamin B2(mg)	0.31	0.30	0.01	0.62	0.31*	0.32*	0.38	0.07	0.08*	0.62	0.31*	0.32*
	Vitamin B6(mg)	0.26	0.24	0.02	0.27	0.01	0.03	0.31	0.05	0.07	0.26	0.00	0.02
	Vitamin B12(µg)	1.10	1.01	0.09	1.00	–0.10	–0.01	-	-	-	1.02	–0.08	0.01
	Folate(µg)	92	67	24*	129	38*	62*	45	–47*	–23*	77	–15	10
Hachee	Calcium(mg)	51	51	0	66	15	15	48	–3	–3	67	16	16
	Vitamin C(mg)	6	5	1	8	2	3	7	1	3	9	3	4
	Vitamin A(µg)	136	129	7	108	–28	–21	123	–13	–6	94	–42*	–34
	Vitamin B1(mg)	0.10	0.06	0.04	0.19	0.09*	0.13*	0.16	0.06*	0.10*	0.19	0.09*	0.13*
	Vitamin B2(mg)	0.19	0.19	0.00	0.21	0.02	0.02	0.24	0.05	0.05	0.25	0.06	0.06
	Vitamin B6(mg)	0.39	0.20	0.19*	0.73	0.34*	0.53*	0.34	–0.05	0.14*	0.73	0.34*	0.53*
	Vitamin B12(µg)	2.95	1.46	1.49*	2.70	–0.25*	1.24*	-	-	-	2.69	–0.26*	1.23*
	Folate(µg)	28	15	13	57	29*	42*	29	1	14	13	–15	–2
# >5% DRI				8		15	14		10	11		12	10
# positive				8		11	12		7	9		5	7

[a] Micronutrient contents of one recipe portion by adding nutrient contents of raw ingredients derived from Dutch food composition database (NEVO). [b] The reference measure where retention factors (RF) were multiplied by each micronutrient content derived from NEVO. *Discrepancy with reference >5% of the Dietary Reference Intake (DRI) which is 5 mg for vitamin C, 49 mg for calcium, 35 µg for vitamin A, 0.06 mg for vitamin B1, 0.08 mg for vitamin B2, 0.08 mg for vitamin B6, 0.20 µg for vitamin B12, and 17 µg for folate.

4. Discussion

The current study evaluated the recipe function that was available in only one-fifth of the popular available food diary apps. We found a varying quality of recipe features across selected apps which were, on average, judged as suboptimal from research perspectives. Furthermore, capturing the true nutrient intake of mixed dishes is a challenge for this innovative dietary assessment method. A comparison of energy, macro-, and micronutrient contents of recipes between apps with a reference standard recipe calculation showed variation in terms of their ability to accurately estimate nutrient contents. In only three apps was micronutrient information available for recipes, and none of these apps included a procedure to take nutrient losses due to recipe processing into account, and the variability in micronutrient content databases was large.

This is the first study to evaluate the recipe function of current popular dietary assessment apps in a standardized way in which the quality assessment was performed using a rubric of assessment which was made prior to the evaluation. The scores of recipe function were discussed by two researchers, which has eliminated mistakes and the bias of scoring. From the quality assessment of the recipe functions, apps were given a mean overall score of 3.0 (out of 5.0) where the highest score was 3.9 and the lowest 2.2. No correlations were found between the scores given in this study and the popularity and user ratings in app stores. This could illustrate that the recipe function was not the main aspect contributing to users' overall app-experiences, or that researchers and users have different needs for dietary apps [9]. Some simplified features might be favored by users since it was observed that the user's time invested for understanding and learning about an app should be small to sustain long-term app usage [30], whereas researchers are more concerned with features that could enable detailed and accurate data collection. This preference gap between the app users and researchers is important to select suitable features to be included in dietary assessment tools for large nutrition monitoring studies.

Although the quality of recipe function in popular apps was not investigated before [13], several features of a recipe function were investigated by others since they are also relevant for recording food intake. In terms of options for searching ingredients in apps from the current study, all apps had a text searching option and the majority of the apps had a barcode function. Barcode scanning has been shown to save time and was favored by users in recording branded food items, however, the resulting nutrient intake estimation depends largely on the quality of the underlying food composition database within the app [31]. An aspect in which these apps differ from many web-based tools is that most of them do not have portion images, which may due to limited space in the user interface. Previous research has found that the incorporation of portion images was preferred by all age groups [9]. However the overall advantage of using portion images remains unknown [17]. In terms of nutrient information, the energy and macronutrient information was more complete in apps than micronutrient information, and this complied with the fact that energy and macronutrients were more closely correlated with weight change, which was the aim for most apps.

Features specific for creating recipes were evaluated. For instance, in addition to other basic features for entering recipes (i.e., add a name, ingredients, and serving number of the recipe), half of the evaluated apps had the capability to enter a photo and cooking explanation. However, this information was not used by the app to estimate nutrient intake. A photo of the recipe could help identify and estimate the amount of food consumed by participants, and could also reduce the extent of underreporting, especially for people with low literacy levels [17], while a cooking explanation provided information relevant for nutrient retention estimation. However, with the extra efforts required in using these features, they might be practical only in small-scale studies. Unlike computer/web-based dietary assessment tools for research purposes [32], all apps lack reminders for frequently forgotten ingredients when creating recipes (e.g., oil, spices, sugar, etc.), which may have partly contributed to the systematic underestimation of macronutrients in most apps found in other studies [33]. Also, current apps did not have pre-defined recipes that could be adapted by users whereas, in some computer-based software, standard recipes could be adapted by switching ingredients or changing the amount of ingredients [32]. However, the practicality of above features to be included in apps or to be used by participants, without the help of researchers, remains questionable. As a simpler alternative, the feature for saving frequently consumed or favorite foods in current apps was shown to save the efforts of users from entering the same recipes repeatedly and searching for food in a comprehensive food list [34].

In the present study, differences in energy, macro-, and micronutrient contents were found between the apps and the reference measure, which could be explained by several reasons. There were substantial differences in the nutrient contents of the recipe ingredients between apps, showing the differences in underlying nutrient databases. Apps were made by companies from different countries and they might have incorporated a nutrient database from their own countries which might have varying nutrient contents for certain foods, due to different cultivating environments [35]. Another source of nutrient values might be input from the app users. This has the benefit of customization of food consumed, however, has shortcomings in the accuracy of nutrients and can lead to quality losses in the food database [14].

The inability to enter exactly the same ingredients across the apps and the limited choice of food amounts may additionally explain part of the variation in nutrient estimation [33]. For example, it was difficult to find an exact match of beef steak in hachee, since there was a large variety of beef steak in different apps, and food amounts in grams were not available in some apps. However, for most other recipe ingredients, this problem did not occur. For micronutrients, the difference was also due to applying retention factors to the reference nutrient values, whereas all apps came up with the nutrient content of recipes by simply adding up the nutrient content of each ingredient without taking nutrient retention into account.

Variations of nutrient content of three recipes between apps and the reference measure were observed in the present study, with fewer variations in energy and macronutrient than in micronutrient contents. Similarly, comparable energy contents across apps were also observed in a study where nutrient contents from the barcode scanning of 100 food products in apps were compared with product labels [31]. Likewise, Griffiths et al. compared the results of five commercial apps with thirty 24 h dietary recalls collected using the Nutrition Data System for Research (NDSR), and found a better validity of energy estimation than nutrients [33]. The mean difference of 22 kcal in energy across all apps and recipes in this study was similar with the 30 kcal mean energy difference of 23 apps compared with the three days' weighed food record in the study of Chen et al. [14]. The wider range of energy difference (−167 to 262 kcal) in Chen's study compared to the energy difference in our study (−118 to 141 kcal) is possibly due to a higher number of apps evaluated, and a larger amount of foods being entered in apps in Chen's study. These findings indicated a relatively reliable energy estimation for both generic and branded food items in the current apps. Still, it was noteworthy that the largest difference of around 345 kcal between apps from both studies could impact the accuracy on both individual and population nutrient intake estimations. A trend of underestimation of energy and macronutrient contents in apps compared to reference in our study was consistent with the study by Griffiths et al. The reason in the study of Griffiths was because the food preparation details were captured by the reference (NDSR), but not in the apps. By contrast, in our study, the food details were equally captured by both the reference and apps, and the reporting bias by participants did not exist since the foods were being entered by researchers. Hence, the main reason of underestimation is the inaccuracy of the nutrition databases within the apps.

A proper way of calculating the nutrient contents within a recipe requires the consideration of nutrient loss during cooking. Currently, the nutrient retention for foods based on different cooking processes is not calculated automatically in any dietary assessment tools, and none of the apps had instructions on using the recipe function. Although existing recipes in food composition tables take the nutrient loss into account, none of the food composition databases cover all the variations on recipes made individually [14]. Alternatively, cooked ingredients could be chosen from the food list. However, the availability of cooked ingredients was incomplete, and this would also require participants to know the amount of the prepared ingredients (which might be smaller due to shrinkage during preparation). Hence, we entered ingredients as raw ingredients, as that is the most logical option for a user.

This is the first study to investigate the discrepancies of nutrient content between raw ingredients in different apps, compared to a more accurate estimation that takes the nutrient loss into account. Only three out of twelve apps had comprehensive micronutrient information, with both actual amounts and percentage of RDA. The large variation in micronutrient content found in this study implied the importance of choosing the right nutrient database, especially when micronutrient intake estimation is part of the study purposes. The input of raw ingredients potentially leads to overestimation of several heat-sensitive micronutrients, which was shown in the micronutrient comparison between NEVO with the reference method in this study. Moreover, the results showed that the extent of difference depends largely on the nutrient contents in the recipe. Therefore, it was suggested that retention factors are most influential when applied to recipes with high micronutrient contents (e.g., boerenkool stamppot).

NEVO was chosen as the reference measure for nutrient estimations, which was a well-maintained food composition database that had all the data on the nutrition values that were assessed and has a standardized food-compiling procedure that follows the guidelines set by EuroFIR [36,37]. Retention factors applied in this study were the most up-to-date values from the harmonization of retention factors provided by 17 EuroFIR partners [38]. However, the results of nutrient differences may lack representativeness in this study, due to a limited recipe selection. To develop a full picture of the importance of recipe calculation, additional studies, that include more recipes and an evaluation on their contribution to population nutrient intake, will be needed. Furthermore, the evaluation was done only from a research perspective in this study, while user perspective was not analyzed for the apps. Especially factors that could affect the individual's ability to accurately enter the recipe consumed were not examined. Further development of an app for large nutrition monitoring studies would benefit from an evaluation on app users' perspectives.

5. Conclusion

In popular food diary apps, the quality of recipe functions is suboptimal from a research perspective. All apps follow a basic nutrition-calculating algorithm, without taking nutrient retention into consideration. This leads to inaccurate nutrient intake estimations in the case that recipes are an important source of micronutrients which are vulnerable to the effects of food processing. Moreover, across apps, there is large variability in nutrient databases. From a research perspective and out of interest regarding micronutrient intake, a balance between user-friendliness and completeness of the recipe function is important. In order to obtain more insight into the need for more complex recipe functionalities, further studies on their potential impact on the nutrient intake estimations in large nutrition-monitoring studies and users' perspective are needed.

References

1. Cade, J.E. Measuring Diet in the 21st Century: Use of New Technologies. P Nutr Soc 2017, 76, 276-282.
2. Shim, J.-S.; Oh, K.; Kim, H.C. Dietary Assessment Methods in Epidemiologic Studies. Epidemiology and Health 2014, 36, e2014009.
3. Klesges, R.C.; Eck, L.H.; Ray, J.W. Who Underreports Dietary Intake in a Dietary Recall? Evidence from the Second National Health and Nutrition Examination Survey. Journal of consulting and clinical psychology 1995, 63, 438-444.
4. Livingstone, M.B.; Robson, P.J.; Wallace, J.M. Issues in Dietary Intake Assessment of Children and Adolescents. The British journal of nutrition 2004, 92 Suppl 2, S213-222.
5. Schober, M.F. Asking Questions and Influencing Answers. Questions about questions: Inquiries into the cognitive bases of surveys 1992, 15.
6. Boushey, C.J.; Spoden, M.; Zhu, F.M.; Delp, E.J.; Kerr, D.A. New Mobile Methods for Dietary Assessment: Review of Image-Assisted and Image-Based Dietary Assessment Methods. The Proceedings of the Nutrition Society 2017, 76, 283-294.
7. Ramachandran, K. Nutrition Monitoring and Surveillance. NFI BULLETIN 2006, 27, 5.
8. Ocke, M.C.; Van Rossum, C.T.M.; de Boer, E.J. Dietary Surveys: National Food Intake. In The Encyclopedia of Food and Health, Caballero, B.; Finglas, P.; Toldrà, F., Eds. Academic Press: Oxford, 2016; Vol. 2, pp 432-438.
9. Carter, M.C.; Albar, S.A.; Morris, M.A.; Mulla, U.Z.; Hancock, N.; Evans, C.E.; Alwan, N.A.; Greenwood, D.C.; Hardie, L.J.; Frost, G.S., et al. Development of a UK Online 24-H Dietary Assessment Tool: Myfood24. Nutrients 2015, 7, 4016-4032.
10. Boshuizen, H.C.; Viet, A.L.; Picavet, H.S.; Botterweck, A.; van Loon, A.J. Non-Response in a Survey of Cardiovascular Risk Factors in the Dutch Population: Determinants and Resulting Biases. Public health 2006, 120, 297-308.
11. Conrad, J.; Nothlings, U. Innovative Approaches to Estimate Individual Usual Dietary Intake in Large-Scale Epidemiological Studies. P Nutr Soc 2017, 76, 213-219.
12. Carter, M.C.; Burley, V.J.; Nykjaer, C.; Cade, J.E. 'My Meal Mate' (Mmm): Validation of the Diet Measures Captured on a Smartphone Application to Facilitate Weight Loss. The British journal of nutrition 2013, 109, 539-546.
13. Xu, W.; Liu, Y. Mhealthapps: A Repository and Database of Mobile Health Apps. JMIR Mhealth Uhealth 2015, 3, e28.
14. Chen, J.; Cade, J.E.; Allman-Farinelli, M. The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment. JMIR Mhealth Uhealth 2015, 3, e104.
15. Aguilar-Martinez, A.; Sole-Sedeno, J.M.; Mancebo-Moreno, G.; Medina, F.X.; Carreras-Collado, R.; Saigi-Rubio, F. Use of Mobile Phones as a Tool for Weight Loss: A Systematic Review. Journal of Telemedicine and Telecare 2014, 20, 339-349.

16. Dennison, L.; Morrison, L.; Conway, G.; Yardley, L. Opportunities and Challenges for Smartphone Applications in Supporting Health Behavior Change: Qualitative Study. J Med Internet Res 2013, 15, e86.
17. Lieffers, J.R.L.; Hanning, R.M. Dietary Assessment and Self-Monitoring: With Nutrition Applications for Mobile Devices. Canadian Journal of Dietetic Practice and Research 2012, 73, e253-e260.
18. Dietary Assessment: A Resource Guide to Method Selection and Application in Low Resource Settings.; FAO: Rome, 2018.
19. 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.
20. Reinivuo, H.; Bell, S.; Ovaskainen, M.-L. Harmonisation of Recipe Calculation Procedures in European Food Composition Databases. J Food Compos Anal 2009, 22, 410-413.
21. Maringer, M.; van't Veer, P.; Klepacz, N.; Verain, M.C.D.; Normann, A.; Ekman, S.; Timotijevic, L.; Raats, M.M.; Geelen, A. User-Documented Food Consumption Data from Publicly Available Apps: An Analysis of Opportunities and Challenges for Nutrition Research. Nutrition Journal 2018, 17, 59.
22. Franco, Z.R.; Fallaize, R.; Lovegrove, A.J.; Hwang, F. Popular Nutrition-Related Mobile Apps: A Feature Assessment. JMIR Mhealth Uhealth 2016, 4, e85.
23. Severi, S.; Bedogni, G.; Zoboli, G.P.; Manzieri, A.M.; Poli, M.; Gatti, G.; Battistini, N. Effects of Home-Based Food Preparation Practices on the Micronutrient Content of Foods. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation 1998, 7, 331-335.
24. Slimani, N.; Deharveng, G.; Charrondiere, R.U.; van Kappel, A.L.; Ocke, M.C.; Welch, A.; Lagiou, A.; van Liere, M.; Agudo, A.; Pala, V., et al. Structure of the Standardized Computerized 24-H Diet Recall Interview Used as Reference Method in the 22 Centers Participating in the Epic Project. European Prospective Investigation into Cancer and Nutrition. Computer methods and programs in biomedicine 1999, 58, 251-266.
25. van Rossum, C.T.M.; Fransen, H.P.; Verkaik-Kloosterman, J.; Buurma, E.M.; Ocke, M.C. Dutch National Food Consumption Survey 2007-2010: Diet of Children and Adults Aged 7 to 69 Years; RIVM-report 350070006.; RIVM: Bilthoven, 2011.
26. RIVM. Nevo-Table. Available online: https://nevo-online.rivm.nl/ (accessed on 23 July 2018),
27. Bognár, A. Tables on Weight Yield of Food and Retention Factors of Food Constituents for the Calculation of Nutrient Composition of Cooked Foods (Dishes). BFE Karlsruhe: Karlsruhe, Germany, 2002.
28. Reference Intake. Available online: https://www.voedingscentrum.nl/encyclopedie/referentie-inname.aspx (accessed on 26 December 2018),

29. Barr, S.I.; Murphy, S.P.; Poos, M.I. Interpreting and Using the Dietary References Intakes in Dietary Assessment of Individuals and Groups. Journal of the American Dietetic Association 2002, 102, 780-788.
30. Spriestersbach, A.; Springer, T. Quality Attributes in Mobile Web Application Development. Lect Notes Comput Sc 2004, 3009, 120-130.
31. Maringer, M.; Wisse-Voorwinden, N.; Veer, P.V.; Geelen, A. Food Identification by Barcode Scanning in the Netherlands: A Quality Assessment of Labelled Food Product Databases Underlying Popular Nutrition Applications. Public health nutrition 2018, 1-8.
32. Slimani, N.; Casagrande, C.; Nicolas, G.; Freisling, H.; Huybrechts, I.; Ocke, M.C.; Niekerk, E.M.; van Rossum, C.; Bellemans, M.; De Maeyer, M., et al. The Standardized Computerized 24-H Dietary Recall Method Epic-Soft Adapted for Pan-European Dietary Monitoring. Eur J Clin Nutr 2011, 65 Suppl 1, S5-15.
33. Griffiths, C.; Harnack, L.; Pereira, M.A. Assessment of the Accuracy of Nutrient Calculations of Five Popular Nutrition Tracking Applications. Public health nutrition 2018, 21, 1495-1502.
34. Rusin, M.; Årsand, E.; Hartvigsen, G. Functionalities and Input Methods for Recording Food Intake: A Systematic Review. International journal of medical informatics 2013, 82, 653-664.
35. Marconi, S.; Durazzo, A.; Camilli, E.; Lisciani, S.; Gabrielli, P.; Aguzzi, A.; Gambelli, L.; Lucarini, M.; Marletta, L. Food Composition Databases: Considerations About Complex Food Matrices. Foods 2018, 7.
36. Castanheira, I.; Roe, M.; Westenbrink, S.; Ireland, J.; Møller, A.; Salvini, S.; Beernaert, H.; Oseredczuk, M.; Calhau, M.A. Establishing Quality Management Systems for European Food Composition Databases. Food chemistry 2009, 113, 776-780.
37. Westenbrink, S.; Oseredczuk, M.; Castanheira, I.; Roe, M. Food Composition Databases: The Eurofir Approach to Develop Tools to Assure the Quality of the Data Compilation Process. Food chemistry 2009, 113, 759-767.
38. Vasquez-Caicedo, A.L.; Bell, S.; Hartmann, B. Report on Collection of Rules on Use of Recipe Calculation Procedures Including the Use of Yield and Retention Factors for Imputing Nutrient Values for Composite Foods; European Food Information Resource: Brussels, Belgium, 2014.

