Exploring Meal-Based Dietary Intake Assessment: Development of Statistical Analysis Strategies and Development of a Meal-Based Dietary Intake Assessment Tool

Cathal O'Hara

Student Number 19209325

This thesis is submitted to University College Dublin in fulfilment of the requirements for the degree of Doctor of Philosophy in the School of Agriculture and Food Science.

Head of School: Prof. Frank Monahan
Supervisor: Prof. Eileen Gibney
Research Studies Panel: Dr. Claire Timon, Prof. Nial Friel, and Dr. Aifric O'Sullivan
Submitted in January 2024

PRELIMINARY INFORMATION

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

24HR	24-hour recall
AIC	Akaike information criterion
ANOVA	Analysis of variance
ASA24	Automated Self-Administered 24-Hour Dietary Assessment Tool
BIC	Bayesian information criterion
BMI	Body mass index
BMR	Basal metabolic rate
BW	Body weight
CDC	Center for Disease Control and Prevention
CI	Confidence interval
CL	Cluster
CoFID	Composition of Foods Integrated Dataset
CRP	C-reactive protein
DASH	Dietary Approaches to Stop Hypertension
DLW	Doubly labelled water
DR	Diet record
DXA	Dual energy x-ray absorptiometry
EDTA	Ethylenediaminetetraacetic acid
EFSA	European Food Safety Authority
EI	Energy intake
EU	European Union
FCQ	Food combination questionnaire
FFQ	Food frequency questionnaire
FNDDS	Food and Nutrient Database for Dietary Studies
FTO	Fat mass and obesity-associated gene
GraFFS	Graphical Food Frequency System
HDL	High density lipoprotein
HEI	Healthy Eating Index
HFSS	High fat, sugar, or salt
HR	Hazard ratio
LC	Latent class
LCA	Latent class analysis
LDL	Low density lipoprotein
LOA	Limits of agreement
MMM	MyMealMate

MRC	Medical Research Council
MTHFR	Methylene tetrahydrofolate reductase
MUFA	Monounsaturated fatty acid
NANS	National Adult Nutrition Survey
NDNS	National Diet and Nutrition Survey
NHANES	National Health and Nutrition Examination Survey
NHS	National Health Service
NRF	Nutrient rich foods
OR	Odds ratio
PAM	Partitioning around the medoids
PC	Principal component
PCA	Principal component analysis
PR	Prevalence ratio
PUFA	Polyunsaturated fatty acid
RI	Reference intake
RTÉ	Raidió Teilifís Éireann
SD	Standard deviation
SFA	Saturated fatty acid
TEI	Total energy intake
UCD	University College Dublin
UK	United Kingdom
USA	United States of America
UV	Ultraviolet
WHO	World Health Organization
WHR	Waist-to-hip ratio

GENERAL ABSTRACT

There is limited investigation of meal patterns in nutrition research, that is, meal-based aspects of dietary intake including the timing or distribution of meals across a day, the combinations of foods consumed within meals, or the combinations of meals over a given period. Meal pattern research has the potential to further our understanding of the link between diet and disease and complement existing food-based dietary guidelines and inform personalised nutrition. Despite an increasing research focus on meal patterns in recent years, there is no research comparing the various statistical techniques used to identify meal patterns, limited research on the relationship between meal patterns and health, and an absence of a meal-based method of dietary intake assessment. The objectives of this thesis were therefore to review existing meal pattern research, refine the framework for meal pattern research, identify relationships between meal patterns and health, and develop a new meal-based method of dietary intake assessment.

Analysis of data from the Irish National Adult Nutrition Survey (NANS) carried out for this thesis condensed the 27336 meals consumed by participants into 63 generic, or characteristic, meals that provide an accurate estimate of nutrient intakes. Using data from the National Health and Nutrition Examination Survey (NHANES) in the USA, this process was reproduced, and further analysis was conducted to identify the different combinations of generic meals consumed by the participants (meal patterns). This was carried out using three different statistical approaches for comparison: partitioning around the medoids clustering, principal component analysis (PCA), and latent class analysis (LCA). Differences arose among the methods with respect to the number of patterns identified, the identification of meal skipping in the patterns, and the different combinations of generic meal that were consumed. An exploratory analysis of the meal patterns arising from clustering identified differences among the meal patterns for diet quality but not for health variables. A novel meal-based dietary assessment tool, based on the generic meals identified, was found to have moderate agreement with 24-hour recalls in estimating nutrient intakes, with stronger agreement for some nutrients compared to others.

In conclusion, a reproducible clustering approach can identify generic meals in dietary intake datasets. Meal patterns can be identified using different statistical methods; however, differences arise in the meal patterns identified by different methods, and their strengths and weaknesses should be considered when choosing a method. Some associations between meal patterns and diet quality were identified. A meal-based method of dietary intake assessment is a feasible method of collecting dietary intake data. Further research is required to further develop this tool and improve accuracy across a range of nutrients.

STATEMENT OF ORIGINAL AUTHORSHIP

I hereby certify that the submitted work is my own work, was completed while registered as a candidate for the degree stated on the Title Page, and I have not obtained a degree elsewhere on the basis of the research presented in this submitted work.

_____

Cathal O'Hara

COLLABORATORS

Professor Eileen Gibney secured the funding to conduct the research described in this thesis, developed the initial research questions, and supervised and provided feedback on all aspects of the research process. In addition to Prof. Gibney, the research studies panel that oversaw the PhD research (comprised of Dr. Claire Timon, Prof. Nial Friel, and Dr. Aifric O'Sullivan) provided feedback on research plans, analysis, and writing.

PhD candidate, Cathal O'Hara, carried out the planning, organisation, and day-to-day operation and management of the research described in this thesis. This included the development and implementation of study methods, plans, and schedules; the development of data analysis plans, data collection, and data analysis; and writing, editing, and submission of manuscripts for publication.

Three chapters in this thesis describe the secondary analysis of datasets that had previously been created by other research groups. Data from the National Adult Nutrition Survey (NANS) were used in Chapter 4, having been first collected from 2008 to 2010 by a team of researchers from the Irish Universities Nutrition Alliance. Data from the National Health and Nutrition Examination Survey (NHANES) were used in Chapters 5 and 6 having been originally collected during 2017 and 2018 by a team of researchers from the National Center for Health Statistics in the USA.

PUBLICATIONS ARISING FROM THIS THESIS

Peer-Reviewed Journal Articles

O'Hara C and Gibney ER (2024) Dietary Intake Assessment Using a Novel Generic Meal-Based Recall and a 24-Hour Recall: A Comparison Study. *J Med Internet Res*, **26**, e48817.

O'Hara C, O'Sullivan A, and Gibney ER (2022) A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data. *J Nutr*, **152**(10), 2297-2308

O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

Conference Publications

O'Hara C and Gibney ER (2023) A Comparison of Different Methods for Meal Pattern Analysis [Oral Presentation]. *14th European Nutrition Conference FENS 2023*. 14–17 November, Belgrade, Serbia.

