Accepted manuscript

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

Liangzi Zhang [1,2], Hendriek Boshuizen[1,2] and Marga Ocké [1,2,]*

[1]   National Institute for Public Health and the Environment (RIVM), Bilthoven 3721 MA, The Netherlands;

[2]   Division of Human Nutrition, Wageningen University & Research, Wageningen 6708 PB, The Netherlands;

*   Correspondence: Marga Ocké; PO Box 1, 3700 BA Bilthoven; marga.ocke@rivm.nl; Tel.: +31302743814.

Short title：Recipe Simplification and Nutrient Intakes

Accepted manuscript

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.

Accepted manuscript

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.

Accepted manuscript

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].

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/m²) and males (26.0 kg/m²) were similar with those for girls (18.1 kg/m²) and females (26.6 kg/m²), 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.

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.

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.

Authorship

MO formulated the research question, MO, HB and LZ designed the study, LZ carried out the simulation study, LZ analysed data, HB provided guidance on methodology and data analysis, LZ wrote the manuscript, MO and HB edited the manuscript, LZ had primary responsibility for final content. All authors have read and approved the final manuscript.

Accepted manuscript

Figures
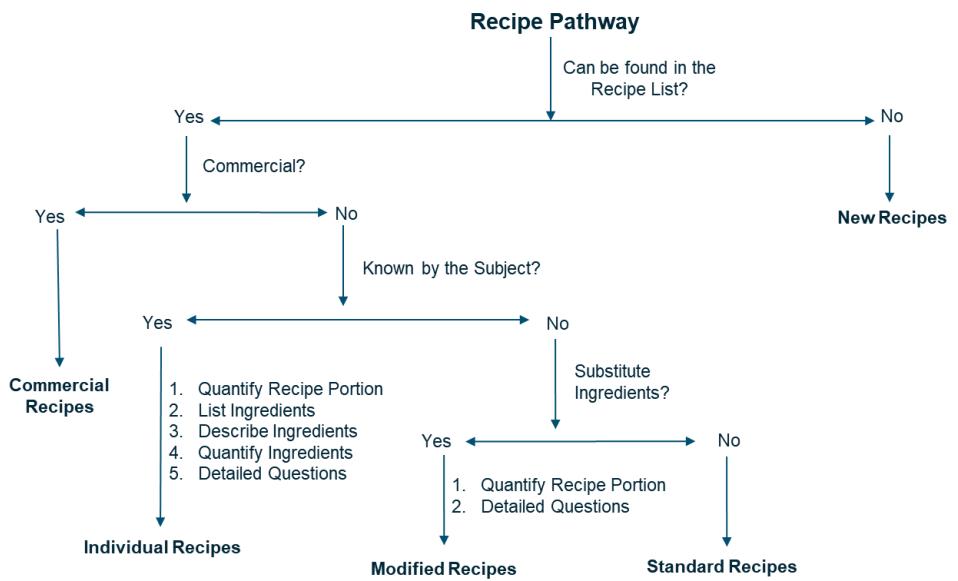
Figure 1. The flow chart of the mixed meal pathway in GloboDiet. Dishes were defined as homemade dishes if they could be found in the pre-defined recipe list and were not derived from commercial sources. Individual recipes were defined when people knew the information, they could substitute the predefined 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.