Table S1: Energy and macronutrient differences between app and NEVO for stamppot

	Fat (g)	Protein (g)	Carbohydrate (g)	Energy (kcal)	Energy difference with reference	Fat difference with reference	Protein difference with reference	Carbohydrate difference with reference
Reference raw	10.9	16.95	70.375	472.25	0	0	0	0
Myfitnesspal	10.7	17	70.3	476	3.75	-0.2	0.05	-0.075
Fatsecret	5.78	17.39	70.66	430	-42.25	-5.12	0.44	0.285
YAZIO	10.5	17.7	82.2	482	9.75	-0.4	0.75	11.825
Lose it!	7.2	17.8	71.6	456	-16.25	-3.7	0.85	1.225
Lifesum	9.9	11.8	67.5	403	-69.25	-1	-5.15	-2.875
MyPlate	10.33	15.26	80.53	444	-28.25	-0.57	-1.69	10.155
MyNetDiary	10.75	16.75	55.25	419	-53.25	-0.15	-0.2	-15.125
Calories!	10	11.7	56.3	379.6	-92.65	-0.9	-5.25	-14.075
The Secret of Weight	-	-	-	356.25	-116	-	-	-
Virtuagym Food	11.8	15.03	59.02	410	-62.25	0.9	-1.92	-11.355
Nutracheck	17.5	5.9	61.4	531	58.75	6.6	-11.05	-8.975
HI	8	0	64.3	428.1	-44.15	-2.9	-16.95	-6.075

Table S2: Energy and macronutrient differences between app and NEVO for pizza salami, tomato, mushroom

	Fat (g)	Protein (g)	Carbohydrate (g)	Energy (kcal)	Energy difference with reference	Fat difference wit reference	Protein difference with reference	Carbohydrate difference with reference
Reference raw	25.925	22.075	38.8	482.75	0	0	0	0
Myfitnesspal	23.3	19.8	38.9	447	-35.75	-2.625	-2.275	0.1
Fatsecret	25.64	20.87	39.41	478	-4.75	-0.285	-1.205	0.61
YAZIO	26.2	21.9	39.1	481	-1.75	0.275	-0.175	0.3
Lose it!	23	19.4	40.7	440.4	-42.35	-2.925	-2.675	1.9
Lifesum	25.6	21.3	40.7	478	-4.75	-0.325	-0.775	1.9
MyPlate	21.48	19.49	50.63	448	-34.75	-4.445	-2.585	11.83
MyNetDiary	25.25	17	38	482.75	0	-0.675	-5.075	-0.8
Calories!	24.3	21.1	36	459.2	-23.55	-1.625	-0.975	-2.8
The Secret of Weight	-	-	-	476	-6.75	-	-	-
Virtuagym Food	25.85	21.13	38.42	475	-7.75	-0.075	-0.945	-0.38
Nutracheck	20.5	19.5	43	436	-46.75	-5.425	-2.575	4.2
HI	23	18.3	36	442.1	-40.65	-2.925	-3.775	-2.8

Table S3: Energy and macronutrient differences between app and NEVO for hachee

	Fat (g)	Protein (g)	Carbohydrate (g)	Energy (kcal)	Energy difference with reference	Fat difference with reference	Protein difference with reference	Carbohydrate difference with reference
Reference raw	17.85	23.325	13.7	315.5	0	0	0	0
Myfitnesspal	20	22.4	15.4	330	14.5	2.15	-0.925	1.7
Fatsecret	13.59	22.7	17.51	273	-42.5	-4.26	-0.625	3.81
YAZIO	13.4	22.3	17.4	269	-46.5	-4.45	-1.025	3.7
Lose it!	9.1	12.1	17.5	197	-118.5	-8.75	-11.225	3.8
Lifesum	20.3	22.5	12.8	322	6.5	2.45	-0.825	-0.9
MyPlate	19.51	2.05	15.95	327	11.5	1.66	-21.275	2.25
MyNetDiary	26.25	22	9	390.25	74.75	8.4	-1.325	-4.7
Calories!	17.6	35.8	9.6	347.5	32	-0.25	12.475	-4.1
The Secret of Weight	-	-	-	373	57.5	-	-	-
Virtuagym Food	12.72	24.59	13.17	270	-45.5	-5.13	1.265	-0.53
Nutracheck	28.6	32.3	16.8	457	141.5	10.75	8.975	3.1
HI	20.2	23.2	12	334.4	18.9	2.35	-0.125	-1.7

Table S4. Calcium calculation for three recipes

		Raw weight in recipe (g)	Ca mg in recipe(NEVO)	Ca mg in recipe (Mynetdiary)	Ca mg in recipe (Calories!)	Ca mg in recipe(Virtue)	First and second cooking procedure	First and second retention factor	Ca mg in one portion (NEVO)
Stamppot	kale	250.00	450.00	377.50	530.00	337.50	stew	1.00	450.00
	potatoes	312.50	18.75	28.13	18.75	28.13	stew	1.00	18.75
	butter	10.00	1.70	2.40	1.30	2.40	boil	1.00	1.70
	semi-skimmed milk	18.80	23.12	22.94	23.12	23.31	boil	1.00	23.12
	Total		493.57	430.96	573.17	391.34			493.57
Pizza	Flour	50.00	7.50	7.50	7.80	7.50	bake in oven	1.00	7.50
	Olive oil	4.50	0.00	0.05	0.04	0.05	bake in oven	1.00	0.00
	Tomato puree	11.25	5.51	2.03	4.95	2.03	bake in oven	1.00	5.51
	Yeast	0.88	0.70	0.26	0.24	0.26	bake in oven	-	0.70
	Oregano	0.38	5.91	5.99	0.00	5.99	bake in oven	-	5.91
	Mature cheese 48+	37.50	306.00	270.38	270.00	270.38	bake in oven	1.00	306.00
	Tomato	25.00	2.75	2.50	2.50	2.50	bake in oven	1.00	2.75
	Salami	25.00	8.75	3.25	4.00	3.25	bake in oven	1.00	8.75
	Mushroom	25.00	1.50	0.75	0.75	0.75	bake in oven	1.00	1.50
	Total		338.62	292.70	290.28	292.70			338.62
Hachee	Hachee meat	100.00	5.00	13.00	4.00	14.00	fry in pan and stew	1.00 * 1.00	5.00
	Onion	100.00	29.00	23.00	23.00	23.00	fry in pan and stew	1.00 * 1.00	29.00
	Vinegar	4.50	0.14	1.22	0.72	1.35	stew	1.00	0.14
	Flour	7.50	0.75	1.13	1.20	1.13	stew	1.00	0.75
	Pepper	1.25	5.46	5.54	5.30	5.54	stew	1.00	5.46
	Butter	12.50	2.13	3.00	1.60	3.00	stew	1.00	2.13
	Laurel	1.25	0.00	10.43	0.00	10.43	stew	1.00	0.00
	Cloves	1.25	8.08	7.90	9.10	8.08	stew	1.00	8.08
	Bouillon powder	1.25	0.00	0.75	2.80	0.09	stew	1.00	0.00
	Total		50.55	65.95	47.72	66.60			50.55

4

Table S5. Vitamin C calculation for three recipes

		Raw weight in recipe (g)	VC mg in recipe(NEVO)	VC mg in recipe (Mynetdiary)	VC mg in recipe (Calories!)	VC mg in recipe(Virtue)	First and second cooking procedure	First and second retention factor	VC mg in one portion (NEVO)
Stamppot	kale	250.00	250.00	299.00	252.50	300.00	stew	0.60	150.00
	potatoes	312.50	43.75	28.30	50.00	61.50	stew	0.85	37.19
	butter	10.00	0.00	0.00	0.00	0.00		-	0.00
	semi-skimmed milk	18.80	0.19	0.00	0.00	0.00	boil	0.70	0.13
	Total		293.94	327.30	302.50	361.50			187.19
Pizza	Flour	50.00	0.00	0.00	0.00	0.00	bake in oven	0.70	0.00
	Olive oil	4.50	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Tomato puree	11.25	1.46	1.24	4.19	1.19	bake in oven	0.80	1.17
	Yeast	0.88	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Oregano	0.38	0.00	0.00	0.00	0.01	bake in oven	-	0.00
	Mature cheese 48+	37.50	0.00	0.00	0.00	0.00	bake in oven	0.70	0.00
	Tomato	25.00	3.75	3.50	3.43	3.43	bake in oven	0.80	3.00
	Salami	25.00	0.00	0.00	0.00	0.00	bake in oven	0.80	0.00
	Mushroom	25.00	1.00	0.53	0.53	0.53	bake in oven	0.80	0.80
	Total		6.21	5.26	8.14	5.15			4.97
Hachee	Hachee meat	100.00	0.00	0.00	0.00	0.00	fry in pan and stew	0.75 * 2	0.00
	Onion	100.00	5.00	7.40	7.40	7.40	fry in pan and stew	0.85 * 2	3.61
	Vinegar	4.50	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Flour	7.50	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Pepper	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Butter	12.50	0.00	0.00	0.03	0.03	stew	1.00	0.00
	Laurel	1.25	0.00	0.58	0.00	0.58	stew	1.00	0.00
	Cloves	1.25	1.01	0.00	0.00	1.01	stew	1.00	1.01
	Bouillon powder	1.25	0.00	0.00	0.00	0.01	stew	1.00	0.00
	Total		6.01	7.98	7.43	9.02			4.63

Table S6. Vitamin A calculation for three recipes

		Raw weight in recipe (g)	VA ug in recipe(NEVO)	VA ug in recipe (Mynetdiary)	VA ug in recipe (Calories!)	VA ug in recipe(Virtue)	First and second cooking procedure	First and second retention factor	VA ug in one portion (NEVO)
Stamppot	kale	250.00	1677.50	2475.00	2250.00	0.00	stew	0.90	1509.75
	potatoes	312.50	3.13	7.50	0.00	0.00	stew	0.90	2.81
	butter	10.00	90.50	74.97	70.00	73.80		-	90.50
	semi-skimmed milk	18.80	3.20	0.00	0.00	0.00	boil	1.00	3.20
	Total		1774.32	2557.47	2320.00	73.80			1606.26
Pizza	Flour	50.00	0.00	0.35	0.00	0.00	bake in oven	0.90	0.00
	Olive oil	4.50	0.18	0.00	8.24	0.00	bake in oven	1.00	0.18
	Tomato puree	11.25	31.16	19.13	22.50	0.00	bake in oven	0.90	28.05
	Yeast	0.88	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Oregano	0.38	2.59	2.13	0.00	0.00	bake in oven	-	2.59
	Mature cheese 48+	37.50	129.38	125.25	136.13	96.75	bake in oven	1.00	129.38
	Tomato	25.00	19.25	58.25	37.48	0.00	bake in oven	0.90	17.33
	Salami	25.00	5.25	0.00	0.00	0.00	bake in oven	1.00	5.25
	Mushroom	25.00	0.00	0.00	0.00	0.00	bake in oven	1.00	0.00
	Total		187.81	205.10	204.34	96.75			182.76
Hachee	Hachee meat	100.00	20.00	1.98	34.80	2.20	fry in pan and stew	0.80 * 0.90	12.80
	Onion	100.00	0.00	0.66	0.40	0.00	fry in pan and stew	0.80 * 0.90	0.00
	Vinegar	4.50	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Flour	7.50	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Pepper	1.25	2.38	2.26	0.00	0.00	stew	1.00	2.38
	Butter	12.50	113.13	103.09	87.50	92.26	stew	1.00	113.13
	Laurel	1.25	0.00	0.03	0.00	0.00	stew	1.00	0.00
	Cloves	1.25	0.63	0.33	0.70	0.00	stew	1.00	0.63
	Bouillon powder	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Total		136.13	108.34	123.40	94.46			128.93

4

Table S7. Vitamin B1 calculation for three recipes

		Raw weight in recipe (g)	VB1 mg in recipe(NEVO)	VB1 mg in recipe (Mynetdiary)	VB1 mg in recipe (Calories!)	VB1 mg in recipe(Virtue)	First and second cooking procedure	First and second retention factor	VB1 mg in one portion (NEVO)
Stamppot	kale	250.00	0.50	0.25	0.25	0.28	stew	0.90	0.45
	potatoes	312.50	0.13	0.06	0.31	0.22	stew	0.90	0.11
	butter	10.00	0.00	0.00	0.00	0.00	boil	1.00	0.00
	semi-skimmed milk	18.80	0.03	0.00	0.01	0.00	boil	0.90	0.03
	Total		0.66	0.32	0.57	0.49			0.60
Pizza	Flour	50.00	0.10	0.40	0.05	0.40	bake in oven	0.75	0.08
	Olive oil	4.50	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Tomato puree	11.25	0.02	0.00	0.02	0.00	bake in oven	0.90	0.02
	Yeast	0.88	0.02	0.10	0.01	0.10	bake in oven	-	0.02
	Oregano	0.38	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Mature cheese 48+	37.50	0.00	0.00	0.01	0.01	bake in oven	0.75	0.00
	Tomato	25.00	0.00	0.00	0.01	0.01	bake in oven	0.90	0.01
	Salami	25.00	0.05	0.23	0.18	0.23	bake in oven	0.90	0.04
	Mushroom	25.00	0.02	0.03	0.02	0.02	bake in oven	0.90	0.02
	Total		0.21	0.75	0.29	0.77			0.18
Hachee	Hachee meat	100.00	0.05	0.10	0.10	0.08	fry in pan and stew	0.60 * 0.60	0.02
	Onion	100.00	0.04	0.00	0.05	0.05	fry in pan and stew	0.90 * 0.90	0.03
	Vinegar	4.50	0.00	-	0.00	0.00	stew	1.00	0.00
	Flour	7.50	0.01	0.08	0.01	0.06	stew	0.75	0.00
	Pepper	1.25	0.00	0.01	0.00	0.00	stew	1.00	0.00
	Butter	12.50	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Laurel	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Cloves	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Bouillon powder	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Total		0.10	0.19	0.16	0.19			0.06

Table S8. Vitamin B2 calculation for three recipes

		Raw weight in recipe (g)	VB2 mg in recipe(NEVO)	VB2 mg in recipe (Mynetdiary)	VB2 mg in recipe (Calories!)	VB2 mg in recipe(Virtue)	First and second cooking procedure	First and second retention factor	VB2 mg in one portion (NEVO)
Stamppot	kale	250.00	0.05	0.25	0.75	0.33	stew	0.95	0.05
	potatoes	312.50	0.38	0.31	0.00	0.09	stew	0.95	0.36
	butter	10.00	0.00	0.00	0.01	0.00	boil	1.00	0.00
	semi-skimmed milk	18.80	0.01	0.00	0.03	0.06	boil	0.95	0.01
	Total		0.43	0.56	0.79	0.48			0.41
Pizza	Flour	50.00	0.03	0.25	0.05	0.25	bake in oven	1.00	0.03
	Olive oil	4.50	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Tomato puree	11.25	0.01	0.01	0.01	0.01	bake in oven	0.95	0.01
	Yeast	0.88	0.04	0.04	0.02	0.04	bake in oven	-	0.04
	Oregano	0.38	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Mature cheese 48+	37.50	0.11	0.15	0.15	0.14	bake in oven	0.95	0.10
	Tomato	25.00	0.00	0.00	0.00	0.01	bake in oven	0.95	0.00
	Salami	25.00	0.05	0.08	0.05	0.08	bake in oven	1.00	0.05
	Mushroom	25.00	0.08	0.10	0.10	0.10	bake in oven	0.95	0.07
	Total		0.31	0.62	0.38	0.62			0.30
Hachee	Hachee meat	100.00	0.16	0.20	0.20	0.17	fry in pan and stew	1.00 * 1.00	0.16
	Onion	100.00	0.02	0.00	0.03	0.03	fry in pan and stew	0.95 * 0.95	0.02
	Vinegar	4.50	0.00	-	0.00	0.00	stew	1.00	0.00
	Flour	7.50	0.00	0.00	0.00	0.04	stew	1.00	0.00
	Pepper	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Butter	12.50	0.00	0.00	0.01	0.00	stew	1.00	0.00
	Laurel	1.25	0.00	0.01	0.00	0.01	stew	1.00	0.00
	Cloves	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Bouillon powder	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Total		0.19	0.21	0.24	0.25			0.19