O'Hara C and Gibney ER (2023) Comparison of a Novel Meal-Based Method of Dietary Assessment and a 24-Hour Recall [Oral Presentation]. *International Conference on Diet and Activity Methods (ICDAM)*. 26–29 June, Limerick, Ireland.

O'Hara C and Gibney ER (2023) An Exploratory Analysis of Meal Patterns in NHANES [Oral Presentation]. *The Nutrition Society Irish Section 32nd Annual Postgraduate Conference*. 8–10 February, Cork, Ireland.

O'Hara C and Gibney ER (2022) Use of Generic Meal Images to Assess Dietary Intakes [Poster]. *The International Union of Nutrition Societies' 22nd International Congress of Nutrition (IUNS-ICN)*. 6–11 December, Tokyo, Japan.

O'Hara C and Gibney ER (2022) A Meal-Based Approach to the Classification of Individuals as Low, Adequate, or High Consumers of Nutrients [Poster]. *The Nutrition Society Irish Section Conference*. 15–17 June, Cork, Ireland.

O'Hara C and Gibney ER (2022) Meal Patterns and Image-Based Dietary Assessment [Oral Presentation]. *The Nutrition Society Irish Section 31st Annual Postgraduate Conference*. 9–11 February, Portrush, United Kingdom.

ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor, Prof. Eileen Gibney, for her support. Eileen gave me the time and space to explore and learn but was always there to provide clarity and guidance when that exploration and learning became muddled. I owe her a huge debt of gratitude for the numerous opportunities she has given me during my PhD. More importantly, however, Eileen has been an incredibly friendly, fun, and thoughtful person to work with over the past four years.

Both Eileen and the other members of my Research Studies Panel—Dr. Aifric O'Sullivan, Dr. Claire Timon, and Prof. Nial Friel—gave so generously of their expertise and their time, providing essential advice and alternative perspectives on my work. The depth and breadth of their knowledge and their ability to challenge me to think about things in different ways has enhanced my work and, by extension, this thesis immeasurably. I am most thankful to them for all their help.

I am incredibly grateful to all those that I had the good fortune to share an office with during my PhD. Their friendship, humour, and kindness (whether I wanted it or not!) were a consistent source of joy. Without them, this PhD would have been an altogether bleaker experience.

My friends deserve special thanks for giving me regular breaks from all things PhD related with a steady supply of chats, laughs, craic, and adventures. On top of all that, they also regularly went out of their way to ensure that I stayed grounded despite my lofty academic aspirations!

My family—Stephen, Attracta, Stephen B., and Caoimhe—have been an unwavering source of support and inspiration throughout my life. Their encouragement and optimism gave me the motivation to return to education in pursuit of this PhD. They remain the greatest supporters of all my endeavours, however unconventional they may be! Words cannot describe how grateful I am to have them in my life.

Gabhaim buíochas ó chroí libh go léir.

CHAPTER 1

Introduction

1.1 Background

Suboptimal dietary intake contributes to the global burden of chronic non-communicable diseases including cardiovascular disease, type 2 diabetes, and certain cancers [1]. Traditionally, diet-disease epidemiological studies have focused on linking disease, or risk of disease, with individual nutrients, individual foods and food groups, or dietary patterns (i.e. the combinations of foods or food groups consumed habitually that represents the diet as a whole) [2-5]. These approaches have revealed valuable insights and have contributed to the development of both nutrient and food-based dietary guidelines [2, 6, 7]. For example, food-based dietary patterns associated with a reduced risk of cardiovascular disease include the Mediterranean [8] and Dietary Approaches to Stop Hypertension (DASH) [9] patterns. Among other features, both encourage consumption of fruit, vegetables, and wholegrains etc., while discouraging processed meats, salt, and sugar-sweetened beverages etc., but do not place any emphasis on specific meals.

While the food-based approach is typical of most guidelines [6], those in the United States of America (USA) and Brazil acknowledge the potential for meal-based information to complement what is already known about nutrients, foods, and dietary patterns by referring to the context in which meals are consumed, the frequency or number of meals consumed [10], and the combinations of foods that can be used to make up an individual meal [11]. However, it is highlighted that there is a need for further research on the place of meals in food-based guidance to provide a stronger evidence base to support such guidelines [10].

The area of meal-based research is complex and in its infancy. To start, there are no widely agreed definitions of a meal. People consume individual foods or drinks, either alone or in combination, as part of eating occasions, which are often described as meals [12]. Various approaches have been taken in the literature, basing definitions on participants' own description of eating occasions using labels such as breakfast, lunch, dinner, or snack; the time of day that eating occasion occurred; the time that elapses between meals; the energy content of a meal; or in some instances based on a combination of the above approaches [12, 13]. Meals are an integral part of our dietary intake, and even though they are composed of individual foods, people typically conceptualise their dietary intakes as a collection of meals rather than a collection of foods or nutrients. Despite this, dietary intake data are mostly captured at the food or food group level, i.e., individuals are asked to report the individual foods or food groups that they consume [14, 15].

The assessment of dietary intake is an essential element of identifying diet-disease associations, as well as being a crucial part of developing and providing both healthy eating guidelines for the general public and personalised advice to support appropriate change of

dietary behaviours for individuals [16, 17]. The essence of dietary intake assessment is the recording or monitoring of what people eat and drink either prospectively or retrospectively, and either at a food, food group, or meal level [16, 18, 19]. It has been proposed that a meal-based approach may be useful for the collection of information about dietary intakes, for analysing and interpreting those data, and for the provision of dietary advice for personalised nutrition. A meal-based approach to assessment and interpretation of dietary intakes could also support the use of internet and mobile technology to collect dietary information thereby reducing user burden with regard to data input [20, 21]. A focus on meal level information in dietary intake assessment might allow for improved meal-based dietary advice and enhance existing food-based dietary guidelines, with practical benefits for individuals relating to meal planning and preparation [22]. However, in order to develop meal-based methods of dietary assessment, it is first necessary to identify what meals and meal patterns are consumed in a given population [20, 23].

1.2 Meal Patterns

Meal patterns are typically categorised into one of three constructs (**Figure 1.1**). Temporal patterns refer to the timing and distribution of meals [24-29], content patterns refer to the combinations of foods within a meal or the combinations of those meals over a day [30-33], and context patterns refer to the external elements of the meal such as location, activities while eating, and the presence or absence of others at the meal [34]. It is also possible to examine combined patterns which explore multiple meal pattern constructs simultaneously [34].

The meal pattern research and analysis within this thesis will focus primarily on content meal patterns. In order to derive content meal patterns, the vast number of unique meals consumed by participants in a given study must first be condensed to a smaller but "representative" number that acts as a qualitative (i.e., the types of foods in the meals) and a quantitative (i.e., the nutritional content of the meals) summary of the meals consumed by a given sample population. These representative meals are often termed generic meals, as they are a summary, or combination, of many meals rather than a specific description of an individual meal. Content meal patterns, therefore, refer to the various combinations of these generic, or representative, meals that are consumed over a day [30].