Accepted manuscript
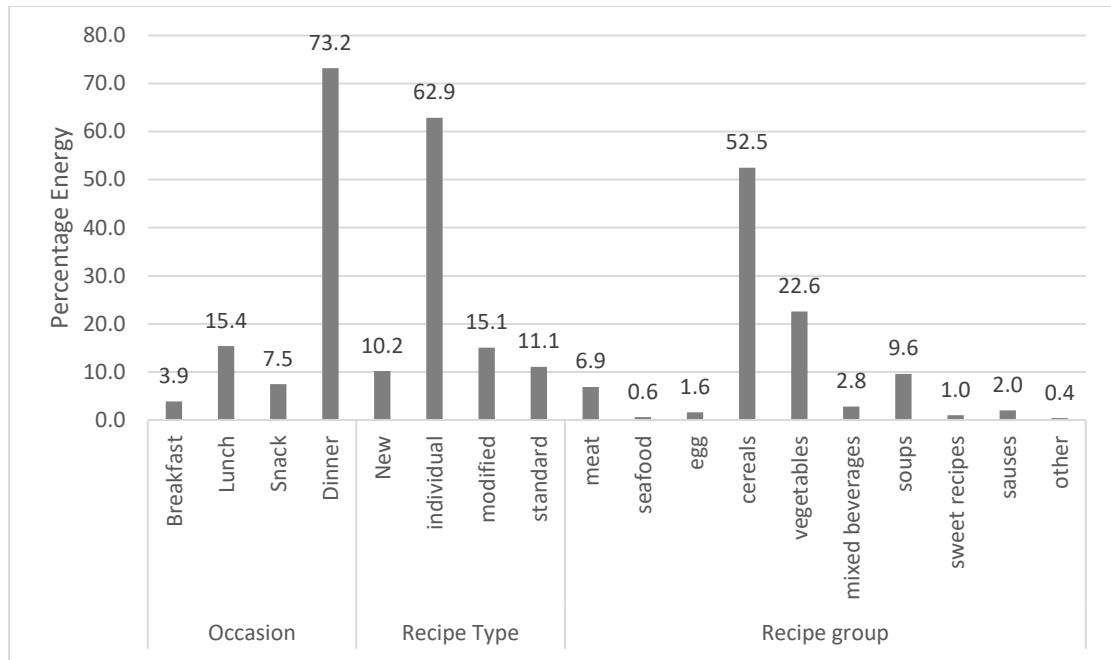
Figure 2. Energy consumed through mixed dishes partitioned (%) by different occasions, recipe types and recipe groups from the Dutch National Food Consumption Survey 2012-2016.
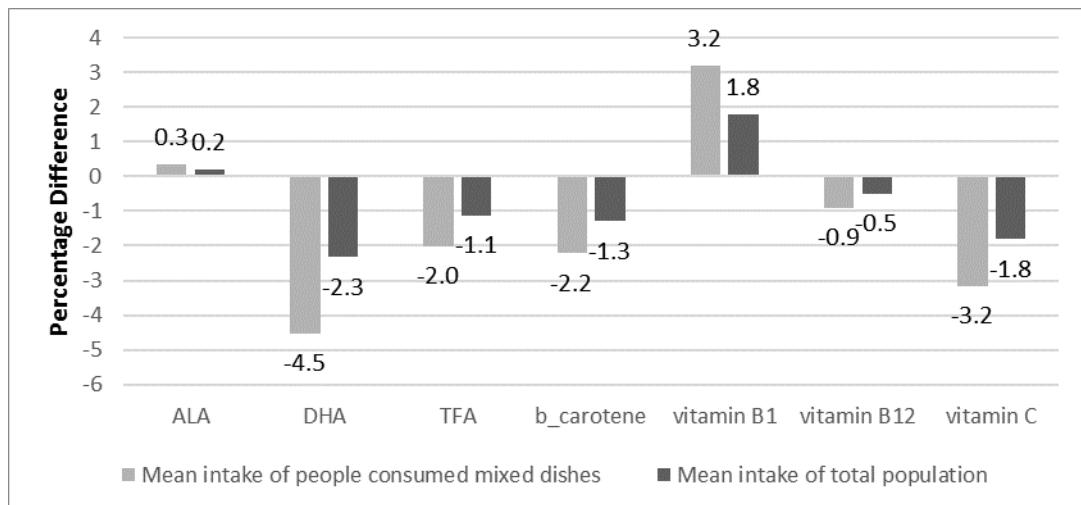
Figure 3a. The percentage difference of the mean intake of 7 nutrients of the total population and within people who consumed mixed dishes between the combined scenario and the original dataset.
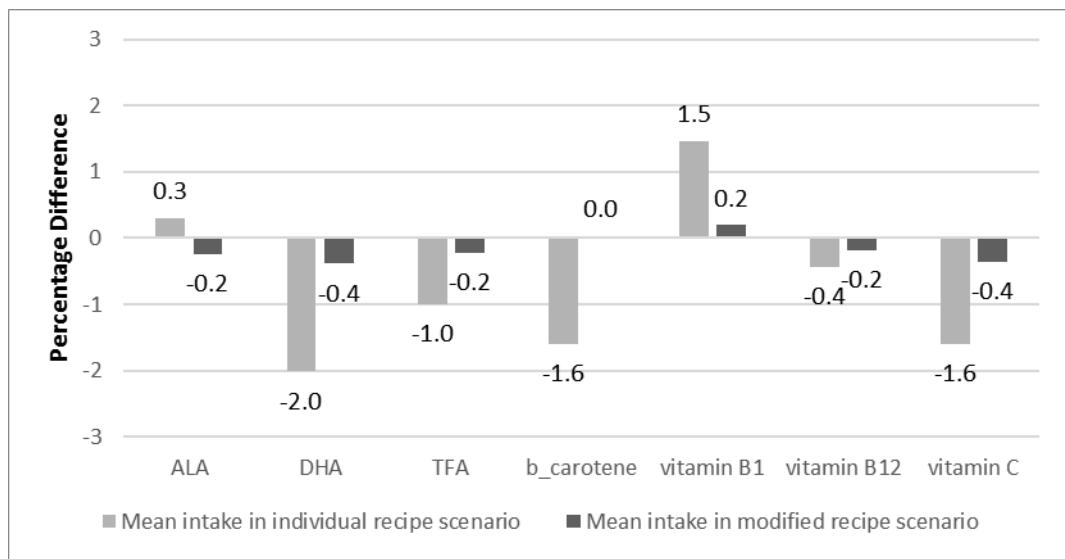
Figure 3b. The percentage difference of the mean intake of 7 nutrients of the total population between each scenario and the original dataset.

Accepted manuscript

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.

Accepted manuscript

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.

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.

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

Nutrients	Percentage Difference (%)						Amount Difference		
	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

Accepted manuscript

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.