4

Table S9. Vitamin B6 calculation for three recipes

		Raw weight in recipe (g)	VB6 mg in recipe(NEVO)	VB6 mg in recipe (Mynetdiary)	VB6 mg in recipe (Calories!)	VB6 mg in recipe(Virtue)	First and second cooking procedure	First and second retention factor	VB6 mg in one portion (NEVO)
Stamppot	kale	250.00	0.55	0.75	0.75	0.68	stew	0.90	0.50
	potatoes	312.50	0.94	0.63	0.94	0.63	stew	0.90	0.84
	butter	10.00	0.00	0.00	0.00	0.00	boil		0.00
	semi-skimmed milk	18.80	0.01	0.00	0.01	0.00	boil	0.80	0.01
	Total		1.49	1.38	1.70	1.30			1.34
Pizza	Flour	50.00	0.13	0.00	0.05	0.02	bake in oven	0.90	0.11
	Olive oil	4.50	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Tomato puree	11.25	0.03	0.01	0.05	0.01	bake in oven	0.90	0.03
	Yeast	0.88	0.02	0.01	0.01	0.01	bake in oven	-	0.02
	Oregano	0.38	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Mature cheese 48+	37.50	0.01	0.04	0.04	0.03	bake in oven	0.75	0.01
	Tomato	25.00	0.02	0.03	0.02	0.02	bake in oven	0.90	0.02
	Salami	25.00	0.03	0.15	0.13	0.14	bake in oven	0.90	0.02
	Mushroom	25.00	0.03	0.03	0.03	0.03	bake in oven	0.90	0.03
	Total		0.26	0.27	0.31	0.26			0.24
Hachee	Hachee meat	100.00	0.26	0.60	0.20	0.57	fry in pan and stew	0.60 * 0.60	0.09
	Onion	100.00	0.12	0.10	0.12	0.12	fry in pan and stew	0.90 * 0.90	0.10
	Vinegar	4.50	0.00	-	0.00	0.00	stew	1.00	0.00
	Flour	7.50	0.01	0.00	0.02	0.00	stew	0.80	0.01
	Pepper	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Butter	12.50	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Laurel	1.25	0.00	0.02	0.00	0.02	stew	1.00	0.00
	Cloves	1.25	0.00	0.01	0.00	0.01	stew	1.00	0.00
	Bouillon powder	1.25	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Total		0.39	0.73	0.34	0.73			0.20

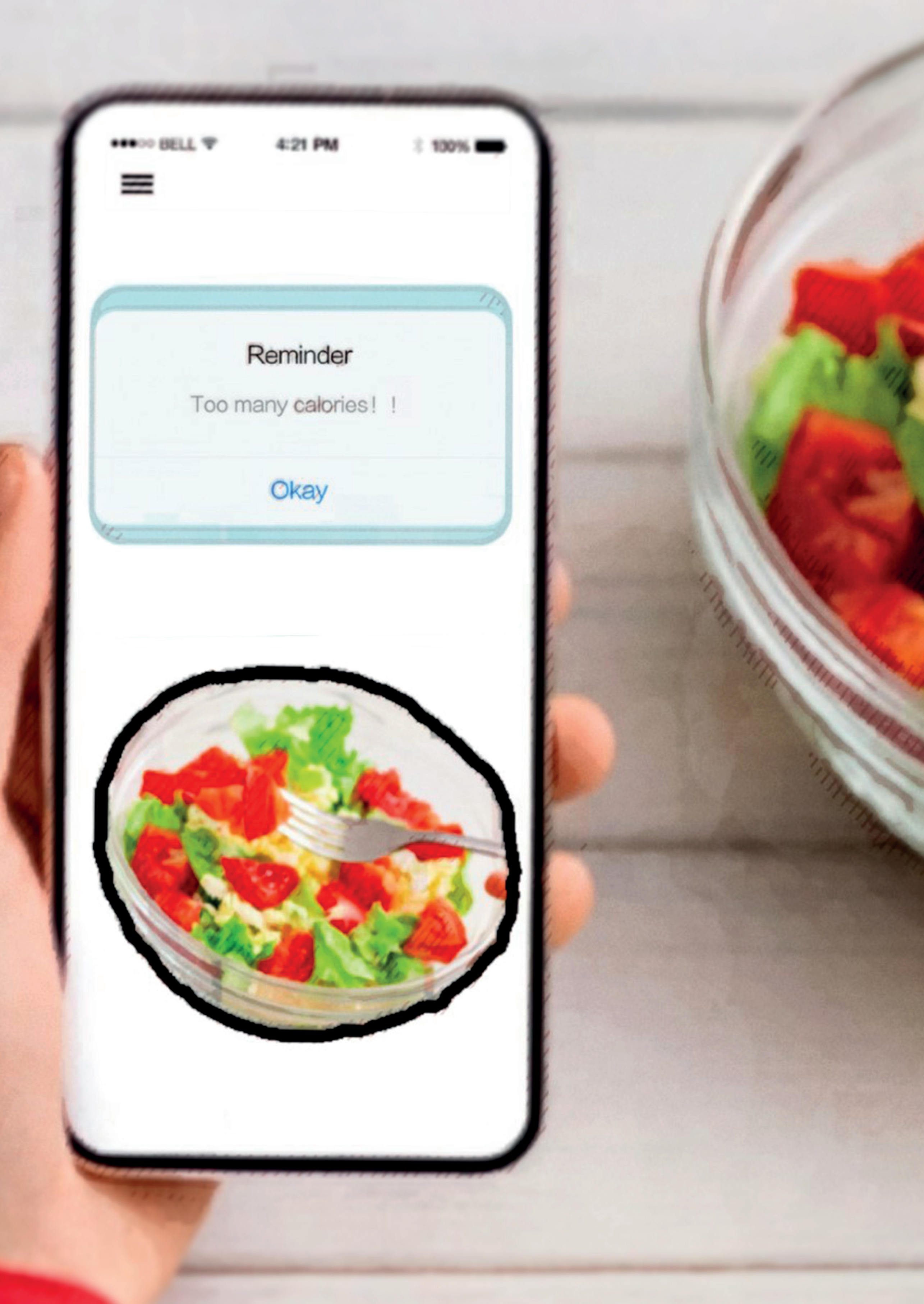
Table S10. Vitamin B12 calculation for three recipes

		Raw weight in recipe (g)	VB12 ug in recipe(NEVO)	VB12 ug in recipe (Mynetdiary)	VB12 ug in recipe (Calories!)	VB12 ug in recipe(Virtue)	First and second cooking procedure	First and second retention factor	VB12 ug in one portion (NEVO)
Stamppot	kale	250.00	0.00	0.00	-	0.00	stew	-	0.00
	potatoes	312.50	0.00	0.00	-	0.00	stew	-	0.00
	butter	10.00	0.03	0.02	-	0.02		1.00	0.03
	semi-skimmed milk	18.80	0.08	0.41	-	0.17	boil	0.90	0.08
	Total		0.11	0.43	-	0.19			0.11
Pizza	Flour	50.00	0.00	0.00	-	0.00	bake in oven	1.00	0.00
	Olive oil	4.50	0.00	0.00	-	0.00	bake in oven	-	0.00
	Tomato puree	11.25	0.00	0.00	-	0.00	bake in oven	0.70	0.00
	Yeast	0.88	0.00	0.00	-	0.00	bake in oven	-	0.00
	Oregano	0.38	0.00	0.00	-	0.00	bake in oven	-	0.00
	Mature cheese 48+	37.50	0.75	0.30	-	0.31	bake in oven	0.90	0.68
	Tomato	25.00	0.00	0.00	-	0.00	bake in oven	0.70	0.00
	Salami	25.00	0.35	0.70	-	0.70	bake in oven	0.95	0.33
	Mushroom	25.00	0.00	0.00	-	0.01	bake in oven	0.70	0.00
	Total		1.10	1.00	-	1.02			1.01
Hachee	Hachee meat	100.00	2.91	2.53	-	2.66	fry in pan and stew	0.70 * 0.70	1.43
	Onion	100.00	0.00	0.00	-	0.00	fry in pan and stew	0.70 * 0.70	0.00
	Vinegar	4.50	-	0.00	-	0.00	stew	1.00	0.00
	Flour	7.50	0.00	0.00	-	0.00	stew	0.95	0.00
	Pepper	1.25	0.00	0.00	-	0.00	stew	1.00	0.00
	Butter	12.50	0.04	0.03	-	0.02	stew	1.00	0.04
	Laurel	1.25	0.00	0.00	-	0.00	stew	1.00	0.00
	Cloves	1.25	0.00	0.00	-	0.00	stew	1.00	0.00
	Bouillon powder	1.25	0.00	0.14	-	0.00	stew	1.00	0.00
	Total		2.95	2.70	-	2.69			1.46

4

Table S11. Folate calculation for three recipes

		Raw weight in recipe (g)	Folate ug in recipe(NEVO)	Folate ug in recipe (Mynetdiary)	Folate ug in recipe (Calories!)	Folate ug in recipe(Virtue)	First and second cooking procedure	First and second retention factor	Folate ug in one portion (NEVO)
Stamppot	kale	250.00	125.00	350.00	0.00	0.00	stew	0.70	87.50
	potatoes	312.50	71.88	56.25	93.75	0.00	stew	0.75	53.91
	butter	10.00	0.00	0.30	0.30	0.00		-	0.00
	semi-skimmed milk	18.80	1.22	0.00	0.00	0.00	boil	0.50	0.61
	Total		198.10	406.55	94.05	0.00			142.02
Pizza	Flour	50.00	27.00	91.50	8.00	77.00	bake in oven	0.50	13.50
	Olive oil	4.50	0.00	0.00	0.00	0.00	bake in oven	-	0.00
	Tomato puree	11.25	4.64	1.24	3.68	0.00	bake in oven	0.70	3.24
	Yeast	0.88	35.00	20.48	8.22	0.00	bake in oven	-	35.00
	Oregano	0.38	0.00	0.89	0.00	0.00	bake in oven	-	0.00
	Mature cheese 48+	37.50	9.38	6.75	12.38	0.00	bake in oven	0.50	4.69
	Tomato	25.00	3.93	3.75	9.75	0.00	bake in oven	0.70	2.75
	Salami	25.00	0.63	0.50	0.50	0.00	bake in oven	0.80	0.50
	Mushroom	25.00	11.00	4.25	2.00	0.00	bake in oven	0.70	7.70
	Total		91.56	129.35	44.52	77.00			67.38
Hachee	Hachee meat	100.00	2.90	3.30	13.00	0.00	fry in pan and stew	0.80 * 0.80	1.86
	Onion	100.00	23.70	20.90	14.60	0.00	fry in pan and stew	0.70 * 0.70	11.61
	Vinegar	4.50	0.00	0.00	0.00	0.00	stew	1.00	0.00
	Flour	7.50	1.43	28.60	0.80	12.71	stew	0.80	1.14
	Pepper	1.25	0.00	0.23	0.00	0.00	stew	1.00	0.00
	Butter	12.50	0.00	0.41	0.40	0.00	stew	1.00	0.00
	Laurel	1.25	0.00	2.48	0.00	0.00	stew	1.00	0.00
	Cloves	1.25	0.00	0.34	0.00	0.00	stew	1.00	0.00
	Bouillon powder	1.25	0.00	0.44	0.00	0.00	stew	1.00	0.00
	Total		28.03	56.70	28.80	12.71			14.61



BELL
4:21 PM
100%
Reminder
Too many calories! !
Okay

Chapter 5

A Systematic Review and Meta-analysis of Validation Studies Performed on Food Record Apps

Liangzi Zhang, Andreja Misir, Hendriek Boshuizen, Marga Ocké

Submitted for publication

Abstract: Mobile food record apps have been increasingly validated by studies with various study designs. This review aims to evaluate the overall accuracy of FR apps in measuring the intake of energy, macro- and micronutrients, food groups in real-life settings and to provide a summary of the study designs used in these studies. We systematically searched online databases for mobile FR validation studies published during 2013-2019. We identified 14 studies for the systematic review, of which 11 studies were suitable for meta-analyses on energy intake and eight for meta-analysis on macronutrient intake. Mean differences and SDs for each outcome were pooled using a random-effects model. All apps underestimated energy intake when compared to their reference methods with a pooled effect of -202 kcal (-319 to -85 kcal). After stratification, studies which used the same food composition tables for both the app and the reference method had no heterogeneity with a pooled effect of -57 kcal (-116 to 2 kcal). In eight studies that investigated macronutrient intake, after excluding outliers, the heterogeneity of carbohydrate, fat, and protein was 54%, 73% and 80%, with the pooled effect of -18.8 g/day, -12.7 g/day, and -12.2 g/day respectively. Micronutrients from six studies and food groups from four studies were - mostly statistically insignificantly - underestimated by the apps. Alcohol was significantly overestimated by one app while significantly underestimated by another app. This review concluded that FR apps seem to underreport dietary intake slightly more than traditional dietary assessment methods. Better quality validation studies should be conducted in the future, i.e. by applying biomarkers as the reference; testing in larger and more representative study populations for longer periods; avoiding learning effect of each method; comparing food groups and micronutrients with both raw data and adjusted values.

Introduction

Diet has been recognized as one of the determinants for developing non-communicable diseases such as cardiovascular disease, diabetes, and cancer (1). An accurate assessment of dietary intake is fundamental for carrying out nutritional studies (2). Self-reported dietary intake is the most commonly used method in large scale nutritional studies, which could assess all food and nutrients and has a better trade-off between cost, response and accuracy than objective measures (e.g., biomarkers) (3). However, self-reported intake may be subject to response error (inaccurate recall, under- and overreporting) and portion size error (inaccurate portion size assessment) (4, 5). Retrospective methods such as 24-hour recall (24HR) are subject to memory loss, while prospective methods such as food records are subject to reactivity bias (6), but are better in estimating portion sizes (7).

Due to the error-prone nature and burdensome procedures in available dietary assessment methods, technology advancement has favored the use of digital applications in assessing dietary intake in large-scale studies (8-10). Most interesting is mobile phone ownership that has grown exponentially in the past two decades, providing a convenient platform for recording dietary intake (11). Specifically, mobile applications were constructed based on the theory of traditional dietary assessment methods are among the main instruments investigated in nutritional studies nowadays (11). Most mobile dietary apps have an underlying mechanism of food records, due to the portable nature of smartphones, and the ability to incorporate real-time recording features like barcode and photo recognition to assist in food searching and portion size estimation (12).

New methods (and technologies) need to be validated to ensure accuracy in estimating dietary intake before being applied in large-scale research. Validation studies assess the degree to which a new method measures what it is intending to measure by comparing with a reference method (13). The reference method should have a higher degree of demonstrated validity and have uncorrelated errors with the test method (14). Currently, most research-based apps have been validated with a well-established dietary assessment method, while only few commercial apps have been validated (8).

The quality of existing validation studies depends on the resources and methodologies that researchers can access (8). There are no recent reviews on the results of validation studies that specifically focused on food record apps. A review study by Sharp et al. focused on evaluating the validity, feasibility, and acceptability of a broader range of technologies, including both dietary apps and image-based technologies. They concluded that these technologies showed similar, but not superior validity when compared with conventional methods (9). It is expected that after this review, which dates from 2014, many new apps were developed and validated. Apart from reviewing the new evidence from these validation studies, a meta-analysis on results across different validation studies, along with a critical evaluation of the study designs, could provide more information on the accuracy of using food record apps in real-life situations.

Thus, this systematic review aims to evaluate the current state of the overall accuracy of recent mobile phone dietary apps in estimating the intake of energy, macronutrients, micronutrients, and food groups, using a meta-analysis if applicable. Also, it aims to review the applied designs and methodological aspects of validation studies on mobile phone food record apps.

Methods

Studies published in English were identified from the online databases Web of Science, Medline, and PubMed, using the following search strategy from Jan. 1st, 2013 to Oct. 31st, 2019: [("smartphone" OR "phone" OR "telephone" OR "mobile" OR "app" OR "mobile app*") AND ("diet* record" OR "dietary assessment" OR " food intake" OR "dietary measurement" OR "energy intake" OR "caloric intake" OR "nutrient intake" OR "nutrition assessment" OR "diet tracking" OR "food tracking") AND ("valid*" OR "accuracy" OR "compar*" OR "evaluat*") in abstract or title]. We also scrutinized citations from already detected studies and review articles.

Study identification and data extraction

Studies were potentially eligible for inclusion in this systematic review if they satisfied all of the following criteria: (1) exclusively self-reported mobile phone apps that simulate food records; (2) included a validation that compares the app to an objective method (e.g. biomarker or accelerometer) or with a reference dietary assessment method (e.g. 24HR, FFQ, etc.); (3) studies with a "real life" setting (a sample of participants entering all consumptions they consumed on a day in a free-living situation); (4) Validation studies covering any segments of the global population and all genders. Two researchers (AM, LZ) performed study screening independently and blinded in the web application Rayyan (15). After the first screening looking at titles and abstracts, agreement on the list of selected papers was reached between the reviewers. Full articles were then retrieved and were further assessed for eligibility, independently and blinded, by the two researchers. The final decision on the inclusion of studies was based on a consensus between the two researchers and discussed with MO (supervisor), if necessary. This systematic review protocol was developed following the Preferred Reporting Items for Systematic Reviews (PRISMA) statement (16).

The features and results of each validation study were extracted consecutively by two researchers (AM extracted the data, and LZ checked the data for accuracy and vice versa). General characteristics of the validation studies, such as the type of reference method, the choice of a timeframe, the sequence and spacing of test and reference methods, the selection and the number of subjects, and the applied statistical tests were extracted. Mean differences in energy and macronutrient intake were extracted between the test method (app) and the reference method for further meta-analysis. Energy intake was transformed into kcal if it was only available in kilojoules. For studies in which multiple days were compared, only the average of the total period or only data where the number of participants satisfied the power calculation for studies was taken into account (e.g., Chen et al.). The correlation coefficients (Pearson r and Spearman rho) and limits of agreement (LOA) were collected where available. The correlation coefficients were categorized based on Chan (17) and Akoglu (18) into strong if $r \geq 0.80$, moderate if $0.60 \leq r < 0.80$, fair if $0.30 \leq r < 0.60$, poor if $r < 0.30$. For studies where other nutrients and food groups were measured, correlation coefficients and under- or overreporting between the app and the reference methods are presented.

Meta-Analysis

The meta-analysis of energy and macronutrients was performed on studies that had enough uniformity of available data for the dietary component under analysis. Studies were included for meta-analysis if they presented a mean and standard deviation for the app and the reference method (so-called raw effect size data that was most consistent between reviewed studies), and their units for macronutrient were in grams. Pooled mean differences (and 95% confidence intervals) between the app and the reference method were calculated using Hartung-Knapp-Sidik-Jonkman (HKSJ) random effect model. HKSJ has fewer false positives with a small number of studies than the more common DerSimonian-Laird estimator (19). X^2 test (20) at the significance level of $p<0.05$ was performed with the I^2 statistic, in which cut-offs in between 25% to 50%, 50% to 75% and more than 75% indicate low, moderate, and high heterogeneity, respectively (21).