Content patterns have provided new insights to the combinations of foods that form meals and the combinations of meals over the course of a day [30-32, 35, 36]. For example, two separate studies in the same Irish cohort found similar meal patterns while using different statistical methods. Both studies identified patterns consisting of a cereal/toast breakfast, a

sandwich-based light meal, and protein-carbohydrate-based main meal [30, 32]. In a separate Japanese cohort the majority of breakfast meals were vegetable- and rice-based [31], yet none of the breakfasts in the Irish cohort were likely to contain these foods [30, 32]. Bread-based breakfasts, on the other hand, were consumed in both cohorts. Summaries of studies examining various meal pattern constructs are given in **Table 1.1**.

Uzhova *et al.* [32] went on to identify an association between a weekend meal pattern consisting of a cooked breakfast, a skipped light meal, and a protein-carbohydrate main meal and increased likelihood of higher diastolic blood pressure compared to a meal pattern consisting of cereal and/or toast for breakfast, sandwich light meal and protein-carbohydrate main meal. However, other studies have not examined the extent to which content meal patterns may or may not influence health. In studies of temporal meal patterns the most consistent finding is the association between a pattern consisting of three meals per day and high dietary quality or positive health outcomes [24, 25, 29, 37, 38]. No study, however, has examined combined meal patterns incorporating both temporal and content constructs. It is necessary to establish what relationships exist between meal patterns and health to determine their role in complementing existing food-based dietary advice [10, 11].

While advances have been made in the identification of and use of meal patterns in nutrition research, some limitations still exist. Current generic meal approaches do not account for the portion size of meals and consider only the food groups that make up those meals [30, 31]. In previous research, pre-existing food groups from non-meal-related research topics have been used rather than creating food groups for the specific research question at hand as previously suggested by Newby *et al.* [39]. In addition, the generic meals identified in previous studies did not include information on portion size, i.e., all participants were assumed to have consumed the same portion size. To identify generic meals in those previous studies, the original meals consumed by individuals were grouped, using the frequent itemsets data mining method, based on whether they were similar with regard to the food groups that they contained; the similarity of the meals with regard to nutrient profiles was not considered [30, 31]. Further investigation is required to overcome these gaps in meal pattern research to refine the analytical approach so that consideration is given to food groups specific to meal-patterns, portion sizes, and the nutrient profile of meals.

Furthermore, to date different statistical techniques have been used to identify meal patterns. For example, latent class analysis (LCA) [32] and principal component analysis (PCA) [30, 31] have been used to identify content meal patterns, while clustering has been used to identify temporal and context meal patterns [24, 34] (**Table 1.1**). In the related field of dietary patterns, comparison studies and literature reviews have identified that, while

differences exist across studies, broadly speaking there is consistency in the type of dietary patterns identified when using different statistical techniques [4, 40-43]. Despite this, however, no study has examined the extent to which the statistical approach chosen influences the outcome for meal pattern research. This type of comparison would yield important methodological insights, as has been done for the more frequently researched area of dietary patterns, where combinations of foods or food groups are used to describe intakes. This knowledge gap in meal patterns will need to be addressed to understand the strengths and limitations of this approach to dietary data analysis and whether it represents a feasible approach for future use in dietary intake studies, dietary intake assessment, and population or individual dietary recommendations.

Finally, despite the progress that has been made in meal-based research in recent years, the identification of meal patterns still relies on traditional food-based methods of dietary intake assessment. For example, the meal pattern studies described above use a variety of methods to collect the dietary data that they use to identify meal patterns such as 24-hour recalls [24, 27, 28] and diet records [30-33]. No study has been identified that provides a meal-based method of dietary intake assessment capable of assessing content meal patterns.

1.3 Assessing Dietary Intake

In considering how to use dietary intake data to determine meal patterns it is important to reflect on the methods of assessment and their limitations. There are various well-established methods used to assess dietary intakes, each with their own strengths and limitations. Objective methods include the duplicate diet method and the food consumption record. Subjective methods include the 24-hour recall (24HR), diet history, diet record (DR), and food frequency questionnaire (FFQ) [16, 18, 19] (**Table 1.2**).

Objective methods of assessing diet(s) involve observation of an individual's dietary intake by a trained researcher. Specifically, the duplicate diet approach involves collecting duplicate samples of each food or drink consumed, while the food consumption record involves observing individuals as they prepare and consume food and drink while recording intakes. These methods are considered more costly and time consuming than other methods and are therefore not used commonly in practice [16, 18, 19].

Subjective methods (24HR, DR, and FFQ) rely on individuals themselves to provide the information about their own dietary intake. The 24HR involves an individual being interviewed regarding their dietary intake from the previous day or the preceding 24 hours

[18, 19]. During a 24HR, detailed information is gathered regarding the foods and drinks consumed, brands of foods consumed, ingredients used, and cooking methods. Portion sizes are usually estimated based on common household measurements (such as tablespoons and cups etc.), standard measuring cups or spoons, food models, or food photographs [44, 45]. The 24HR places a relatively low burden on participants, but the quality of the information gathered depends greatly on their ability to recall their intakes [18, 19]. The diet history is a more detailed extension of the 24HR where participants are asked about their habitual dietary intake rather than solely their intakes in the preceding day [16, 46].

The DR method for assessing intakes does not depend on recall ability, as participants are asked to record their own intakes at time of consumption, usually in the form of a food diary. Portion sizes are typically estimated based on common household measurements or by weighing of the individual foods of a meal. DRs are typically carried out for a duration of between 3 and 7 days**.** This method, however, places a high time and effort burden on participants for recording of information [47, 48]. Due to the open-ended nature of the 24HR and DR a wide range of information can be collected from a wide range of individuals from differing cultures and with differing dietary habits. However, this also increases the time and effort required to collect, interpret, and analyse the information. Both methods focus on short-term dietary intake rather than habitual intake. Multiple recalls or records may be carried out to determine habitual intake, however, this is resource intensive and may influence participants' intakes [19, 49] .

A Food Frequency Questionnaire (FFQ) presents participants with a list of foods and/or beverages that are indicative of foods consumed and asks the respondent to select from pre-determined options how often and how much they ate of each item over a given time period. Being self-administered and containing closed questions, this method is considered relatively simple and resource efficient. However, these tools need to be developed and validated specifically for their intended use, as the purpose of the dietary assessment and the population whose dietary intakes are being assessed will influence the content of an FFQ [50-52] (**Table 1.2**).

The choice of dietary intake assessment method, within any research, public health, or clinical use, must ultimately consider all features of the various methods while accounting for the purpose of dietary assessment and the resources and skills of both the participant and the researcher [53, 54]. Even when these factors are considered, it is important to be aware of the limitations of the various methods and the data arising from them.

1.4 Limitations of Dietary Intake Assessment

Whilst these methods are commonly used, and they provide important information for nutrition research, nutrition policy, and other areas, they all have known limitations. Subjective methods of dietary intake assessment are subject to both random and systematic error [19]. Random error can be minimised through the use of adequate sample sizes and repeated measures. Systematic error, on the other hand, is more difficult to account for [55]. Subjective methods of dietary intake assessment are thought to typically underestimate energy intake based on comparisons with doubly labelled water (DLW). The degree of underestimation varies according to the method, with a recent systematic review finding underestimates of between 11% to 41% for diet records, 8% to 30% for 24HRs, and 4.6% to 42% for FFQs [56]. The reasons for this are numerous and include, poor memory of intakes, difficulties estimating portion sizes, underreporting, undereating, and inaccuracies present in food composition databases [55].

While underreporting is particularly prevalent for energy intakes, the assessment of intakes of other nutrients such as protein and potassium tend to be somewhat more accurate, with underestimates of protein intake of 5% have been reported for the 24HR and of 11% to 22% for the FFQ [57]. Underestimates of potassium intake of 0% to 4% being reported for the 24HR and 5% to 6% for the FFQ [58]. The values reported above refer to values that are not energy adjusted, i.e., they represent differences between the dietary assessment methods and urinary biomarkers for actual intakes of the nutrients and not the ratio of nutrient intake to energy intake. The 24HR and the FFQ tend to overestimate energy adjusted values of protein and potassium due to the underestimation of energy intake [57, 58].

In more recent years, the use of metabolomics is being considered to mitigate against some of the limitations associated with subjective methods of dietary intake assessment. The objective measurement of metabolites in saliva, blood, or urine can be used as biomarkers to estimate or benchmark dietary intakes of various nutrients, foods, food groups, or dietary patterns [59]. However, the use of metabolites as biomarkers of dietary intakes also has limitations. A limited number of metabolites have a direct relationship with the absolute intakes of the nutrients for which they act as biomarkers; 24-hour urinary sodium, potassium, and nitrogen are well established and accurate biomarkers of intakes of sodium, potassium, and protein respectively [60]. This is not the case, however, for other metabolites that have been investigated to date. A recent review identified 347 urinary metabolites reported for 67 different foods or food groups, indicating a lack of consensus on the most appropriate metabolites to act as biomarkers for many foods [61]. Many of the metabolites

that have been identified are not specific to a single food group. For example, urinary hippurate has been identified as a biomarker for tea intake, fruit intake, vegetable intake, and cocoa intake [61]. One potential solution to improve specificity is to identify combinations of metabolites associated with intakes of particular foods or dietary patterns, however, this will give rise to a more labour intensive and costly process [60, 61].

Given the high costs and the absence of sensitive and specific biomarkers for many foods and dietary patterns these methods are not suitable to replace existing methods of dietary assessment. Their value instead lies in their use as a complementary tool to those methods to validate or calibrate subjective measures of dietary assessment, or to estimate adherence to dietary interventions [59, 62, 63].

1.5 Technology and Dietary Intake Assessment

While the strength of using biomarkers may be in combination with existing methods, the use of digital technologies aims to reduce the cost and burden of dietary intake assessment and expand its potential reach [64]. Advances in digital technology have allowed for FFQs, 24HRs, and DRs to be adapted for use in computer-based, web-based, and mobile application formats, allowing large parts of the processes to be automated (**Table 1.3**). The approach of these methods is the same as their paper-based counterparts, but they are administered by computer technology rather than by a person [64]. These digital formats show good agreement with their traditional counterparts with differences in estimated energy intake between the two formats ranging from 0% to 10%. Their accuracy as judged by comparison with DLW is also similar to that seen for the traditional methods; they also tend to provide underestimates of energy intake ranging from 6.4% to 36.3% [64].

Despite being subject to similar error seen with the traditional methods discussed above, the digital methods have additional benefits including supporting automated or semi-automated analysis and providing rapid feedback to users. These methods therefore are typically less resource intensive for researchers compared to manual methods [64-67]. These new methods do, however, present their own unique challenges. Much of the burden for recording dietary intake using digital 24HRs and DRs is placed on the user [68]. In order to record their own dietary intake, these methods require users to search through food lists derived from food composition tables and select the specific food they consumed. This step is then repeated for each food at each meal throughout the duration of the recording period [69].

One solution postulated to reduce the burden of data input is the use of image-based dietary assessment, i.e., the use of image recognition software to automatically identify foods and portions sizes from individuals' photographs of their dietary intake [70-73]. The complex process that facilitates this can be described in four broad steps. Firstly, the software segments the image to differentiate non-food from food items and to differentiate among the different food items present [74]. Secondly, the individual food segments are then classified as one of the foods from a database. Thirdly, the portion size of the foods is estimated based on the estimation of their volume. Finally, the nutritional content of the meal is estimated by combining the information in the previous steps with the relevant data from a food composition database [74, 75].

Accurately identifying foods from images using computer vision is challenging given the vast variety of foods in meals which is further complicated by the fact that some foods or ingredients may be hidden and that the same food can have different appearances in different contexts [74, 76]. The accuracy of in correctly classifying food images has been reported to range from 54.7% to 93.0% [76] depending on the input variables used to train the image recognition algorithm, the number of foods that it is required to recognise, and the type of algorithm used. The most accurate methods are computationally intensive, require large datasets, high numbers of variables, and have long running times for training [74]. The variable accuracy with these systems places the burden back on the user rather than on the technology [67, 70]. During the classification and portion size estimation steps, the user will be prompted to review for accuracy and amend as necessary [70]. The foods that are incorrectly identified must be amended by the user by searching through a food list manually. To date, all image-based dietary assessment methods have aimed to imitate the already existing paper-based methods by being food-based. That is, they aim to segment a meal and identify the individual food components of that meal. Consideration has not been given to meal-based dietary intake assessment in general or its potential role in image-based dietary intake assessment.

1.6 Meal-Based Dietary Intake Assessment

The traditional methods of dietary intake assessment, and their digital versions, aim to capture the most accurate information possible required for epidemiological studies. This detailed food-based approach is unlikely to be the most appropriate solution for personalised nutrition given the resources required and the burden for collection of data that is placed on the end user [20]. A generic meal method similar to that that has previously been used in meal pattern analysis [30], but with portion size and a greater focus on nutrient

composition incorporated, could facilitate a meal-based approach to dietary intake assessment whereby individuals are presented with images of the generic meals for each meal type and choose the one that most represents their intake. While it unlikely that this approach could be used to provide point estimates of nutrient intakes, it is conceivable that it could categorise individuals based on whether their intakes were adequate, high, or low according to nutrient-based dietary guidelines [20]. This approach of categorising intakes has previously been used to personalise the nutrition feedback messages that an individual receives following dietary assessment depending on whether they are adequate, high, or low consumers of certain nutrients [77].

A meal-based approach need not exist in isolation from other methods of dietary intake assessment, but rather form part of a suite of methods used interchangeably to provide the best outcome for each user depending on their preferences and degree of motivation. While the more burdensome methods of dietary intake assessment may have a role to play in habit-forming during initial engagement with personalised nutrition technologies, allowing users to change from one dietary assessment method to another may aid in maintaining such engagement over a prolonged period [78].