When the test showed significant heterogeneity, the sources of heterogeneity were explored with a stratification analysis by two characteristics of the validation study, i.e., the reference method used in the study and whether the same food composition table was used in the app and the reference method. Stratification was performed only on the validation of dietary components if the number of validation studies was ten or more.

Sensitivity analyses were conducted to examine the impact of outlier studies. The outliers were identified by: first, if the individual study's confidence interval did not overlap with the confidence interval of the pooled effect. Second, the Graphic Display of Heterogeneity (Gosh) Plot method was used to detect potential outliers, in case there were borderline studies that nearly non-overlapping with pooled confidence intervals (22). The test could detect studies which might potentially contribute to the heterogeneity. Sensitivity analysis was performed for the intake of both energy and macronutrients by omitting the outlier study.

In the case of 10 or more contributing studies, the potential of publication bias was analyzed with Egger's test (23) for publication bias. Data were analyzed with the statistical program R-Studio® ver.1.2.5019, R® ver. 3.6.1., R packages used include meta, metaphor, esc, and dmetar.

Results

The database searches yielded 825 publications when search results were combined, and two additional articles were identified through other sources (search alerts in searched databases). After duplicate records were removed, the title and abstract of 582 studies were screened, which resulted in the exclusion of 518 studies. After applying inclusion and exclusion criteria, 14 studies were selected for the systematic review, of which 11 studies were selected for meta-analysis on energy intake, and eight studies were selected for meta-analysis on macronutrient intake (see Figure 1).

Table 1 shows different app characteristics and design aspects regarding each validation study. The 14 studies focussed on 12 different apps, of which 7 provided feedback on nutrient intake (24-32) and 5 others did not (12, 33-36). Most validation studies included young adults as their sample population or advertised in a university setting, while two studies explicitly mentioned to include a wider age range of participants (26, 35). Most validation studies had a medium to small sample size (from 18 to 81 participants), while two studies had a larger sample size of 362 and 189 participants (26, 33). The period of app use ranged mostly from 2 to 7 days and contained at least one weekend day for most studies, while two studies asked participants to record every day for three months (24, 26). The app use was on non-consecutive days for three of the studies (12, 27, 34). Ten studies used 24HR as the only reference method for two days (n=6) (24, 25, 29, 32, 35, 36) or three days (n=4) (30, 31, 33, 34). One study used a food frequency questionnaire (FFQ) (26), one study used food records (27), two studies used an accelerometer (to measure energy expenditure) (12, 28), and one study used a combination of accelerometer, 24HR, and food records (32). Among studies with different days of the app and the reference method, most studies compared the mean of each method averaged across all corresponding days (24, 30-32, 36). Apart from two studies using accelerometers exclusively (12, 28), five studies used different food composition databases (FCDs) for the app and the reference method (24-27, 33), and seven used the same FCD. Ten studies investigated the energy and macronutrient intake, while six of them also compared micronutrient intake (24, 26, 30, 33, 34, 36). Four studies looked at food group intakes (31, 32, 34, 35). In terms of statistical parameters and tests, the frequency of using pair t-test was the highest (n=12), followed by correlation coefficient (n=11) and Bland-Altman limits of agreement (n=11), Thirteen studies used at least two statistical parameters, eight studies used all three parameters in their studies, while Lee only used the t-test (24).

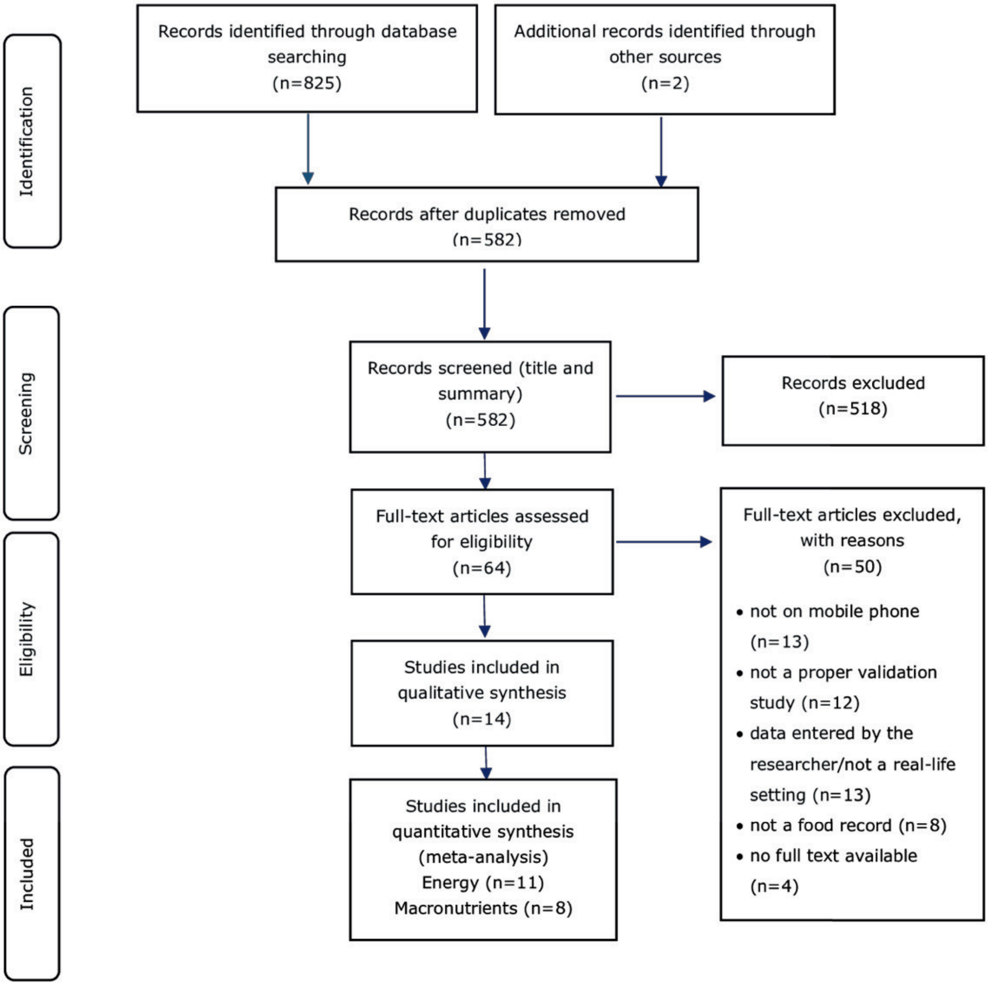


Figure 1. PRISMA Flow Diagram indicating the number of articles included at each phase

Table 1. General characteristics of the 14 food diary apps and their validation studies.

First author, year App	Country n	Study Sample	App days	Reference method days	Feedback on nutrients?	Same FCD in two methods?	Energy	Macronutrients[4]	Micronutrients	Food groups	Significance test	Correlation	LOA
Lee, 2017 **Diet-A (1)**	Korea n=21	High school students	Every day in 3 months	2x24HR:1 pre-app, 1 post-app	Yes	No	√	5	3		√		
Chen, 2019 **MyFitnessPal (MFP)[1] (2)**	Australia n=45[2]	University students and staff	4 consecutive days (including weekend)	2X24HR:unannounced with app	Yes	No	√	4			√	√	√
Recio-Rodriguez, 2019 **EVIDENT II (3)**	Spain n=362[2]	Adult (age:18-70)	Every day in 3 months	FFQ: before the interview	Yes	No	√	9	20		√	√	√
Wellard-Cole, 2019 **Eat and Track (EaT) (4)**	Australia n=189	Young adults	3 consecutive days, starting days staggered across the population (some include weekend)	3x24HR:one day after the app	No	No	√	6	1		√	√	√
Teixeira, 2018 **MyFitnessPal (MFP)[1] (5)**	Brazil n=30	University students	2 non-consecutive days (at the end of a weekday and a weekend)	Food Record (paper):at the time of the consumption	Yes	No	√	5			√	√	√
Pendergast, 2017 **FoodNow (6)**	Australia n=56	Young adults	4 non-consecutive days (1 weekend)	Accelerometer:7 days with app	No	-	√					√	√
Svensson, 2015 **no name (7)**	Sweden n=81[2]	Adolescents	3 consecutive days (some include weekend)	Accelerometer: with the app	Yes	-	√				√	√	√

First author, year App	Country n	Study Sample	App days	Reference method days	Feedback on nutrients?	Same FCD in two methods?	Energy	Macronutrients[4]	Micronutrients	Food groups	Significance test	Correlation	LOA
Mescoloto, 2017 **Nutrabem[1] (8)**	Brazil n=40[2]	University students	3 non-consecutive days (1 weekend); two weeks interval in between	3x24HR:one day after the app	No	Yes	√	4	3	12	√	√	
Bucher Della Torre, 2017 e-CA (9)	Switzerland n=18	Adults (age:20-60)	5 consecutive days including at least 1 weekend	2X24HR:unannounced with app	No	Yes	√	4		2	√		√
Ambrosini, 2018 **Research Food Diary (RFD)[1] (10)**	Australia n=50	University students and staff	4 consecutive days (including weekend)	2X24HR:one with app, one on weekend within 7 days of app use	No	Yes	√	8	2			√	√
Carter, 2013 **My Meal Mate[1] (11)**	UK n=41	University students and staff	7 consecutive days (including weekend)	2X24HR:unannounced with app	Yes	Yes	√	4			√	√	√
Rangan, 2015 **e-DIA (12)**	Australia n=80	University students	5 consecutive days (3 week days, 2 weekend days)	3x24HR:unannounced with app	Yes[3]	Yes	√	10	14		√	√	√
Rangan, 2016 **e-DIA (13)**	Australia n=80	University students	5 consecutive days (3 week days, 2 weekend days)	3x24HR:unannounced with app	Yes[3]	Yes				8	√	√	√
Lozano-Lozano, 2018 **BENECA (14)**	Spain n=20[2]	Breast cancer survivors	6 consecutive days (including weekend)	Accelerometer:8 days with app 2x24HR:unannounced with app 4xFR:with app	Yes	Yes		1		1	√		

[1] App that can be downloaded from Apple/Google store.
[2] With power analysis.
[3] Nutrients feedback deleted before doing the 24HR.
[4] Includes subgroups of macronutrients, such as saturated fat, fibre, sugar, etc.

5

Meta-analysis was performed on 11 studies for energy intake and eight studies for macronutrient intake. Figure 2A shows the pooling of the mean difference in energy. All apps underreported mean energy intake when compared to the reference method with a pooled effect of -202 kcal (95% CI:-319 to -85 kcal). Heterogeneity expressed as I^2 was 72%, which fell into the upper-moderate to high heterogeneity group. Stratification was first performed between the eight studies that used 24HR as a reference method and the three studies that used all “other” reference methods. In the 24HR group, a lowered pooled mean difference of -186 kcal (95% CI: -334 to -37 kcal) was found, with a lowered heterogeneity (I^2 = 59%). Then stratification was performed on 12 studies that either used “the same” or “different” FCDs for the app and the reference method. The pooled mean difference in the group of studies with the same FCD decreased to -57 kcal (95% CI: -116 to 2 kcal), the heterogeneity dropped to 0%. Heterogeneity was also explored with sensitivity analysis to exclude outlying studies. No outliers were detected by looking at the overlapping of confidence intervals (CIs) of each study with the pooled effect. Using the Gosh Plots method the EVIDENT II app (26) was detected as an outlier. The pooled effect dropped to -171 kcal (95% CI: -288 to -54 kcal), and the heterogeneity dropped to I^2 = 52% after deleting the outlier. Egger’s test (p = 0.17) indicated no evidence of study bias.

The pooling of the effect sizes on carbohydrate, fat, and protein intake was performed on eight studies (see Figure 2B, 2C, 2D). The pooled effects were negative for all three macronutrients. High heterogeneity of carbohydrate (I^2 = 86%) and protein (I^2 = 80%) was found, with the pooled effect of -26.9 g/day and -12.2 g/day, respectively. Similar to energy, the EVIDENT II app was detected as an outlier for carbohydrate (26). After deleting the data of the outlier, the heterogeneity dropped to moderate for carbohydrate (I^2 = 54%), with the pooled effect of -18.8 g/day. The heterogeneity of fat was slightly lower than carbohydrate and protein (I^2 = 73%), with a pooled effect of -12.7 g/day. In all eight studies, the app underreported mean fat intake when compared to the reference method.

When looked at the performance of each app, e-DIA had a relatively lower mean difference and variance in the intake of energy and all macronutrients than other apps (30). The app e-CA had the lowest mean difference for both carbohydrate and protein (35). However, the standard deviation of the differences was the highest among all studies for energy, carbohydrate, and fat. Diet-A and MFP (Chen) had the highest mean difference across the energy, fat and protein (24, 25). Together with EVIDENT II app, which is the outlier for energy and carbohydrate, these three studies used different FCD for the app and the reference methods.

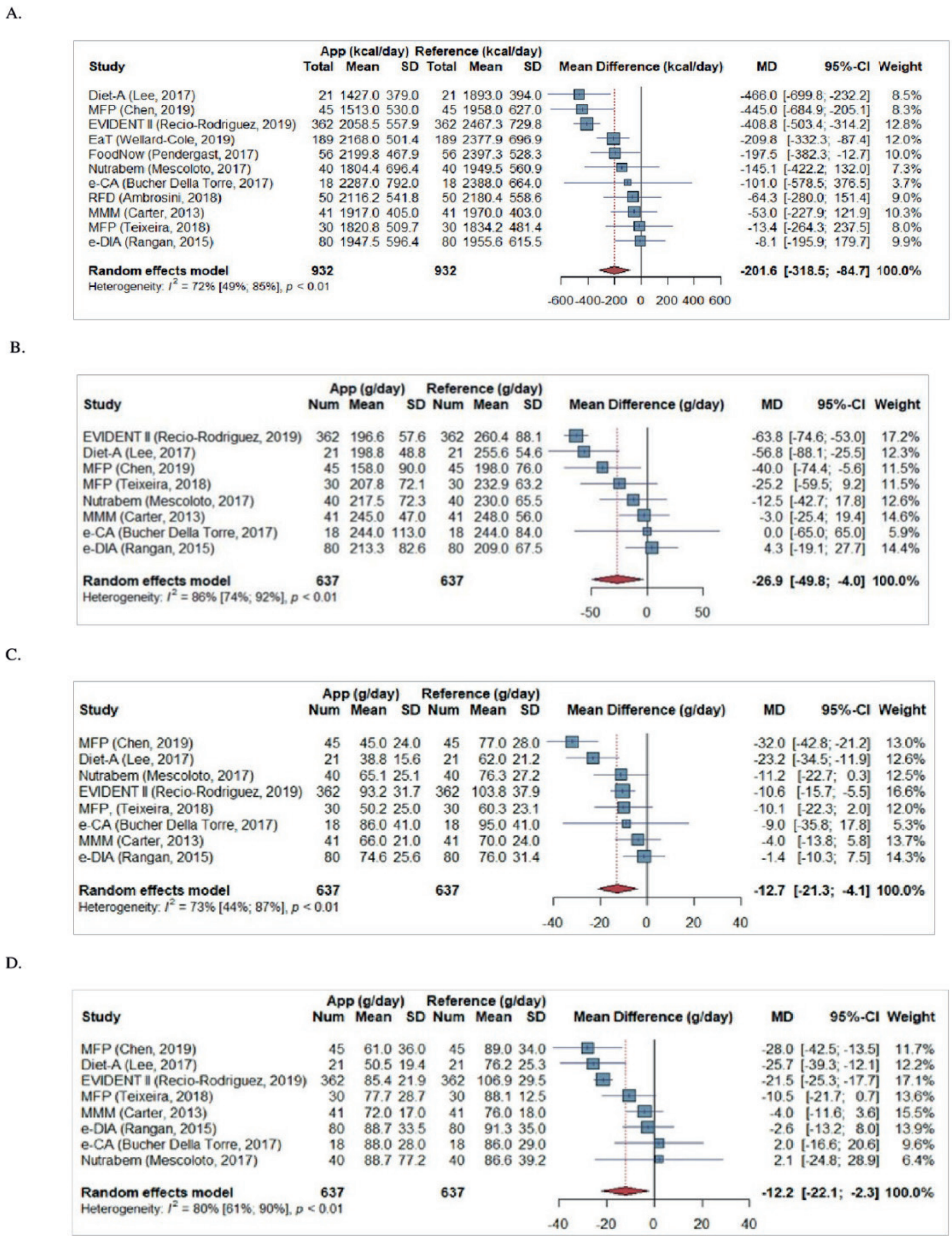


Figure 2. Forest plot for the mean difference in energy and macronutrient intake between the app and the reference method in included validation studies. A. Energy, B. Carbohydrate, C. Fat, D. Protein.

5

Table 2 illustrates the correlation coefficient and limits of agreement (LOA) between the apps and the reference methods for the intake of energy and macronutrients. The column with LOA represents the distance between the upper and the lower limit. Five studies reported both correlation and LOA for energy and all macronutrients. For energy, the three studies that had a weak correlation between two methods, had larger LOAs than other studies (25, 26, 28). Most studies had a moderate correlation with a range of 0.60 to 0.80. The distances of LOA were mostly within 2000 kcal, with one exception of 2223 kcals. Nutrabem had the highest correlation for energy, carbohydrate, and protein (34). MMM had the highest correlation in fat (29). The app e-Dia had similar correlations for energy and all macronutrients from 0.64 to 0.79 (30). EVIDENT II had weak correlations for all macronutrients and energy (26). The average correlation across studies was 0.54 to 0.60, energy and fat intake were both the lowest at 0.54. The average across energy and macronutrients in each study ranged from 0.23 to 0.78, with majority studies in the moderate category. The expression of macronutrient intake differed between studies, with grams, energy percentages, and natural logarithms.

Tables 3 lists other nutrients that were most commonly assessed in the included studies. In most studies, the app underestimated nutrient intakes. Calcium and sodium intake in Diet-A, fiber, and alcohol in EVIDENT II were significantly underestimated while the rest of the underestimated nutrients were all non-significant. Alcohol intake was significantly overestimated in RFD. Rangan compared all nutrients in this table and had the second-highest average correlation among the nutrients, while EVIDENT II had the lowest average correlation across most nutrients, except alcohol. EaT had the highest average correlation among the included nutrients, mainly due to the strong correlation for sugar intake.

Food groups were only validated for four apps (e-CA, Nutrabem, BENECA, e-Dia). A different categorization of food groups was found across studies, differences in dairy, fruits, vegetables, meat, and grain intake, were most commonly reported. Food group intakes were mostly insignificantly underestimated by apps. In the BENECA-app vegetables and fruits were mostly forgotten by participants. Among studies investigated correlations, the highest correlation found for Nutrabem-app was poultry (r=0.85) and lowest in nuts (r=0.31) and vegetable oils (r=0.37). The app e-DIA had relatively stronger correlations among all included food groups, from 0.75 to 0.88, and has an equal number of under- and overestimations.

Table 2. Summary of the correlation coefficients and limits of agreement for energy and macronutrient comparisons between apps and reference methods.

First author, year	App	n	Energy		CHO		Fat		Protein		Mean r (energy & macronutrients)
			r^2	LOA[1]	r	LOA	r	LOA	r	LOA	
Chen, 2019 (1)	MFP	45	0.29	2727 kcal	0.41	357g	0.16	131g	0.43	136g	0.32
Recio-Rodriguez, 2019 (2)	EVIDENT II	362	0.23	3263 kcal	0.27	-	0.23	-	0.20	-	0.23
Wellard-Cole, 2019 (3)	EaT	189	0.67	2223 kcal	0.79	21%	0.56	25%	0.73	14%	0.69
Teixeira, 2018 (4)	MFP	30	0.67	1345 kcal	0.41	123g[3]	0.58	67g[3]	0.43	32g	0.52
Pendergast, 2017 (5)	FoodNow	56	0.75	1383 kcal	-	-	-	-	-	-	-
Svensson, 2015 (6)	no name	81	0.13	2639 kcal	-	-	-	-	-	-	-
Mescoloto, 2017 (7)	Nutrabem	40	0.77	-	0.82	-	0.71	-	0.83	-	0.78
Bucher Della Torre, 2017 (8)	e-CA	18	-	447 kcal	-	266g	-	104g	-	75g	-
Ambrosini, 2018 (9)	RFD	50	0.52	2126 kcal	0.72	23%	0.63	22%	0.79	12%	0.67
Carter, 2013 (10)	MMM	41	0.68	1065 kcal	0.57	-	0.75	-	0.57	-	0.64
Rangan, 2015 (11)	e-Dia	80	0.66	1965 kcal	0.64	274g	0.68	92g	0.79	88g	0.69
	Average		0.54	1918 kcal	0.58		0.54		0.60		

[1] LOA, limits of agreement.

[2] r: Pearson/Spearman's correlation, strong if $r \geq 0.80$, moderate if $0.60 \leq r < 0.80$, fair if $0.30 \leq r < 0.60$, poor if $r < 0.30$.

[3] Back-transformed value.

5

Table 3. Summary of the under-/overestimation of apps and correlation coefficients with the reference methods for macronutrient subgroups and micronutrients.

First author, year	App, n=	Compare with reference	Calcium	Iron	Sodium	Vitamin C	Saturated fat	Sugar	Fibre	Alcohol	Mean r
Lee, 2017 (12)	Diet-A n= 21	Under- or over-reporting	Under *	Under	Under *	-	Under	-	-		-
		r[1]	-	-	-	-	-	-	-		-
Recio-Rodriguez, 2019 (2)	EVIDENT II n=362	Under- or over-reporting	-	-	-	-	Under	-	Under *	Under *	
		r	0.32	0.26	0.15	0.32	0.31	-	0.31	0.68	0.34
Wellard-Cole, 2019 (3)	EaT n=189	Under- or over-reporting	-	-	Under	-	Under	Under	-	-	
		r	-	-	0.56	-	0.59	0.82	-	-	0.66
Mescoloto, 2017 (7)	Nutrabem n=40	Under- or over-reporting	Under	Under	-	Under	-	-	-	-	
		r	0.57	0.66	-	0.6	-	-	-	-	0.61
Ambrosini, 2018 (9)	RFD n=50	Under- or over-reporting	Over	Over	-	-	Under	Under	Under	Over *	
		r	0.45	0.42	-	-	0.60	0.68	0.66	0.65	0.58
Rangan, 2015 (11)	e-Dia n=80	Under- or over-reporting	Under	Over	Under	Under	Under	Under	Over	Over	
		r	0.75	0.57	0.60	0.68	0.75	0.56	0.54	0.77	0.65

[1] r: Pearson/Spearman's correlation, strong if $r \geq 0.80$, moderate if $0.60 \leq r < 0.80$, fair if $0.30 \leq r < 0.60$, poor if $r < 0.30$.
* Significant estimation

Discussion

This paper aimed to assess the overall accuracy of dietary intake measurements in validated mobile phone food record apps. Apps from more than half of the 14 included studies were validated in university settings, were small scale with a duration of 2 to 7 consecutive days, used 24HR as the reference method, and used the same FCDs for the test and the reference method. The meta-analysis on results for 8-10 apps found that food record apps underreported energy and macronutrients relative to classical dietary assessment methods. Moderate heterogeneity was reached when an outlier study was excluded from the meta-analysis for energy and carbohydrate. Studies using the same food composition database for the apps and the reference methods had no heterogeneity for energy intake and had a lowered pooled effect of -57 kcal. Studies that observed smaller differences in energy intake between the app and the reference method also had smaller differences in macro- and/or micronutrients and food groups.

Intentional/Unintentional Underreporting

Underreporting of energy intake in the app compared to the reference method was found in all studies. An even larger extent of under-reporting was expected for studies that used an objective reference method as the reference, because underreporting is also often observed in the 24HR (8), which most studies have used. The tendency of underreporting when using the app or other self-reporting methods may either be unintentional and intentional (11). The effect of unintentional underreporting could potentially be alleviated by adding adequate prompts and improving technological add-ins (36). Intentional underreporting is more challenging to eliminate when participants deliberately omit the input of certain foods out of social acceptability or convenience or temporarily change their eating behaviour (37). In the current study, a larger extent of underestimation in carbohydrate and fat intake was found as compared to protein, which is in line with the findings from another review on a technology-based dietary assessment tool by Eldridge et al. (8). Bucher Della Torre et al. and Chen et al. found that people tend to underreport fat, alcohol, discretionary foods and beverages (high in fat/sugar) intake unless prompted by interviewers (25, 35), while Rangan et al. indicated the underreporting of added sugar and alcohol might be due to intentional underreporting of foods containing added sugars or the reduced alcohol or sugar intake while using the app (30).

Approximately half of the errors in energy intake estimations from dietary records administered on technological devices have been attributed to wrong portion size estimations (38). Participants were asked to refer to a provided food model booklet to assist with the estimation of portion sizes during 24HRs, while most apps provide metric weights (e.g., g, mL) or household measure options (e.g., cups) with no images accompanied (39). Bucher

Della Torre et al. found that participants tended to choose the app proposed portions even if their real portions are different, especially with drinks (35). Mobile technologies with the assistance of digital photographs have shown less extent of underestimation than regular food records in a free-living situation compared to doubly labelled water (DLW) (40-42). These studies were not included in the current review because they were not exclusively self-reporting, and required a large involvement of dietitians to identify foods and amounts from photos correctly. Automatic food recognition and volume estimation could potentially outperform portion sizes estimated by individuals, but validations are needed to verify their applicability in large-scale studies (43).

Some studies conducted the 24HR the next day of using the app, which might have caused a memory effect and lessened the recalling bias of 24HR (29). Besides, access to the nutrient feedbacks from some apps could enhance health consciousness and induce changes in the food intake of those who are motivated for weight reduction (44, 45). Both study designs could lead to an increased agreement and augmented correlation between the two methods (35). The learning effect could be reduced if the app and the reference methods are used on separate days, with the app used first (46). Differences due to day-to-day variation in the data could be evened out with repeated measurements or corrected with statistical modelling (47). However, the source of variation (e.g., food omission, portion underestimation) could be investigated better if both methods were conducted on the same day. Moreover, unannounced 24HR is preferred to avoid behavioural change (48). Ambrosini et al. conducted the second 24HR unannounced on a different day within seven days of app use (36). In this way, both the app and reference method are measuring dietary intake to a similar extent while limiting the possible influence of each method.

Explanations on High Heterogeneity

We observed a higher mean difference in studies where different FCDs were incorporated into the app and the reference method. In studies where the same food items are entered by researchers into different apps, disagreements between apps is mainly due to the different FCDs embedded in each app (8, 49-53). Thus, the "human components", that were mainly accounted for in validation studies of methods rather than nutrient content, should be distinguished from different FCD use. If using the same FCD is impractical, comparing differences in food groups or food items between two methods could be a solution. Moreover, insight in validity of food groups can give some clues on specific foods that are easily forgotten, like the fat used for frying. Besides, advocacy to move from nutrient focus towards food-based research in nutrition epidemiology has stressed the importance of food group validation using new methods (54). Unfortunately, only four of the included studies validated food groups, and none of the studies that used different FCDs have considered comparing food groups. Moreover, studies with food group comparisons used different food categorizations and statistical tests, which limited the comparisons of food group differences across studies.

Our results indicated that the choice of the reference method was also one of the determining factors for heterogeneity. The absolute validity was not reported in smartphone application validations, possibly due to the high cost associated with recovery biomarkers and the availability for limited nutrients. When investigating the relative validity of a method it is desirable to use a reference method with uncorrelated errors and better accuracy, for example, comparing food records with 24HR. One study in the meta-analysis used FFQ as the reference, which has a lower level of accuracy with a limited frequency of consumption options and food lists in the FFQ tool (55). Furthermore, FFQs estimate nutritional intake over a longer time period (usual consumption) while more diverse food item options are influenced by seasonality of different foods. Conversely, Teixeira et al. tested their app with a paper-based food record measuring the food consumption of the same days. Here an overestimation of correlation was expected because two methods share the same embedded errors (27). Two studies used an accelerometer to assess energy expenditure, which is an objective measure less burdensome than DLW (12). However, accelerators have shown over- and underestimation of energy expenditure when different types of physical activities were performed (56).

Most studies used a diverse range of statistical techniques that could facilitate a balanced interpretation of results (30). Correlation coefficients indicate the ability of the app to rank individuals and the strength of the association. Bland-Altman plots reveal the presence, direction, and extent of bias at the group level and the extent of measurement error at the individual level (57). A wide LOA found in most studies was expected because the reference measure itself might have potential errors and is not reflecting true intakes (29). Besides, only a few days of intake were collected for most studies. Rangan et al. found a smaller difference and a higher correlation with values adjusted for within-person variation. Garden et al. also found that the heterogeneity of FFQ validation studies decreased if de-attenuated/energy-adjusted values were used (58). Because the majority of studies in this review did not adjust for the nutrient intake, only studies with raw data were compared. Hence, presenting data in several ways is necessary for cross-study comparisons and in obtaining insight into different types of error, i.e., systematic and random error (59).

Limited information was provided by included studies on whether they aimed at validating current or usual dietary intake. Although a single day food intake can be useful for many studies, usual intake is of primary interest for studies on surveillance, epidemiology and intervention (60). To measure the ability of an app to capture usual intake, studies that used 24HR and FFQ as the reference should be conducted on non-consecutive days, including both weekdays and weekend days (51, 55), which might capture more variations in diet and occasionally consumed foods, such as alcohol (12). A higher reporting accuracy of food records has been found when a weekend day was included. It was speculated that participants have more time during the weekend to complete a food record (61). In half of the included studies, participants used the app for less than four days, which was not sufficient to estimate usual micronutrients intake accurately and to capture habitual diet (62, 63), especially with a sample size less than fifty (64). To be fair, the limited number of studies that investigated and compared micronutrient intake indicated that it is still too premature to get insight in validity of micronutrient intake of apps. The inclusion of mostly young adults from university settings limited the generalizability of the validation results. Furthermore, in the case where people with low technological literacy used the apps, they probably provide less reliable data (48, 63, 65).

Strength and Limitations

This study is the first meta-analyses of the validations of food record apps in free-living conditions; it provides a detailed comparison of the study design, and it includes results on micronutrients and food groups. For this study, a systematic search strategy for three electronic databases was adopted in searching for eligible papers, and we have not found no evidence of publication bias among the included studies. Still, we could not rule out the possibility that other eligible papers that are not in English or not available via electronic databases were missed. The exclusion of image-based mobile technologies (entered by dietitians) helped us to better understand the suboptimal performances among individuals using apps in naturalistic setting compared to studies entered by dietitians. The narrowed study selection criteria promoted a higher quality of reporting validity of dietary apps and allowed an easier comparison between studies. Another strength of the study was that heterogeneity on energy intake was explained by the stratification analysis, unfortunately, due to the small number of studies, testing for publication bias and exploring heterogeneity with stratification was only possible for energy intake. Moreover, the limited number of studies might lowered the power of the meta-analysis (58).

Conclusions

Food record apps underreport energy intake, as well as intake of macronutrients. No specific conclusions could be made on micronutrient and food group comparisons due to limited and incomparable data. Future validation studies should consider applying biomarkers as the reference method next to repeated 24HRs; include larger and more representative study populations, and should try to provide insight in the source of the measurement error by also looking at the validity of food groups and micronutrients.

References

1. Arens-Volland AG, Spassova L, Bohn T. Promising approaches of computer-supported dietary assessment and management-Current research status and available applications. Int J Med Inform. 2015 Dec;84(12):997-1008.
2. Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Res. 2017;6:926.
3. EFSA. General principles for the collection of national food consumption data in the view of a pan-European dietary survey. EFSA Journal. 2009;7(12):1435.
4. Hedrick VE, Dietrich AM, Estabrooks PA, Savla J, Serrano E, Davy BM. Dietary biomarkers: advances, limitations and future directions. Nutr J. 2012;11:109.
5. Shim J-S, Oh K, Kim HC. Dietary Assessment Methods in Epidemiologic Studies. Epidemiology and Health. 2014;36:e2014009.
6. Coulston AM, Boushey CJ. Nutrition in the prevention and treatment of disease. Amsterdam: Academic Press; 2008.
7. Dietary Assessment: A resource guide to method selection and application in low resource settings. Rome: FAO, 2018.
8. Eldridge AL, Piernas C, Illner AK, Gibney MJ, Gurinovic MA, de Vries JHM, et al. Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients. 2018 Dec 28;11(1).
9. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition. 2014;30(11):1257-66.
10. Allman-Farinelli M. Using digital media to measure diet. CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources. 2018;13.
11. Illner AK, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol. 2012 Aug;41(4):1187-203.
12. Pendergast FJ, Ridgers ND, Worsley A, McNaughton SA. Evaluation of a smartphone food diary application using objectively measured energy expenditure. Int J Behav Nutr Phys Act. 2017 Mar 14;14(1):30.
13. Serra-Majem L, Frost Andersen L, Henrique-Sanchez P, Doreste-Alonso J, Sanchez-Villegas A, Ortiz-Andrelluchi A, et al. Evaluating the quality of dietary intake validation studies. Br J Nutr. 2009 Dec;102 Suppl 1:S3-9.
14. Willett W, Lenart E. Reproducibility and validity of food-frequency questionnaires. Nutritional epidemiology. 2013:96-141.
15. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Systematic reviews. 2016 Dec 5;5(1):210.
16. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009 Oct;62(10):1006-12.

17. Chan YH. Biostatistics 104: correlational analysis. Singapore medical journal. 2003 Dec;44(12):614-9.
18. Akoglu H. User's guide to correlation coefficients. Turkish journal of emergency medicine. 2018 Sep;18(3):91-3.
19. Rover C, Knapp G, Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. BMC medical research methodology. 2015 Nov 14;15:99.
20. Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res. 1993;2(2):121-45.
21. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.
22. Olkin I, Dahabreh IJ, Trikalinos TA. GOSH - a graphical display of study heterogeneity. Research synthesis methods. 2012 Sep;3(3):214-23.
23. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Bmj. 1997 Sep 13;315(7109):629-34.
24. Lee JE, Song S, Ahn JS, Kim Y, Lee JE. Use of a Mobile Application for Self-Monitoring Dietary Intake: Feasibility Test and an Intervention Study. Nutrients. 2017 Jul 13;9(7).
25. Chen J, Berkman W, Bardouh M, Ng CYK, Allman-Farinelli M. The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. Nutrition. 2019;57:208-16.
26. Recio-Rodriguez JI, Rodriguez-Martin C, Gonzalez-Sanchez J, Rodriguez-Sanchez E, Martin-Borras C, Martinez-Vizcaino V, et al. EVIDENT Smartphone App, a New Method for the Dietary Record: Comparison With a Food Frequency Questionnaire. JMIR Mhealth Uhealth. 2019 Feb 8;7(2):e11463.
27. Teixeira V, Voci SM, Mendes-Netto RS, da Silva DG. The relative validity of a food record using the smartphone application MyFitnessPal. Nutr Diet. 2018 Apr;75(2):219-25.
28. Svensson A, Larsson C. A Mobile Phone App for Dietary Intake Assessment in Adolescents: An Evaluation Study. JMIR Mhealth Uhealth. 2015 Nov 3;3(4):e93.
29. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr. 2013 Feb 14;109(3):539-46.
30. Rangan AM, O'Connor S, Giannelli V, Yap MLH, Ming L, Roy R, 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 2015--2015;3(4):3-14.
31. Rangan AM, Tieleman L, Louie JC, Tang LM, Hebden L, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): relative validity of a mobile phone application to measure intake of food groups. Br J Nutr. 2016 Jun;115(12):2219-26.