Despite the number of studies examining meal patterns increasing in recent years [22], only three have developed a meal-based method of dietary intake assessment rather than using existing food-based methods [25, 26, 79]. Two of these methods, however, while simple to complete, only provide information on the meal types consumed and their timing and do not provide sufficient detail to determine the different combinations of foods being consumed in meals, the different combinations of meals being consumed over a day, or the nutrient intakes arising from that consumption [25, 26]. The third method involves participants reporting the frequency of consumption of combinations of food groups and foods at specified meal types (breakfast, morning snack, lunch, afternoon snack, dinner, night snack). This approach does allow for the identification of meal patterns and nutrient intakes, but still requires individuals to report intakes at the food level [79]. None of the previous studies, however, allow for the reporting of meal portion sizes or capture information from the previous 24 hours in the form of a recall. No study has been identified that allows individuals to record their dietary intakes at the meal level by reporting intakes of whole meals rather than individual foods or food groups. Further research is therefore required to determine the feasibility and accuracy of a meal-based method for dietary intake assessment.

This thesis will aim to address the research gaps that have been highlighted in this introduction. Specifically, the lack of a critical review in the area of advanced statistical

techniques in meal pattern research; the absence of the use of portion sizes and nutrient profiles of meals when identifying generic meals; the lack of knowledge regarding comparability of different statistical techniques for identifying meal patterns; the limited information regarding relationships between content meal patterns and health; and the lack of a meal-based tool with which to collect dietary intake data. These gaps will be addressed through the following aims and objectives of this thesis.

1.7 Aims

To further develop statistical approaches to meal pattern analysis for use in a meal-based dietary intake assessment tool.

1.8 Objectives

1. Identify existing meal pattern research, research gaps, challenges, and opportunities.
2. Refine the framework for identifying generic meals and meal patterns for use in dietary intake assessment.
3. Identify relationships between adherence to identified meal patterns and health parameters.
4. Develop new meal-based method of dietary intake assessment and compare with established method.

1.9 Tables

Table 1.1: Summary of meal pattern studies and their findings.

Study	Population	Meal Pattern Construct	Meal Patterns
Chau *et al.* [24]	4,508 adults in Taiwan	Temporal	5 patterns based on dietary intake during 6 4-hour time periods.
Englund-Ögge *et al.* [25]	65,487 pregnant women in Norway	Temporal	3 patterns based weekly frequency of 8 meal types, e.g., breakfast, morning snack, lunch etc.
Wilson *et al.* [26]	1,304 adults in Australia	Temporal	3 patterns based on dietary intake during 7 time periods.
Leech *et al.* [27]	5,242 adults in Australia	Temporal	3 patterns based on dietary intakes in each hour of the day.
Khanna *et al.* [28]	7,565 adults in the USA	Temporal	4 patterns based on dietary intake at each hour of the day.
Khanna *et al.* [28]	7,565 adults in the USA	Temporal	4 patterns based on dietary intake at each hour of the day.
Woolhead *et al.* [30]	1,500 adults in Ireland	Content	12 patterns based on combinations of foods consumed at each meal.
Murakami *et al.* [31]	242 adults in Japan	Content	11 patterns based on combinations of foods consumed at each meal.
Uzhova *et al.* [32]	1,500 adults in Ireland	Content	4 weekday and 3 weekend patterns based on combinations of foods consumed at each meal.

Study	Population	Meal Pattern Construct	Meal Patterns
Hearty and Gibney [33]	1,379 adults in Ireland and the UK	Content	Daily patterns not given. 10 food combinations (generic meals) identified that were likely to predict whether an individual was in the first or fifth quintile for the Healthy Eating Index.
Riou *et al.* [34]	2,994 adults in France	Temporal and Context	5 patterns based on number and location of meals, and activities and others present during meals.

Table 1.2: Summary of traditional methods of dietary intake assessment.

Method	Features	Strengths	Limitations
		Objective methods	
Duplicate diet record	Duplicate samples of each food and drink consumed are collected and analysed.	• Does not rely on memory. • Food composition is specific foods consumed rather than averages from composition tables.	• Costly, time-consuming, and resource intensive. • Provides limited information for meals consumed outside the home. • Intrusive for participant.
Food consumption survey	Individuals are observed while preparing and consuming food and this information is recorded.	• Does not rely on memory. • Suitable for those with low literacy.	• Costly, time-consuming, and resource intensive. • Provides limited information for meals consumed outside the home. • Intrusive for participants.
		Subjective Methods	
24-hour recall (24HR)	Use of open-ended questions by a trained interviewer regarding dietary intake in the previous 24 hours or previous day.	• Can provide detailed dietary data. • Suitable for those with low literacy. • Relatively low participant burden.	• Subject to recall bias, social acceptability bias, and interviewer bias. • Multiple recalls required to assess habitual intake.

Method	Features	Strengths	Limitations
Diet history	An extension of the 24-hour recall where the interviewer also asks questions about habitual intake.	• Can assess habitual intake. • Suitable for those with low literacy.	• Subject to recall bias, social acceptability bias, and interviewer bias. • Can be burdensome on participant given heavy reliance on both long-term and short-term memory.
Diet record (DR)	Use of open-ended questions in self-administered questionnaires often referred to as food diaries.	• Not subject to recall bias. • Provides detailed dietary data. • No limited food lists.	• High participant burden may lead to underreporting. • Relies on participant literacy and numeracy. • Multiple days required to assess habitual intake.
Food frequency questionnaire (FFQ)	Questionnaire with closed questions on predefined foods or food groups.	• Can assess habitual dietary intakes. • Cost-effective and relatively resource efficient. • Low burden.	• Relies on participant literacy and numeracy. • Specific to given research question. • Questionnaires must be carefully designed and validated. • Subject to recall bias.

Table 1.3: Summary and examples of digital methods of dietary intake assessment.

Method	Features	Examples
Food Frequency Questionnaire (FFQ)	• The FFQ can be self-administered by the participant via a website or mobile app. • Participants choose from a list the food groups they consumed during the time period in question. They are then presented with follow up questions regarding portion size and frequency of intake. • Portion sizes can be based on household measures or food photographs. • Preliminary data analysis, such as estimated food group intakes, can be automated.	• Food4Me [80] • Graphical Food Frequency System (GraFFS) [81] • Web-FFQ [82]
24-Hour Recall	• Can be self-administered by the participant via a website or mobile app. • The back end contains a food list with associated nutrient compositions. • Participants text search for the foods that they consume and select the most appropriate option from the search results. • Portion sizes can be based on household measures or food photographs. • Preliminary data analysis, such as estimated dietary intakes, can be automated.	• Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA24) [83] • Foodbook24 [69] • Intake24 [84]
Diet Record	• Can be self-administered by the participant via a website or mobile app. • The back end contains a food list with associated nutrient compositions. • Participants text search for or scan the barcode of the foods that they consume and select the most appropriate option from the search results. • Portion sizes can be based on household measures or food photographs. • Preliminary data analysis and outputs, such as estimated intakes of food groups or nutrients, can be automated.	• MyMealMate (MMM) [85] • Mijn Eetmeter [86]