32. Lozano-Lozano M, Galiano-Castillo N, Martin-Martin L, Pace-Bedetti N, Fernandez-Lao C, Arroyo-Morales M, et al. Monitoring Energy Balance in Breast Cancer Survivors Using a Mobile App: Reliability Study. JMIR mHealth and uHealth. 2018 2018--Mar;6(3):e67-e.
33. Wellard-Cole L, Chen J, Davies A, Wong A, Huynh S, Rangan A, 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 Mar 14;11(3).
34. Mescoloto SB, Caivano S, Domene SMA. Evaluation of a mobile application for estimation of food intake. Revista De Nutricao-Brazilian Journal of Nutrition. 2017 Jan-Feb;30(1):91-8.
35. 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 Jan 18;9(1).
36. Ambrosini GL, 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 Jan 9;17(1):5.
37. Scagliusi FB, Polacow VO, Artioli GG, Benatti FB, Lancha AH, Jr. Selective underreporting of energy intake in women: magnitude, determinants, and effect of training. J Am Diet Assoc. 2003 Oct;103(10):1306-13.
38. Beasley J, Riley WT, Jean-Mary J. Accuracy of a PDA-based dietary assessment program. Nutrition. 2005 Jun;21(6):672-7.
39. Subar AF, Crafts J, Zimmerman TP, Wilson M, Mittl B, Islam NG, 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 Jan;110(1):55-64.
40. Martin CK, Han HM, Coulon SM, Allen HR, Champagne CM, Anton SD. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. British Journal of Nutrition. 2009 Feb;101(3):446-56.
41. Rollo ME, Ash S, Lyons-Wall P, Russell AW. Evaluation of a mobile phone image-based dietary assessment method in adults with type 2 diabetes. Nutrients. 2015;7(6):4897-910.
42. Delisle Nystrom C, Forsum E, Henriksson H, Trolle-Lagerros Y, Larsson C, Maddison R, et al. A Mobile Phone Based Method to Assess Energy and Food Intake in Young Children: A Validation Study against the Doubly Labelled Water Method and 24 h Dietary Recalls. Nutrients. 2016 Jan 15;8(1).
43. Daugherty BL, Schap TE, Ettienne-Gittens R, Zhu FM, Bosch M, Delp EJ, et al. Novel technologies for assessing dietary intake: evaluating the usability of a mobile telephone food record among adults and adolescents. J Med Internet Res. 2012 Apr 13;14(2):e58.

44. Afshin A, Babalola D, McLean M, Yu Z, Ma W, Chen CY, et al. Information Technology and Lifestyle: A Systematic Evaluation of Internet and Mobile Interventions for Improving Diet, Physical Activity, Obesity, Tobacco, and Alcohol Use. Journal of the American Heart Association. 2016 Aug 31;5(9).
45. Forster H, Walsh MC, Gibney MJ, Brennan L, Gibney ER. Personalised nutrition: the role of new dietary assessment methods. Proc Nutr Soc. 2016 Feb;75(1):96-105.
46. Nelson M. The validation of dietary assessment. Design concepts in nutritional epidemiology. 1997;2:241-56.
47. Freedman LS, Midthune D, Dodd KW, Carroll RJ, Kipnis V. A statistical model for measurement error that incorporates variation over time in the target measure, with application to nutritional epidemiology. Stat Med. 2015 Nov 30;34(27):3590-605.
48. Walton J. Dietary Assessment Methodology for Nutritional Assessment: A Practical Approach. Top Clin Nutr. 2015;30(1):33-46 10.1097/TIN.0000000000000018.
49. Ferrara G, Kim J, Lin S, Hua J, Seto E. 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 Uhealth. 2019 May 17;7(5):e9232.
50. Conrad J, Nothlings U. Innovative approaches to estimate individual usual dietary intake in large-scale epidemiological studies. Proc Nutr Soc. 2017 Aug;76(3):213-9.
51. Franco RZ, Fallaize R, Lovegrove JA, Hwang F. Popular Nutrition-Related Mobile Apps: A Feature Assessment. JMIR Mhealth Uhealth. 2016 Aug 01;4(3):e85.
52. Chen J, Cade JE, Allman-Farinelli M. The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment. JMIR Mhealth Uhealth. 2015 Dec 16;3(4):e104.
53. Zhang L, Nawijn E, Boshuizen H, Ocke M. Evaluation of the Recipe Function in Popular Dietary Smartphone Applications, with Emphasize on Features Relevant for Nutrition Assessment in Large-Scale Studies. Nutrients. 2019 Jan 19;11(1).
54. Jacobs DR, Jr., Tapsell LC. Food, not nutrients, is the fundamental unit in nutrition. Nutr Rev. 2007 Oct;65(10):439-50.
55. Hooson J, Hutchinson J, Warthon-Medina M, Hancock N, Greathead K, Knowles B, et al. A systematic review of reviews identifying UK validated dietary assessment tools for inclusion on an interactive guided website for researchers: www.nutritools.org. Crit Rev Food Sci Nutr. 2019:1-25.
56. Bhammar DM, Sawyer BJ, Tucker WJ, Lee JM, Gaesser GA. Validity of SenseWear (R) Armband v5.2 and v2.2 for estimating energy expenditure. J Sports Sci. 2016 Oct;34(19):1830-8.
57. Lombard MJ, Steyn NP, Charlton KE, Senekal M. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J. 2015 Apr 22;14:40.

58. Tabacchi G, Filippi AR, Amodio E, Jemni M, Bianco A, Firenze A, et al. A meta-analysis of the validity of FFQ targeted to adolescents. Public Health Nutr. 2016 May;19(7):1168-83.
59. Ambrosini GL, de Klerk NH, Musk AW, Mackerras D. Agreement between a brief food frequency questionnaire and diet records using two statistical methods. Public Health Nutr. 2001 Apr;4(2):255-64.
60. Kirkpatrick SI, Baranowski T, Subar AF, Tooze JA, Frongillo EA. Best Practices for Conducting and Interpreting Studies to Validate Self-Report Dietary Assessment Methods. J Acad Nutr Diet. 2019 Nov;119(11):1801-16.
61. Svensson A, Waling M, Backlund C, Larsson C. Overweight and Obese Children's Ability to Report Energy Intake Using Digital Camera Food Records during a 2-Year Study. Journal of nutrition and metabolism. 2012;2012:247389.
62. Basiotis PP, Welsh SO, Cronin FJ, Kelsay JL, Mertz W. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. J Nutr. 1987 Sep;117(9):1638-41.
63. Lancaster R, Radd-Vagenas S, Fiatarone Singh M, Noble Y, Daniel K, Mavros Y, et al. Electronic food records among middle-aged and older people: A comparison of self-reported and dietitian-assisted information. Nutrition & Dietetics. 2020.
64. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires - a review. Public Health Nutr. 2002 Aug;5(4):567-87.
65. Ngo J, Gurinovic M, Frost-Andersen L, Serra-Majem L. How dietary intake methodology is adapted for use in European immigrant population groups - a review. Br J Nutr. 2009 Jul;101 Suppl 2:S86-94.

Chapter 6

General discussion

Summary of the results from previous chapters

This thesis investigated approaches to improve the efficiency and accuracy in collecting and handling large-scale dietary data, specifically, for the Dutch National Food Consumption Surveys (DNFCS). Firstly, we investigated possible simplifications of the method that is currently applied in the survey by removing unnecessary steps, such as less important food details, and a more simple recipes function. Secondly, new possibilities for incorporating self-administered tools (i.e. smartphone apps) in the future surveys were explored. The collected chapters provided multidimensional evidence for constructing a self-administered smartphone app for dietary data collections in DNFCS. The learned lessons can also be useful for other large-scale dietary studies that are interested in collecting information on all foods and beverages consumed.

Simplifications on methodologies used in current DNFCS were evaluated in chapter 2 and 3. In chapter 2, the impact of less detailed characterization of consumed foods was simulated. One third of the total food descriptors used in the data collection were identified as less important in determining the nutrient intake distributions of the population. The deletion of those descriptors could potentially contribute to around 1000 hours reduction in the collection and handling of dietary data. The majority (80%) of the differences between percentile estimates of the population nutrient intake distributions ranged from 0% to 1% before and after facet deletion. On the other hand, chapter 3 addressed the methodological simplification of collecting information on mixed dish consumption. The direct use of standard recipes without asking for details on deviations from the standard could greatly reduce the complexity of the recipe pathway, and avoid the appearance of detailed questions of which the participants that do not cook themselves often do not know the answer. A minor impact on the nutrient intake and food group consumption distributions was observed primarily due to the relatively low mixed meal consumption in the Netherlands.

Chapter 4 and 5 focused on evaluating technological developments and validations on recent smartphone food record applications. In chapter 4, the content analysis of recipe functions in popular commercial apps showed a varied functional design and displayed differences in nutrient contents of selected recipes among the apps, which were mostly due to the underlying food composition databases (FCDs). Moreover, the lack of application of yield and retention factors affected the intake estimation of heat-sensitive nutrients in certain dishes. A systematic review and meta-analysis of validation studies conducted on smartphone apps in chapter 5 revealed that energy intake derived from self-administered by apps in real-life settings were underestimated compared to more-established reference methods in general. Differences in energy intake were smaller in validation studies in which the same FCD was used in the app and reference method.

Reflections on the study findings

1. **Importance of food descriptions**

Chapter 2&3 have illustrated that a reduction of food descriptors and the use of standard recipes without modifications could reduce the length of the interview and ease the data handling, without much impact on population nutrient intake distributions. However, apart from estimating food consumption and nutrient intakes, the collection of dietary data with adequate details makes other use of the data possible, such as assessing exposures to harmful substances (e.g. heavy metals, mycotoxins and acrylamide) (1). Information on the level of chemical exposures provides evidence for risk assessment and management of a safe and healthy food environment (2). Therefore, the usability of NFCS in estimating food safety exposure should also be considered in pursuit of a simpler methodology. Another use of food descriptors is to guide the food selection process completed by the participant themselves in self-administered methods. Findings from previous usability studies suggested that the lack of food descriptors poses difficulties in finding a specific food item within a database, resulting in a higher chance of selecting generic food items (3-5). Moreover, in chapter 5 it was speculated that this lack of guidance in apps is one of reasons for the general underestimation of the food consumption. In summary, the existence of a certain level of food details in self-administered methods enables the acquisition of useful information and provides better guidance for food selection.

Compared to web-based or computer-based dietary assessment methods, food records based on smartphones could benefit from technologies like barcodes or image capturing. For commercial foods or recipes, the detailed questions on brand names, packaging materials or other related information could be automatically captured with a simple scan, which could reduce the time and effort needed for text input, and prevent making mistakes in choosing the foods from a list (6). Note that the successful operation of this automatic linkage is built under the premise of an established pathway between the food items and food product databases (preferably with updated country-specific food products) that stores the barcodes and associated product information (7).

2. **Strategies in recording mixed dish intake**

Calculating nutrient contents of ingredients in cooked meals without yield and retention factors could lead to a large extent of misestimation in the intake of heat-sensitive nutrients (chapter 4). Due to the complexity of incorporating conversion factors into individual ingredients, chapter 3 investigated if standard recipes could be used directly without modifications. By already taking yield and retention factors into account, the source of error would mainly come from the differences in ingredient composition with the mixed dishes actually consumed. The extent of nutrient misquantification and food group misclassification at the population group level depends on the proportion of food consumed through mixed dishes, which was found to be rather low (10%) in the Netherlands, causing an unnoticeable impact to the nutrient distribution at the population level. However, there was a larger difference found in certain food group intake distributions when using the standard recipes, indicating their limited suitability in reflecting the ingredients and quantities people used in real-life situations. Similar to findings from Tucker et al., mixed dishes prepared in real-life settings were much simpler than the recipes from the Internet or cookbooks (8, 9). Hence, the representation of standard recipes could be improved taking a large range of both cookbooks and real-life recipes into account, and incorporate functions that allow a certain level of customisation to the standard recipes (e.g. potential ingredients with checkboxes). The results from chapters 3&4 fill a gap in the limited evidence on the impact of errors when reporting mixed dish intake (10), and hopefully raises the awareness of app developers to take these factors into consideration in future app development.

3. Importance of Food Composition Databases (FCDs)

Commercial food products on the market are evolving rapidly, including the introduction of new products and modifications of existing products (11, 12). The reflections of these factors might differ across different FCDs. FCDs have shown to be one of the most influential determinants in comparing energy and nutrient intakes among different commercial apps (chapter 4), and in explaining discrepancies in apps and their reference methods across validation studies (chapter 5). Energy and macronutrients were underestimated by apps when three recipes were entered in apps and compared to calculations based on the Dutch National FCD (NEVO) by researchers (chapter 4). Griffith et al. also found underestimation of energy and nutrient using apps when compared the US Department of Agriculture's (USDA) National Nutrient Database (13). In contrast, Ferrara et al. found most of the apps tended to overestimate intake greatly (14), while Chen et al. found a balanced over- and underestimation among apps compared to USDA's FCD (15). The meta-analysis of validation studies in chapter 5 revealed that although underreporting (intentionally/unintentionally) by participants was the main contributor for the underestimation of all apps, studies that used different FCDs for the app and the reference method had higher discrepancies in intake estimations than those with the same FCDs.

The explanations for the discrepancies between commercial and national FCDs were that commercial apps might use FCDs from the country where they were developed, they might be more frequently updated in terms of commercial food products than national FCDs, but also have a higher chance of false information from crowdsourcing (15). Furthermore, the availability of micronutrient information in general was limited in apps, while most national FCDs have a rather complete nutrient profile. Proactive approaches to supplement commercial FCDs with nutrient information from national FCDs have been undergoing (16).

Careful considerations should be made when harmonising different FCDs, since they might differ in various aspects. For example, the difference in nutrient expressions are highly likely to exist, which requires inspections and adjustment before integrating (17).

Reflections on methodology:

1. Machine learning and the utility of large datasets

Machine learning has been gradually incorporated into the dietary assessment area in recent years, with the main application in automatic image and spoken language recognition. Alternatively, machine learning could also be a suitable technique to predict important features (e.g. identify and reduce detailed questions) and automate certain tasks (e.g. link food to FCDs) that were mostly done manually, making use of existing data for training the algorithms (chapter 2). Another potential use could be to reduce the number of items available in the food/recipe list according to their popularity from previous surveys. Participants' inclination of answers or food choices could be differentiated based on their socioeconomic status or other personal characteristics. This information would be useful for developing customised survey protocols targeted to different population groups.

A limitation of machine learning is that the results are only applicable to the same instrument components and design. Specifically, the reduction of facets in chapter 2 was only limited to the FCD tested, in this case, the NEVO 2011/3.0. With an updated or different FCD, the proposed facet reduction might lead to different results. Besides, with the addition of new technologies and functions, the convenience of getting specific detailed information also differs. Hence, it is needed to apply a similar study protocol as in chapter 2 with each new addition of technologies and functions.

Exploiting an existing large-scale dataset for potential methodological improvement, as in chapter 2&3, has not often been found in researches developing new dietary assessment methods. Especially national surveys that have the advantage of already collected data from previous survey rounds, could consider manipulating the data somehow in understanding the utilisation of certain features or options. Cautions for manipulating large dataset should be made. Firstly, due to the complex nature of dietary intake and the detailed information collected from dietary surveys, factors that might influence the nutrient outcome should all be carefully considered when simulating procedures that were aimed to be applied automatically in real-life (e.g. apply conversion factors for calculating cooked amount). Secondly, error checking for large datasets can be problematic. Preparing a randomly selected sample dataset can be more efficient in testing the protocol. Thirdly, the limitations and assumptions in the data are reproduced, and it is well-known that reported dietary data collected from the survey always includes error.

2. Best practices for reporting and evaluating new dietary assessment tools

Although a fast-growing industry of technology-assisted dietary assessment provides a wide range of selections for specific study purposes, the accuracy of the new tools is often unknown due to a lack of proper validation studies. A method that has high validity is capable of providing a useful measurement for a given purpose and has an established internal and external validity (18-20). Hence, the validity and reliability of these tools needs to be further explored with a proper evaluation strategy (21). A detailed guideline for reporting validation studies (STROBE-nut, Strengthening the Reporting of Observational Studies in Epidemiology Statement-nutritional epidemiology extension) may improve reporting of epidemiological and validation studies involving dietary assessment methods and enhance the quality of the published evidence (21). A checklist adapted from STROBE-nut with more specific guidance on the study design and results interpretation was proposed by Kirkpatrick et al (19). Another guideline developed by Eldridge et al. based on STROBE-nut consists of aspects that are more specific for reporting and validating technology-based tools (22).

As we found in chapter 5, none of the validation studies of dietary assessment apps used recovery biomarkers as their reference measure. Although they were identified as the optimal approach for measuring true intake, their limitations in cost and available nutrients have led to a reliance on the measurement of relative validity for most studies (23). The complex and dynamic nature of dietary intake contributes to difficulties in evaluating relative validity. For example, both the test and reference method might all be subjected to self-reporting errors, and the effect of using both methods might differ with using the test method only. So careful considerations on the allocated period for each method and overlaps of periods of different methods are needed to avoid learning effects in the test and reference methods. Besides, as we found in chapter 5, it can be difficult for some studies to unify the FCD used in the test and reference method. As described before in chapter 5, a larger difference between the methods using different FCDs than methods with unified FCDs has been found.

Apart from nutrients, insight in the validity of food group recordings is needed if assessing food group consumption is the purpose of the tool. This can for example be the case if consumption needs to be compared to food-based dietary guidelines, which have become more useful compared to nutrient-based guidelines in disseminating healthy eating to the public (24-29). Moreover, insight in the validity of consumption of food groups is also useful to trace back to the underlying cause of limited validity for energy and nutrients. Therefore, it is advised that the future reporting of validation studies should incorporate food group comparisons between the test and the reference method. Moreover, the investigation of omissions and intrusions of specific food items in tools can be an alternative method of comparison and could provide even more detailed insights for the source of measurement error. Also, the discrepancy of portion size estimation by the test and reference methods is another main source of error worth comparing.

Different levels of validity might exist for one tool in different population groups (10). Typically, a wide range of population groups is included in the sample population for NFCS (30). Hence, validation studies in a diverse population for NFCS are needed to establish external validity (31). Specifically, the practicality of technology-assisted method might be limited in segments of the population who have low e-literacy levels or motivation, leading to a weakened capacity to identify food items or portion sizes and a larger drop-out (32, 33).

From the meta-analyses in chapter 5, it became clear that observed variation in differences in intake between methods has been rarely discussed for practical relevance. For instance, the Bland-Altman analysis has been applied more frequently in evaluating new methods in recent years due to its ability in detecting the presence and direction of bias at the group level, and the extent of its variation at the individual level (34). However, most studies focussed on interpreting the average bias at the group level, while the practical relevance of the individual variance (the width of limits of agreement) has rarely been assessed. In the included studies the width of the limits of agreement ranged from 447kcal to 3263kcal across studies. It is therefore advised to define acceptable limits of the variation in both the group and individual level a priori taking the desired use of the dietary assessment method into account.

Before conducting a validation study of a new dietary assessment tool, other types of evaluation are very useful during the development process. For example, the usability study on ASA24 (Automated Self-Administered Dietary Assessment Tool) found that certain usability issues might limit the participation rate in a group of low-income participants. Participant experiences with certain features could be collected from usability studies, such as probes that could exacerbate or reduce social desirability biases (19, 35). A tool designed with feedbacks from users will eventually lead to better cooperation, which will, in turn, translate to a better quality of the data. Therefore, by taking usability issues into account, customised dietary assessment methods based on respondent characteristics (e.g. educational status, physiological status, geographical location, technology use) can be developed and would potentially improve the validity of the test method.

Aspects (not from chapters) that are important to consider when moving from interviewer-administered 24h dietary recalls to self-administered smartphone food records:

1. Trend of smartphone usage/data privacy

ICT-based technologies (computers, smartphones) are more expensive platforms than pen and paper methods for dietary assessment from the user perspectives, and were deemed inaccessible for groups of lower socioeconomic status one decade ago (36). However, the increased coverage and ubiquity of worldwide smartphone ownership in the past ten years indicated that the affordability is of less concern and more digital devices like smartwatches and tablets are also penetrating in our daily life (18). This trend has fostered the increased access to innovative methods for assessing dietary intake. Specifically, 98 percent of Dutch households had internet access in 2018, putting the Netherlands at the forefront within Europe. The Netherlands also ranks among the European top in terms of high-speed broadband connectivity, mobile internet usage and maturity of the Mobile Health market (37). This wide application of the internet and mobile devices provides a relatively convenient start-up for implementing surveys using smartphones. Still, the level of technology-literacy of particular population groups needs to be taken in careful consideration.

Besides, the capacity of apps and other devices in monitoring other health behaviours and indicators (e.g. physical activity, sleep, heart rate, etc.) poses new opportunities for collecting a complete personal lifestyle and health profile (38). The large-scale data serves as a complement to traditional surveillance studies that could reveal new insights about the interrelationships between the environment, society and health behaviours. Data sharing partnerships between research institutions and industries might be needed for certain aims of research (39). However, at the same time, this poses challenges in ethical issues and protecting user privacy. Also the threshold of access to data on the individual level differs across countries, some countries in Europe having stricter privacy laws than other countries (40). Citizen concerns for data security also differ, for example, Swiss citizens are more concerned than citizens in the Netherlands (37). Careful considerations should be made in terms of providing standards for anonymizing activity data and transparent explanations on the use of data to the participants (39).

2. Cost implications

The use of interviewer-assisted food consumption surveys with much detail in food description is labour intensive and costly, which led to an exploration of the development of cost-effective technologies. This requires high investment in the early stage of the app development and testing, depending on available resources in financial, logistical and staff conditions. Once the app is ready, cost and time can be saved in organizing the study, collecting and handling data, as well as calculating dietary intakes, potentially leading to a return on investment (41). The decreased cost of data collection could enable the inclusion of more people into large-scale studies, making the study sample to be more representative of the general population.

Still, despite that removing interviewers might reduce errors related to contact bias, it may introduce additional challenges and different sources of error, causing a declined quality of the collected data (42). As the complexity of interaction with technology increase, it is reasonable to expect additional cost for technical support and training of participants, which have shown to improve user cooperation and proficiency (43). As seen in the self-administered 24HR ASA24, on-demand technical assistance was available to ensure the data quality and participant retention (44). In general, there is a lack of information on costs associated with the development and implementation of new technologies in a survey setting (45). The evaluation of costs with respect to each aspect for a new method in comparison with the traditional methods could provide additional input for decision-making (45).

3. International harmonization

In order to develop collaborative strategies to optimize the health of the European populations, the collection of comparative food consumption data across Europe by a common framework of procedures and tools has been suggested by EFCOSUM (European Food Consumption Survey Methods) project and later validated in the 'European Food Consumption and Validation' (EFCOVAL) project. In addition, European Food Safety Authority (EFSA) emphasised the importance of pan-European dietary exposure assessments from harmonized the food consumption surveys (46). However, the differences in culture, reluctance to change currently used methods, organization structure and budgets for survey conduction are the limiting factors for methodological harmonization across countries. It was suggested that complete standardization should not be strived for at the cost of overall data quality in any individual country (40). Hence, a compromise between the level of harmonization and the practical context within each country should be reached. Although GloboDiet has been suggested as the 'first choice' instrument for data collection, the potential cost of its adaptation has prevented the use of it in some countries. Besides, other methodological aspects have also contributed to incomparability across countries, such as differences in FCD and its included nutrients, age group categories, etc. (47). Hence, the exploration of a more cost-effective method might provide new opportunities for a better future harmonization across countries. Due to the lack of validations and applications to particular population groups, smartphones have not yet been used for dietary data collection in any of the NFCS in European countries (47). The early initiative of collecting dietary data using a smartphone app in the DNFCS, taking advantage of the ever-growing smartphone penetration in the Netherlands, could provide insights for other countries that are aiming at the same direction and have expected increased use of smartphones.

Future directions

The self-reporting bias in traditional dietary assessment methods is the most worrying source of bias. Despite its limitations, self-reported intake could provide necessary detailed information about the complexity of what individuals consume. Such information is critical for providing information about dietary patterns and diet quality in order to evaluate questions such as whether intakes are consistent with recommendations or associated with health outcomes (27). Technology involvement can only solve certain level of unintentional underreporting (41, 48), while intentional underreporting cannot be easily solved. As long as the participants are aware of being monitored, the tendency to alter their diet is inevitable, especially for prospective methods (3). Until now, none of the self-administered methods has shown significant improvements in accuracy, with most of them underestimating dietary intake compared to traditional methods (30). With this in mind, parallel efforts should be put into searching for more convenient technologies and advancing statistical models that could adjust for measurement error, using data from validation studies with objective measures (e.g. recovery biomarkers).

A participation rate of less than 50% was found in the majority of the countries that have conducted NFCS (40, 49), with most countries relying on interviewer-assisted dietary assessment methods for current survey collections. The future participation rate was predicted to drop further if the survey methodology could not keep up with the speed of technological development (50). In general, increased compliance and willingness in using technology-assisted methods has been found in previous usability studies, due to more efficient data input, process, and flexibility in registering intake at their own convenience (51). However, there were varying levels of receptivity in a wide population group using self-administered methods targeted to large-scale data collection (e.g. ASA24)(44, 52, 53). In addition, compared to people who voluntarily use apps for dietary self-monitoring, people who were invited by a third party might not understand the purposes of the study and the importance of correct and precise recording, or have limited knowledge about their food consumption. Continued investigation on incorporating new technologies into large-scale dietary monitoring systems is an essential step for developing sufficiently accurate, cost-saving yet easy to participate future surveys.

A planned methodological change in NFCS would inevitably constitute a change in data collected, affecting the continuity of results from different survey waves, which would impair the estimation of the population intake trends over time (54). Bridging studies that investigate both methods in parallel for a sample of the population could potentially reveal the systematic bias between the methods. Compatible and comparable results from new and old method would also enable the implementation of a multi-modal approach for new survey collections, which would offer the respondents the option of either an interviewer-administered or a self-administered survey (30).

The use of food records is usually associated with behavioural change, which was considered as a disadvantage for nutrition monitoring in many countries. However, with the incorporation of barcode scanning, and potential use of image recognition and analysis, the practicality of using an app for more days of food recording might be possible, and users might be less likely to alter their diet for prolonged periods. However, inconsistent evidence for the long-term use has been found, either more underreporting due to boredom or fatigue (55), or increased familiarity and a better performance in using the app (56). Features that could eliminate repetitive actions (e.g. saving favourite or previous food items) could potentially reduce the extent of underreporting (57). More research on how food recording differs on the progression of the app use is warranted.

The use of currently available technology may not necessarily reduce all respondent burden. Although barcode scanning could automate the data entering to some extent, they are only applicable to branded food products for which the packages are available to the respondent. Text input remains the main method, which might impact the level of convenience in using the apps (58). The current incorporation of image-taking in technology-based tools makes it possible to omit food identification and amount estimation from respondents (22). These functions were proven to be useful in facilitating memory, avoiding underreporting, and ease the recording process, and might benefit most the population groups with low technology literacy (59). Especially automatic image recognition provides obvious advantages by reducing both respondent and researcher burden (60), and might be the future mainstream with the advancement in computer vision and deep learning. However, current development in image recognizing is not mature enough to be fully automated. Enormous amounts of pictures of foods and dishes are required as the input for algorithm training. The intra- and inter-individual variability with which food is prepared, served, and consumed in free-living situations brings up the levels of complexity for accurate recognition (18).

Current image-based apps still require the supplementation of descriptive information and huge data-handling efforts from researchers, meaning that participant burden may not be sufficiently reduced to offset the additional costs of extra time of researchers in the current situation (61). The evaluations of these methods were mostly conducted in controlled settings, the feasibility outside of controlled settings needs further evaluation. Hence, the full transfer from text to images would only be feasible when fully automated image recognition and analysis is achievable (62), and their requirements on the digital environment are within the technological capacity of the average consumer device (10).

Using smartphone apps for dietary intake measurement is a promising future for nutrition monitoring, given the increased penetration of smartphone use worldwide and its capability to incorporate technological features. In Figure 1, a workflow is given for developing and testing a new dietary assessment method for the use of a NFCS based on the current process of app development in the DNFCS. In the beginning phase, factors such as estimated cost and the collection of other dietary components, such as time/place of consumption need to be considered. Furthermore, taking advantages of existing results of previous surveys, as presented in chapter 2 and 3 of this thesis, provides useful evidence for methodological development. Experiences from other studies serve as essential references when integrating new technologies into nutrition surveys. An iterative process of developing and testing will ensure an user-adapted tool to be produced. During the developing phase, the affiliate components such as FCDs should be kept representative of the country-specific diet and equipped with information on both generic and branded food items. An effective participant training program and data cleaning protocol should be prepared. The usability and validity of the tool assessed in different population groups, together with bridging studies between the old and the new method are necessary to ensure the consistency of results before and after the methodological change.

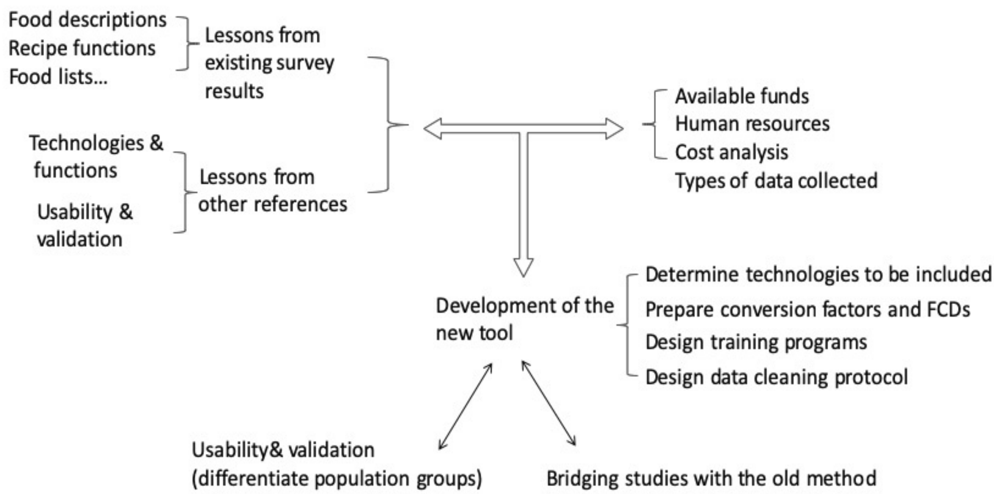


Figure 1. Process of dietary assessment tool development for NFCS

Conclusions

Although being acknowledged as error-prone, self-reported dietary intake at the national level has reaffirmed its value as an essential scientific foundation for developing public health policies, food-based guidelines, and understanding diet and health relationships (54, 63). In order to improve the accuracy and reduce the burden of obtaining dietary data, the dietary assessment field is working on enhancing existing methods, developing innovative instruments using new technologies, and incorporating statistical methods for error adjustments (64-67). This thesis concluded that a reduced amount of food descriptors and a simplified recipe pathway in 24HR does not have a large impact on the population nutrient intake distributions and could potentially reduce the cost of future interviewer-administered 24HRs. The findings thus indicated that the collection of certain details could be omitted for developing smartphone apps built for self-administered food records. On the other hand, whether a self-administered food record tool has sufficient accuracy for dietary monitoring needs to be determined. No biomarker-based validity studies are available yet; and relative to other dietary assessment methods there seems to be more underreporting. Insight in the underlying causes of this underestimation and variations in accuracy is largely lacking.

Smartphone apps for dietary assessment have rarely been tested for large-scale studies, especially for NFCS. Several main reasons might explain the lack of such explorations, including varied acceptability among different population groups, the susceptibility to behavioural change using prospective methods, and insufficient insight in the accuracy of smartphone food records. Still, with the undeniable trend towards more automated procedures in dietary assessment, a self-administered method for NFCS is likely to take over the interviewer-administered method in the near future. Several strategies to cope with the challenges in developing and testing self-administered methods for large surveys exist. Firstly, data mining of previously collected food consumption data is a cost-effective approach that could potentially reveal useful information. Secondly, the sources of errors using the new method should be traced to enable further adjustments of the tool. Thirdly, the design of validation studies should comply with the established recommendations and cover all population groups of interest to the survey.

Reference

1. Mengelers M, de Wit L, Boon P, Franz E, Bouwknegt M, de Jonge R, et al. How safe is our food? Background report to 'What's on our plate? Safe, healthy and sustainable diets in the Netherlands.'. Bilthoven National Institute for Public Health and the Environment 2016.
2. EFSA. General principles for the collection of national food consumption data in the view of a pan-European dietary survey. EFSA Journal. 2009;7(12):1435.
3. Illner AK, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol. 2012 Aug;41(4):1187-203.
4. Zimmerman TP, Hull SG, McNutt S, Mittl B, Islam N, Guenther PM, et al. Challenges in converting an interviewer-administered food probe database to self-administration in the National Cancer Institute Automated Self-administered 24-Hour Recall (ASA24). Journal of food composition and analysis : an official publication of the United Nations University, International Network of Food Data Systems. 2009;22(Suppl 1):S48-S51.
5. Koch SAJ, Conrad J, Hierath L, Hancock N, Beer S, Cade JE, et al. Adaptation and Evaluation of Myfood24-Germany: A Web-Based Self-Administered 24-h Dietary Recall for the German Adult Population. Nutrients. 2020;12(1):160.
6. Steele R. An overview of the state of the art of automated capture of dietary intake information. Crit Rev Food Sci Nutr. 2015;55(13):1929-38.
7. Maringer M, Wisse-Voorwinden N, Veer PV, Geelen A. Food identification by barcode scanning in the Netherlands: a quality assessment of labelled food product databases underlying popular nutrition applications. Public Health Nutr. 2018 Jul 2:1-8.
8. Tucker KL, Bianchi LA, Maras J, Bermudez OI. Adaptation of a food frequency questionnaire to assess diets of Puerto Rican and non-Hispanic adults. American journal of epidemiology. 1998 Sep 1;148(5):507-18.
9. Ngo J, Gurinovic M, Frost-Andersen L, Serra-Majem L. How dietary intake methodology is adapted for use in European immigrant population groups - a review. Br J Nutr. 2009 Jul;101 Suppl 2:S86-94.
10. Prinz N, Bohn B, Kern A, Püngel D, Pollatos O, Holl RW. Feasibility and relative validity of a digital photo-based dietary assessment: results from the Nutris-Phone study. Public Health Nutrition. 2019;22(7):1160-7.
11. Mueller S, Szolnoki G. The relative influence of packaging, labelling, branding and sensory attributes on liking and purchase intent: Consumers differ in their responsiveness. Food Quality & Preference. 2010;21(7):0-783.
12. World food consumption patterns—trends and drivers. European Commission, 2015.
13. Griffiths C, Harnack L, Pereira MA. Assessment of the accuracy of nutrient calculations of five popular nutrition tracking applications. Public Health Nutr. 2018 Jun;21(8):1495-502.

14.Ferrara G, Kim J, Lin S, Hua J, Seto E. 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 Uhealth. 2019 May 17;7(5):e9232.
15.Chen J, Cade JE, Allman-Farinelli M. The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment. JMIR Mhealth Uhealth. 2015 Dec 16;3(4):e104.
16.Carter MC, Hancock N, Albar SA, Brown H, Greenwood DC, Hardie LJ, et al. Development of a New Branded UK Food Composition Database for an Online Dietary Assessment Tool. Nutrients. 2016 Aug 5;8(8).
17.Leclercq C, Valsta LM, Turrini A. Food composition issues--implications for the development of food-based dietary guidelines. Public Health Nutr. 2001 Apr;4(2b):677-82.
18.Rollo ME, Williams RL, Burrows T, Kirkpatrick SI, Bucher T, Collins CE. What Are They Really Eating? A Review on New Approaches to Dietary Intake Assessment and Validation. Current Nutrition Reports. 2016;5(4):307-14.
19.Kirkpatrick SI, Baranowski T, Subar AF, Tooze JA, Frongillo EA. Best Practices for Conducting and Interpreting Studies to Validate Self-Report Dietary Assessment Methods. J Acad Nutr Diet. 2019 Nov;119(11):1801-16.
20.Barry AE, Chaney B, Piazza-Gardner AK, Chavarria EA. Validity and reliability reporting practices in the field of health education and behavior: a review of seven journals. Health education & behavior : the official publication of the Society for Public Health Education. 2014 Feb;41(1):12-8.
21.Lachat C, Hawwash D, Ocke MC, Berg C, Forsum E, Hornell A, et al. Strengthening the Reporting of Observational Studies in Epidemiology - nutritional epidemiology (STROBE-nut): An extension of the STROBE statement. Nutr Bull. 2016 Sep;41(3):240-51.
22.Eldridge AL, Piernas C, Illner AK, Gibney MJ, Gurinovic MA, de Vries JHM, et al. Evaluation of New Technology-Based Tools for Dietary Intake Assessment-An ILSI Europe Dietary Intake and Exposure Task Force Evaluation. Nutrients. 2019 Jan;11(1).
23.Jenab M, Slimani N, Bictash M, Ferrari P, Bingham SA. Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet. 2009 Jun;125(5-6):507-25.
24.Julian-Almarcegui C, Bel-Serrat S, Kersting M, Vicente-Rodriguez G, Nicolas G, Vyncke K, et al. Comparison of different approaches to calculate nutrient intakes based upon 24-h recall data derived from a multicenter study in European adolescents. Eur J Nutr. 2016 Mar;55(2):537-45.
25.Millen BE, Abrams S, Adams-Campbell L, Anderson CA, Brenna JT, Campbell WW, et al. The 2015 Dietary Guidelines Advisory Committee Scientific Report: Development and Major Conclusions. Adv Nutr. 2016 May;7(3):438-44.