1.10 Figures

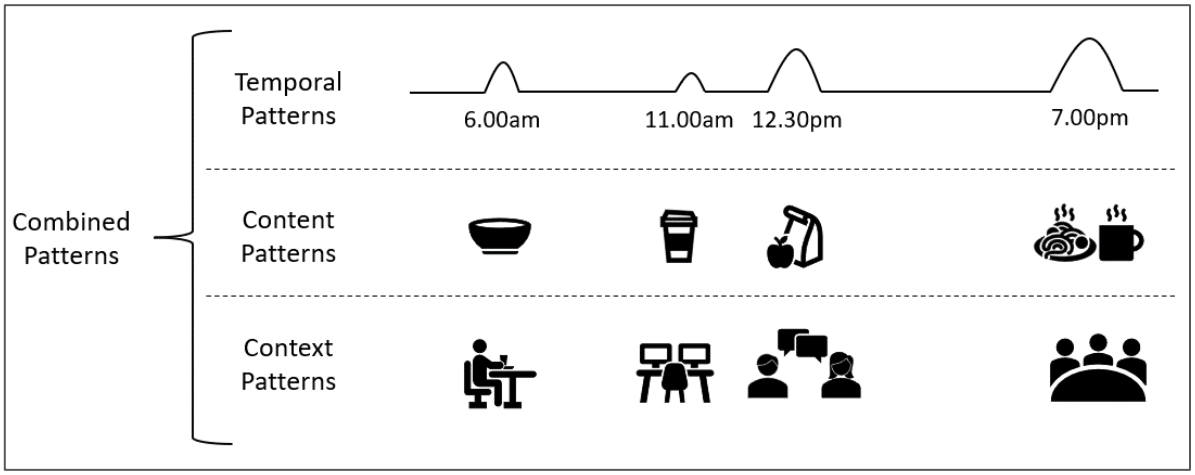


Figure 1.1: Visual representation of the three meal pattern constructs.

1.11 References

1. Afshin A, Sur PJ, Fay KA*, et al.* (2019) Health Effects of Dietary Risks in 195 Countries, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *The Lancet*, **393**(10184), 1958-1972.

2. Liang S, Mijatovic J, Li A*, et al.* (2022) Dietary Patterns and Non-Communicable Disease Biomarkers: A Network Meta-Analysis and Nutritional Geometry Approach. *Nutrients*, **15**(1), 76.

3. Tapsell LC, Neale EP, Satija A*, et al.* (2016) Foods, Nutrients, and Dietary Patterns: Interconnections and Implications for Dietary Guidelines. *Adv Nutr*, **7**(3), 445-454.

4. Jannasch F, Riordan F, Andersen LF*, et al.* (2018) Exploratory Dietary Patterns: A Systematic Review of Methods Applied in Pan-European Studies and of Validation Studies. *Br J Nutr*, **120**(6), 601-611.

5. Edefonti V, De Vito R, Dalmartello M*, et al.* (2020) Reproducibility and Validity of a Posteriori Dietary Patterns: A Systematic Review. *Adv Nutr*, **11**(2), 293-326.

6. Herforth A, Arimond M, Álvarez-Sánchez C*, et al.* (2019) A Global Review of Food-Based Dietary Guidelines. *Adv Nutr*, **10**(4), 590-605.

7. Bechthold A, Boeing H, Tetens I*, et al.* (2018) Perspective: Food-Based Dietary Guidelines in Europe—Scientific Concepts, Current Status, and Perspectives. *Adv Nutr*, **9**(5), 544-560.

8. Martinez-Gonzalez MA, Gea A, and Ruiz-Canela M (2019) The Mediterranean Diet and Cardiovascular Health. *Circ Res*, **124**(5), 779-798.

9. Chiavaroli L, Viguiliouk E, Nishi SK*, et al.* (2019) Dash Dietary Pattern and Cardiometabolic Outcomes: An Umbrella Review of Systematic Reviews and Meta-Analyses. *Nutrients*, **11**(2), 338.

10. Dietary Guidelines Advisory Committee (2020) Scientific Report of the 2020 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture and the Secretary of Health and Human Services. Washington, DC: *U.S. Department of Agriculture, Agricultural Research Service*.

11. Ministry of Health of Brazil (2015) Dietary Guidelines for the Brazilian Population. Brasília/DF: *Secretariat of Health Care, Primary Health Care Department*.

12. Leech RM, Worsley A, Timperio A*, et al.* (2015) Characterizing Eating Patterns: A Comparison of Eating Occasion Definitions. *Am J Clin Nutr*, **102**(5), 1229-1237.

13. Murakami K, Shinozaki N, Livingstone MBE*, et al.* (2020) Meal and Snack Frequency in Relation to Diet Quality in Japanese Adults: A Cross-Sectional Study Using Different Definitions of Meals and Snacks. *Br J Nutr*, **124**(11), 1219-1228.

14. Bisogni CA, Falk LW, Madore E*, et al.* (2007) Dimensions of Everyday Eating and Drinking Episodes. *Appetite*, **48**(2), 218-231.

15. Sproesser G, Ruby MB, Arbit N*, et al.* (2019) Understanding Traditional and Modern Eating: The TEP10 Framework. *BMC Public Health*, **19**(1), 1606.

16. Welch A (2019) Dietary Assessment. In *Manual of Dietetic Practice*, 6th ed., pp. 64-69 [J Gandy, editor]. Hoboken, NJ, USA: John Wiley & Sons, Inc.

17. Celis-Morales C, Livingstone KM, Marsaux CF*, et al.* (2017) Effect of Personalized Nutrition on Health-Related Behaviour Change: Evidence from the Food4Me European Randomized Controlled Trial. *Int J Epidemiol*, **46**(2), 578-588.

18. Shim JS, Oh K, and Kim HC (2014) Dietary Assessment Methods in Epidemiologic Studies. *Epidemiol Health*, **36**, e2014009.

19. Slimani N, Freisling H, Illner A-K*, et al.* (2015) Methods to Determine Dietary Intake. In *Nutrition Research Methodologies*, 1st ed., pp. 50-70 [JA Lovegrove, L Hodson, S Sharma and SA Lanham-New, editors]. Hoboken, NJ, USA: John Wiley & Sons.

20. Gibney MJ and Walsh MC (2013) The Future Direction of Personalised Nutrition: My Diet, My Phenotype, My Genes. *Proc Nutr Soc*, **72**(2), 219-225.

21. O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

22. Leech RM, Worsley A, Timperio A*, et al.* (2015) Understanding Meal Patterns: Definitions, Methodology and Impact on Nutrient Intake and Diet Quality. *Nutr Res Rev*, **28**(1), 1-21.

23. O'Hara C, O'Sullivan A, and Gibney ER (2022) A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data. *J Nutr*, **152**(10), 2297-2308.

24. Chau CA, Pan WH, and Chen HJ (2017) Employment Status and Temporal Patterns of Energy Intake: Nutrition and Health Survey in Taiwan, 2005–2008. *Public Health Nutr*, **20**(18), 3295-3303.