26.Monteiro CA, Cannon G, Moubarac JC, Martins AP, Martins CA, Garzillo J, et al. Dietary guidelines to nourish humanity and the planet in the twenty-first century. A blueprint from Brazil. Public Health Nutr. 2015 Sep;18(13):2311-22.
27.Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. The Journal of nutrition. 2015 Dec;145(12):2639-45.
28.Micha R, Khatibzadeh S, Shi P, Andrews KG, Engell RE, Mozaffarian D, et al. Global, regional and national consumption of major food groups in 1990 and 2010: a systematic analysis including 266 country-specific nutrition surveys worldwide. BMJ Open. 2015 Sep 24;5(9):e008705.
29.Castro-Quezada I, Ruano-Rodriguez C, Ribas-Barba L, Serra-Majem L. Misreporting in nutritional surveys: methodological implications. Nutricion hospitalaria. 2015 Feb 26;31 Suppl 3:119-27.
30.Amoutzopoulos B, Steer T, Roberts C, Cade J, Boushey C, Collins C, et al. Traditional methods v. new technologies–dilemmas for dietary assessment in large-scale nutrition surveys and studies: a report following an international panel discussion at the 9th International Conference on Diet and Activity Methods (ICDAM9), Brisbane, 3 September 2015. Journal of nutritional science. 2018;7.
31.Thompson FE, Subar AF, Boushey CJ, Ferruzzi MG, Delahanty LM. Chapter 1 - Dietary Assessment Methodology A2 - Coulston, Ann M. Nutrition in the Prevention and Treatment of Disease (Fourth Edition): Academic Press; 2017. p. 5-48.
32.Thompson FE, Subar AF. Dietary assessment methodology. Nutrition in the Prevention and Treatment of Disease: Elsevier; 2017. p. 5-48.
33.Tay W, Kaur B, Quek R, Lim J, Henry CJ. Current Developments in Digital Quantitative Volume Estimation for the Optimisation of Dietary Assessment. Nutrients. 2020;12(4):1167.
34.Lombard MJ, Steyn NP, Charlton KE, Senekal M. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J. 2015 Apr 22;14:40.
35.Subar AF, Thompson FE, Potischman N, Forsyth BH, Buday R, Richards D, et al. Formative research of a quick list for an automated self-administered 24-hour dietary recall. J Am Diet Assoc. 2007 Jun;107(6):1002-7.
36.Long JD, Littlefield LA, Estep G, Martin H, Rogers TJ, Boswell C, et al. Evidence review of technology and dietary assessment. Worldviews on Evidence-Based Nursing. 2010;7(4):191-204.
37.De Korver F. The Digital Health market in the Netherlands and Switzerland Switzerland: Embassy of the Kingdom of the Netherlands in Berne, 2019.
38.Svensson A, Larsson C. A Mobile Phone App for Dietary Intake Assessment in Adolescents: An Evaluation Study. JMIR Mhealth Uhealth. 2015 Nov 3;3(4):e93.

39.Hicks JL, Althoff T, Sosic R, Kuhar P, Bostjancic B, King AC, et al. Best practices for analyzing large-scale health data from wearables and smartphone apps. npj Digital Medicine. 2019;2(1):45.
40.De Henauw S, Brants HA, Becker W, Kaic-Rak A, Ruprich J, Sekula W, et al. Operationalization of food consumption surveys in Europe: recommendations from the European Food Consumption Survey Methods (EFCOSUM) Project. Eur J Clin Nutr. 2002 May;56 Suppl 2:S75-88.
41.Shim J-S, Oh K, Kim HC. Dietary Assessment Methods in Epidemiologic Studies. Epidemiology and Health. 2014;36:e2014009.
42.Lancaster R, Radd-Vagenas S, Fiatarone Singh M, Noble Y, Daniel K, Mavros Y, et al. Electronic food records among middle-aged and older people: A comparison of self-reported and dietitian-assisted information. Nutr Diet. 2020.
43.Six BL, Schap TE, Kerr DA, Boushey CJ. Evaluation of the Food and Nutrient Database for Dietary Studies for use with a mobile telephone food record. Journal of food composition and analysis : an official publication of the United Nations University, International Network of Food Data Systems. 2011 Dec 1;24(8):1160-7.
44.Kupis J, Johnson S, Hallihan G, Olstad DL. Assessing the Usability of the Automated Self-Administered Dietary Assessment Tool (ASA24) among Low-Income Adults. Nutrients. 2019;11(1):132.
45.Burley V, Timmins K, Cade J, Greenwood D, Husain F, Gill V, et al. Making the best use of new technologies in the National Diet and Nutrition Survey: a review. 2015 03/31.
46.Guidance on the EU Menu methodology1 European Food Safety Authority.
47.Rippin H, Hutchinson J, Evans C, Jewell J, Breda J, Cade J. National nutrition surveys in Europe: a review on the current status in the 53 countries of the WHO European region. Food & Nutrition Research. 2018;62.
48.Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Research. 2017;6:926.
49.Walton J. Dietary Assessment Methodology for Nutritional Assessment: A Practical Approach. Top Clin Nutr. 2015;30(1):33-46 10.1097/TIN.0000000000000018.
50.Keeble C, Baxter PD, Barber S, Law G. Participation Rates In Epidemiology Studies And Surveys: A Review 2007–2015. The Internet Journal of Epidemiology. 2016 01/27;14.
51.Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med. 2006 Jun;31(3):205-23.
52.Thompson FE, Dixit-Joshi S, Potischman N, Dodd KW, Kirkpatrick SI, Kushi LH, et al. Comparison of Interviewer-Administered and Automated Self-Administered 24-Hour Dietary Recalls in 3 Diverse Integrated Health Systems. American journal of epidemiology. 2015 Jun 15;181(12):970-8.

53.Kirkpatrick SI, Gilsing AM, Hobin E, Solbak NM, Wallace A, Haines J, et al. Lessons from Studies to Evaluate an Online 24-Hour Recall for Use with Children and Adults in Canada. Nutrients. 2017 Jan 31;9(2).
54.Archer E, Hand GA, Blair SN. Validity of U.S. nutritional surveillance:National Health and Nutrition Examination Survey caloric energy intake data, 1971-2010. PLoS One. 2013;8(10):e76632.
55.Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr. 2013 Feb 14;109(3):539-46.
56.Chen J, Berkman W, Bardouh M, Ng CYK, Allman-Farinelli M. The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. Nutrition. 2019;57:208-16.
57.Weber F, Haering C, Thomaschke R. Improving the human-computer dialogue with increased temporal predictability. Hum Factors. 2013 Oct;55(5):881-92.
58.Ming Z-Y, Chen J, Cao Y, Forde C, Ngo C-W, Chua TS, editors. Food Photo Recognition for Dietary Tracking: System and Experiment. International Conference on Multimedia Modeling; 2018: Springer.
59.Gemming L, Doherty A, Kelly P, Utter J, Ni Mhurchu C. Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. Eur J Clin Nutr. 2013 Oct;67(10):1095-9.
60.Rangan AM, O'Connor S, Giannelli V, Yap MLH, Ming L, Roy R, 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 2015--2015;3(4):3-14.
61.Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition. 2014 Nov-Dec;30(11-12):1257-66.
62.Martin CK, Nicklas T, Gunturk B, Correa JB, Allen HR, Champagne C. Measuring food intake with digital photography. J Hum Nutr Diet. 2014 Jan;27 Suppl 1:72-81.
63.Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. Eur J Clin Nutr. 1991 Dec;45(12):583-99.
64.Neuhouser ML, Tinker L, Shaw PA, Schoeller D, Bingham SA, Horn LV, et al. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. American journal of epidemiology. 2008;167(10):1247-59.
65.Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc. 2010 Jan;110(1):48-51.
66.Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. J Natl Cancer Inst. 2011 Jul 20;103(14):1086-92.
67.Prentice RL, Huang Y. Measurement error modeling and nutritional epidemiology association analyses. Can J Stat. 2011 Sep 1;39(3):498-509.

Acknowledgements

Here, after 4 years and 1 month in Wageningen, my PhD journey has finally ended in Nov. 2020, with the pandemic of Covid-19 everywhere around the world. Although it is difficult to meet with people around this time, I still would like to mention and acknowledge those who inspired, guided, and accompanied me throughout this entire journey.

First of all, I would like to give special thanks to my supervisors, Hendriek Boshuizen and Marga Ocke. Marga has been a close contact to me throughout my entire PhD, we had countless meetings regularly, and tackled big or small challenges together. I am deeply grateful to your patience, guidance, and trust in me, which helped me gradually gained confidence and independence in doing research. Your serious attitude towards science will always be a good example for me. Hendriek, as my promoter, provided substantial support to me in both work and life. I truly appreciate and enjoy spending time with you during tutoring hours on statistics whenever I had difficulties. I am grateful for your genuine advice on future careers. Without guidance from both of you, it would take much longer for me to finalize my study. Your understanding and caring for me during the Covid-19 lockdown has greatly comforted me and facilitated me to overcome this difficult time. Despite of your busy schedule, I appreciate that you were actively involved in each step of my study. I feel very privileged to have been reached out by my supervisors and Anouk Geelen to have the opportunity to work in both Wageningen and RIVM. Anouk, although supervised me for only my early period, I gained so much from you, and started to reflect myself occasionally, you also cared a lot with my life outside work and showed great caring to make my life easier, I am so thankful for that!

Special thanks to my reading committee: Prof Pieter van't Veer, Dr Annette Stafleu, Prof Janet Cade, and Dr Simone Eussen, for taking time to review my thesis and to attend my defence. I appreciate that you will be there for my last step of my PhD journey.

During this four years, I received a lot of support from people at the Division of Human Nutrition in Wageningen, I would start from Adrienne Cavelaars, thank you for your help with my financial situations, which have greatly relieved my stress during my research. Also, I would like to thank the chair of my group, Ellen Kampman, who showed a lot of encouragement and trust in me. I am always influenced by your active attitude when you are around. To Kees de Graaf, you are not my direct supervisor, but we had a great time together during the Cabaret experience, best wishes for you! Edith Feskens, even we did not win the dinner game, our memories will always be remembered by me. Alida Melse-Boonstra and Lucy Elburg, thank you very much for thinking about me when you were organizing international activities, which makes me blend in more easily and feel greatly welcomed. Moreover, my appreciations to Jasmijn Mater and Gea Brussen, you are always there for me to solve my problems. Thank you a lot!

To my dearest colleagues in RIVM, Robert Jan de Klein, Zohreh Ghameshlou, Ido Toxopeus, Caroline van Rossum, Susanne Westenbrink, Maryse Niekerk and Arnold Dekkers even though we did not meet regularly, you are always so friendly and welcome. The game night we had in Marga's place was so much fun! Zohreh, and Maryse, thank you very much for your support of my project, and patient in answering my questions! Also, I would like to thank colleagues in the statistic group, especially José Ferreira, your kind support made it possible for me to grasp complicated statistical concepts, it was always a pleasure to chat with you! And Jan van de Kassteele, who is always there for my technical problems, thank you!

To my two dearest Masters students, Eline Nawijn and Andreja Misir. Eline you are such an easy-going person, and did not bother doing the app testing again after your graduation, I greatly appreciate that! And it was nice to see you again during the game night! Hope you have a nice time working in RIVM. Andreja, you are like a superwoman to me, managing to work and finish your thesis at the same time, you are more like a friend to me than a student. I believe you now have much more confidence in your future career in nutrition.

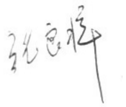
To my earliest friends and colleagues since I joined Wageningen and have left Wageningen already, Martin, Aregash, Masresha, Ibukun, and Apple, thank you for your warmest acceptance to me, blessings to all of you! More people later had contacts with me and we shared a lot of nice memories together, you made this department such a happy place, Maria Duenas, Tsitsi, Inga, Maria Salazar, Tesfaye, Charlotte, Katherine, Eric, Ibukun, Anita, Mahsina, Asrullah, Fusta, Jesca, Giulia, Marijke, Marion, Paulina, Adele, Lowela, Kamalita, Marielle, Pol, Iris, Anniek, Vera, Moniek, Ursula, Dessy, Duong, Aafke. My lovely officemates, Lupita, Maria Jacobo, Umi, Taymara, and Santiago, we spent the most time together, you made my stressful working hours more relaxing through casual chit chats. If I could have such nice officemates as you guys in the future, it will be a blessing for me! To my Chinese gang in HNE, Ruoxuan, Xiaolin, Danny, Max, and Cong. I am so happy to see our family is growing year by year, and feel sad of leaving at this moment. Hope you could enjoy the rest of your study and let's see each other in China someday!

To my second 'parents' in the Netherlands, Ed and Margreet, you provided a home for me for four years, welcomed me every Friday, providing physical and spiritual food. I am so blessed to know you and all of the brothers and sisters in the fellowship, Maaike, EJ, Phil, Lu, Christina, I learned so much from all of you! Meng, Lisheng, Wenzhen, Xiaoqi, Hongli, Yangyang, Yuqing, Dongyao, Yingying, Chen, Shengxin, Weixuan, and Marlon, I was so glad to share thoughts with you, and pray with each of you.

Last but not the least, I appreciate the accompany of Ru, Tania, Sijia, Jiaqi, Danlei, Shan, Yanjun, Yu, Bingyu, Zihan, for countless dinners, trips and parties together. My time in

Wageningen became much more colorful because of all of you! To my long-distance friends, we are connected no matter where we are, Ting, Jingyi, Jing, and Xinyuan. Finally, I wouldn't have come this far without the love and support from my parents and my relatives. My love for you will never end!

With Love,



About the author

Liangzi Zhang was born on June 2nd, 1991 in Datong, China. After completing her high school in Beijing-Concord College of Sino-Canada in 2009, she started majoring in business at Western University, Canada. Two years later, she transferred to the program of nutrition and dietetics, which she was much more interested in. After completing her bachelor in 2013, she worked as a nutrition program coordinator for the university summer camp, then a food journalist in Vancouver. In 2014, She started her master's study in food science and nutrition at the University of Leeds, UK. Her master thesis was about the validation of the MyMealMate food record app, supervised by Dr. Janet Cade. After graduation, she worked as a research assistant in the Nutrition Epidemiology group for the maintenance of the New Branded UK Food Composition Database.

In October 2016, Liangzi started her Ph.D. in the Division of Human Nutrition at Wageningen University and RIVM (National Institute for Public Health and the Environment) in the Netherlands. Her project aimed to enhance the efficiency of the Dutch National Food Consumption Surveys, by eliminating unnecessary steps in the current method, and assess the possibility of incorporating new technologies into future national surveys. Besides her research activities, she supervised several nutrition epidemiology courses and theses of masters students. Liangzi presented her study by posters in WEON conference in RIVM, Max-Rubner conference in Karlsruhe. She presented orally in Dutch Nutrition Science meetings, and at the University of Cambridge and Newcastle University during the 2017 PhD tour in the UK. Apart from her research in the Netherlands, she published three scientific newsfeeds in Chinese and completed a dietary questionnaire for a personalized nutrition program according to the Dietary Guidelines for Chinese Residents (2016).

List of Publications

Published in peer-reviewed journals

Zhang L, Geelen A, Boshuizen HC, Ferreira J, Ocke MC. Importance of details in food descriptions in estimating population nutrient intake distributions. Nutr J. 2019;18(1):17.

Zhang L, Nawijn E, Boshuizen H, Ocke M. Evaluation of the Recipe Function in Popular Dietary Smartphone Applications, with Emphasize on Features Relevant for Nutrition Assessment in Large-Scale Studies. Nutrients. 2019;11(1).

Zhang L, Boshuizen H, Ocke M. How does a simplified recipe collection procedure in dietary assessment tools affect the food group and nutrient intake distributions of the population. Br J Nutr. 2020:1-10.

Publications in preparation

Zhang L, Misir A, Boshuizen H, Ocke M. A Systematic Review and Meta-analysis of Validation Studies Performed on Food Record Apps.

Zhang L*, Sijbrandij J*, Ocke M, Boshuizen H. Validity Measures and Statistical Modeling Choices in Biomarker-based Validation Studies: An In-depth Literature Review.
*Shared first authorship.

Overview of completed training activities

Discipline specific activities	Organizer and location	Year
Courses		
Exposure Assessment	VLAG, Wageningen, NL	2018
Modelling of habitual dietary intake	VLAG, Wageningen, NL	2017
Nutritional Epidemiology Course	Utrecht University, NL	2017
Measurement Error Webinar Series	National Cancer Institute, US	2016
Conferences and meetings		
WEON conference	RIVM, Bilthoven, NL	2018
Dutch Nutritional Science Days	NAV, Heeze, NL	2018-2019
Menu-D meeting	Human Nutrition and Disease, Wageningen, NL	2016-2020
Max-Rubner Conference	Max-Rubner Institute, Karlsruhe, DE	2017
NUTRITION 2020 LIVE ONLINE	American Society for Nutrition, US	2020

General Courses and activities		
VLAG PhD week	VLAG, Baarlo,NL	2016
Scientific writing	WGS, Wageningen, NL	2017
Presentation workshop	Division of Human Nutrition, Wageningen, NL	2019
R course	RIVM, Bilthoven, NL	2019
Career orientation	WGS, Wageningen, NL	2019

Optional courses and activities		
PhD study tour UK	WUR, UK	2017
MSc course Applied Data Analysis	WUR, Wageningen, NL	2016
NAD-Paperclub	Human Nutrition and Disease, Wageningen, NL	2016-2020

Colophon

The research described in this thesis was partly performed at the National institute for Public Health and the Environment (RIVM), Bilthoven.

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

Cover Design and layout: Liangzi Zhang and Mercedes Benjaminse
Printed by DIGIFORCE Proefschriftmaken.nl, Wageningen, the Netherlands