25. Englund-Ögge L, Birgisdottir BE, Sengpiel V*, et al.* (2017) Meal Frequency Patterns and Glycemic Properties of Maternal Diet in Relation to Preterm Delivery: Results from a Large Prospective Cohort Study. *PLoS One*, **12**(3), 1-19.

26. Wilson JE, Blizzard L, Gall SL*, et al.* (2019) An Eating Pattern Characterised by Skipped or Delayed Breakfast Is Associated with Mood Disorders among an Australian Adult Cohort. *Psychol Med*, **50**(16), 2711-2721.

27. Leech RM, Worsley A, Timperio A*, et al.* (2017) Temporal Eating Patterns: A Latent Class Analysis Approach. *Int J of Behav Nutr Phys Act* **14**(1), 1-9.

28. Khanna N, Eicher-Miller HA, Boushey CJ*, et al.* (2011) Temporal Dietary Patterns Using Kernel K-Means Clustering. *IEEE International Symposium on Multimedia*, **2011**, Dana Point, CA, USA, 375-380.

29. Aqeel MM, Guo J, Lin L*, et al.* (2020) Temporal Dietary Patterns Are Associated with Obesity in US Adults. *J Nutr*, **150**(12), 3259-3268.

30. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

31. Murakami K, Livingstone MBE, Sasaki S*, et al.* (2018) Applying a Meal Coding System to 16-D Weighed Dietary Record Data in the Japanese Context: Towards the Development of Simple Meal-Based Dietary Assessment Tools. *J Nutr Sci*, **7**(e29), 1-10.

32. Uzhova I, Woolhead C, Timon CM*, et al.* (2018) Generic Meal Patterns Identified by Latent Class Analysis: Insights from NANS (National Adult Nutrition Survey). *Nutrients*, **10**(3), 310.

33. Hearty ÁP and Gibney MJ (2008) Analysis of Meal Patterns with the Use of Supervised Data Mining Techniques—Artificial Neural Networks and Decision Trees. *Am J Clin Nutr*, **88**(6), 1632-1642.

34. Riou J, Lefèvre T, Parizot I*, et al.* (2015) Is There Still a French Eating Model? A Taxonomy of Eating Behaviors in Adults Living in the Paris Metropolitan Area in 2010. *PLoS One*, **10**(3), 1-18.

35. Shinozaki N, Murakami K, Asakura K*, et al.* (2021) Identification of Dish-Based Dietary Patterns for Breakfast, Lunch, and Dinner and Their Diet Quality in Japanese Adults. *Nutrients*, **13**(1), 67.

36. Shi Z, Riley M, Taylor A*, et al.* (2017) Meal-Specific Food Patterns and the Incidence of Hyperglycemia in a Chinese Adult Population. *Br J Nutr*, **118**(1), 53-59.

37. Leech RM, Timperio A, Worsley A*, et al.* (2019) Eating Patterns of Australian Adults: Associations with Blood Pressure and Hypertension Prevalence. *Eur J Nutr*, **58**(5), 1899-1909.

38. Eicher-Miller HA, Khanna N, Boushey CJ*, et al.* (2016) Temporal Dietary Patterns Derived among the Adult Participants of NHANES 1999-2004 Are Associated with Diet Quality. *J Acad Nutr Diet*, **116**(2), 283-291.

39. Newby PK, Sc D, Tucker KL*, et al.* (2004) Empirically Derived Eating Patterns Using Factor or Cluster Analysis: A Review. *Nutr Rev*, **62**(5), 177-203.

40. Hearty AP and Gibney MJ (2009) Comparison of Cluster and Principal Component Analysis Techniques to Derive Dietary Patterns in Irish Adults. *Br J Nutr*, **101**(4), 598-608.

41. Newby PK, Muller D, and Tucker KL (2004) Associations of Empirically Derived Eating Patterns with Plasma Lipid Biomarkers: A Comparison of Factor and Cluster Analysis Methods. *Am J Clin Nutr*, **80**(3), 759 - 767.

42. Bamia C, Orfanos P, Ferrari P*, et al.* (2005) Dietary Patterns among Older Europeans: The EPIC-Elderly Study. *Br J Nutr*, **94**(1), 100-113.

43. Ocke MC (2013) Evaluation of Methodologies for Assessing the Overall Diet: Dietary Quality Scores and Dietary Pattern Analysis. *Proc Nutr Soc*, **72**(2), 191-199.

44. Amoutzopoulos B, Page P, Roberts C*, et al.* (2020) Portion Size Estimation in Dietary Assessment: A Systematic Review of Existing Tools, Their Strengths and Limitations. *Nutr Rev*, **78**(11), 885-900.

45. de Vlieger NM, Weltert M, Molenaar A*, et al.* (2020) A Systematic Review of Recall Errors Associated with Portion Size Estimation Aids in Children. *Appetite*, **147**(1 April), 104522.

46. Culkin A, Duncan A, Dyson P*, et al.* (2020) Nutrition Assessment. In *Oxford Handbook of Nutrition and Dietetics*, 3rd ed., pp. 37-60 [J Webster-Gandy, A Madden and M Holdsworth, editors]. Oxford, UK: Oxford University Press.

47. Park Y, Dodd KW, Kipnis V*, et al.* (2018) Comparison of Self-Reported Dietary Intakes from the Automated Self-Administered 24-H Recall, 4-D Food Records, and Food-Frequency Questionnaires against Recovery Biomarkers. *Am J Clin Nutr*, **107**(1), 80-93.

48. Bailey RL (2021) Overview of Dietary Assessment Methods for Measuring Intakes of Foods, Beverages, and Dietary Supplements in Research Studies. *Curr Opin Biotechnol*, **70**(August), 91-96.

49. Ma Y, Olendzki BC, Pagoto SL*, et al.* (2009) Number of 24-Hour Diet Recalls Needed to Estimate Energy Intake. *Ann Epidemiol*, **19**(8), 553-559.

50. Cui Q, Xia Y, Wu Q*, et al.* (2021) Validity of the Food Frequency Questionnaire for Adults in Nutritional Epidemiological Studies: A Systematic Review and Meta-Analysis. *Crit Rev Food Sci Nutr*, **63**(12), 1670-1688.

51. Cui Q, Xia Y, Wu Q*, et al.* (2021) A Meta-Analysis of the Reproducibility of Food Frequency Questionnaires in Nutritional Epidemiological Studies. *Int J Behav Nutr Phys Act*, **18**(1), 12.

52. Tucker KL, Smith CE, Lai CQ*, et al.* (2013) Quantifying Diet for Nutrigenomic Studies. *Annu Rev Nutr*, **33**, 349-371.

53. Thompson FE and Subar AF (2017) Dietary Assessment Methodology. In *Nutrition in the Prevention and Treatment of Disease*, 4th ed., pp. 5-48 [A Coulston, C Boushey, M Ferruzzi and L Delahanty, editors]. Amsterdam, Netherlands: Academic Press.

54. Cade JE, Warthon-Medina M, Albar S*, et al.* (2017) Diet@Net: Best Practice Guidelines for Dietary Assessment in Health Research. *BMC Med*, **15**(1), 202.

55. Subar AF, Freedman LS, Tooze JA*, et al.* (2015) Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr*, **145**(12), 2639-2645.

56. Burrows TL, Ho YY, Rollo ME*, et al.* (2019) Validity of Dietary Assessment Methods When Compared to the Method of Doubly Labeled Water: A Systematic Review in Adults. *Front Endocrinol (Lausanne)*, **10**, 850.

57. Freedman LS, Commins JM, Moler JE*, et al.* (2014) Pooled Results from 5 Validation Studies of Dietary Self-Report Instruments Using Recovery Biomarkers for Energy and Protein Intake. *Am J Epidemiol*, **180**(2), 172-188.

58. Freedman LS, Commins JM, Moler JE*, et al.* (2015) Pooled Results from 5 Validation Studies of Dietary Self-Report Instruments Using Recovery Biomarkers for Potassium and Sodium Intake. *Am J Epidemiol*, **181**(7), 473-487.

59. Guasch-Ferre M, Bhupathiraju SN, and Hu FB (2018) Use of Metabolomics in Improving Assessment of Dietary Intake. *Clin Chem*, **64**(1), 82-98.

60. Liang S, Nasir RF, Bell-Anderson KS*, et al.* (2022) Biomarkers of Dietary Patterns: A Systematic Review of Randomized Controlled Trials. *Nutr Rev*, **80**(8), 1856-1895.

61. Clarke ED, Rollo ME, Pezdirc K*, et al.* (2020) Urinary Biomarkers of Dietary Intake: A Review. *Nutr Rev*, **78**(5), 364-381.

62. Scalbert A, Brennan L, Manach C*, et al.* (2014) The Food Metabolome: A Window over Dietary Exposure. *Am J Clin Nutr*, **99**(6), 1286-1308.

63. Brennan L, Hu FB, and Sun Q (2021) Metabolomics Meets Nutritional Epidemiology: Harnessing the Potential in Metabolomics Data. *Metabolites*, **11**(10), 709.

64. 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**(1), 55.

65. 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**(4), 1187-1203.

66. McClung HL, Ptomey LT, Shook RP*, et al.* (2018) Dietary Intake and Physical Activity Assessment: Current Tools, Techniques, and Technologies for Use in Adult Populations. *Am J Prev Med*, **55**(4), e93-e104.

67. Rollo ME, Williams RL, Burrows T*, et al.* (2016) What Are They Really Eating? A Review on New Approaches to Dietary Intake Assessment and Validation. *Current Nutrition Reports*, **5**(4), 307-314.

68. Six BL, Schap TE, Zhu FM*, et al.* (2010) Evidence-Based Development of a Mobile Telephone Food Record. *J Am Diet Assoc*, **110**(1), 74-79.

69. Timon CM, Blain RJ, McNulty B*, et al.* (2017) The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. *J Med Internet Res*, **19**(5), e158.

70. Boushey CJ, Spoden M, Zhu FM*, et al.* (2017) New Mobile Methods for Dietary Assessment: Review of Image-Assisted and Image-Based Dietary Assessment Methods. *Proc Nutr Soc*, **76**(3), 283-294.

71. Zhu F, Bosch M, Khanna N*, et al.* (2015) Multiple Hypotheses Image Segmentation and Classification with Application to Dietary Assessment. *IEEE J Biomed Health Inform*, **19**(1), 377-388.

72. Kong F, He H, Raynor HA*, et al.* (2015) Dietcam: Multi-View Regular Shape Food Recognition with a Camera Phone. *Pervasive and Mobile Computing*, **19**(May), 108-121.

73. Wang Y, He Y, Boushey CJ*, et al.* (2018) Context Based Image Analysis with Application in Dietary Assessment and Evaluation. *Multimed Tools Appl*, **77**(15), 19769-19794.

74. Dalakleidi KV, Papadelli M, Kapolos I*, et al.* (2022) Applying Image-Based Food-Recognition Systems on Dietary Assessment: A Systematic Review. *Adv Nutr*, **13**(6), 2590-2619.

75. Konstantakopoulos FS, Georga EI, and Fotiadis DI (2023) A Review of Image-Based Food Recognition and Volume Estimation Artificial Intelligence Systems. *IEEE Rev Biomed Eng*(Online ahead of print).

76. Lo FPW, Sun Y, Qiu J*, et al.* (2020) Image-Based Food Classification and Volume Estimation for Dietary Assessment: A Review. *IEEE J Biomed Health Inform*, **24**(7), 1926-1939.

77. Forster H, Walsh MC, O'Donovan CB*, et al.* (2016) A Dietary Feedback System for the Delivery of Consistent Personalized Dietary Advice in the Web-Based Multicenter Food4Me Study. *J Med Internet Res*, **18**(6), e150.

78. Turner-McGrievy GM, Yang C-H, Monroe C*, et al.* (2021) Is Burden Always Bad? Emerging Low-Burden Approaches to Mobile Dietary Self-Monitoring and the Role Burden Plays with Engagement. *Journal of Technology in Behavioral Science*, **6**(3), 447-455.

79. Murakami K, Shinozaki N, McCaffrey TA*, et al.* (2021) Data-Driven Development of the Meal-Based Diet History Questionnaire for Japanese Adults. *Br J Nutr*, **126**(7), 1056-1064.

80. Fallaize R, Forster H, Macready AL*, et al.* (2014) Online Dietary Intake Estimation: Reproducibility and Validity of the Food4Me Food Frequency Questionnaire against a 4-Day Weighed Food Record. *J Med Internet Res*, **16**(8), e190.

81. Kristal AR, Kolar AS, Fisher JL*, et al.* (2014) Evaluation of Web-Based, Self-Administered, Graphical Food Frequency Questionnaire. *J Acad Nutr Diet*, **114**(4), 613-621.

82. Labonte ME, Cyr A, Baril-Gravel L*, et al.* (2012) Validity and Reproducibility of a Web-Based, Self-Administered Food Frequency Questionnaire. *Eur J Clin Nutr*, **66**(2), 166-173.

83. 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**(1), 233-240.

84. Simpson E, Bradley J, Poliakov I*, et al.* (2017) Iterative Development of an Online Dietary Recall Tool: Intake24. *Nutrients*, **9**(2), 118.

85. Carter MC, Burley VJ, Nykjaer C*, et al.* (2013) 'My Meal Mate' (MMM): Validation of the Diet Measures Captured on a Smartphone Application to Facilitate Weight Loss. *Br J Nutr*, **109**(3), 539-546.

86. Ocké M, Dinnissen C, Stafleu A*, et al.* (2021) Relative Validity of Mijneetmeter: A Food Diary App for Self-Monitoring of Dietary Intake. *Nutrients*, **13**(4), 1135.

CHAPTER 2

Methods

2.1 Chapter Overview

This chapter provides an outline of the methodology for this thesis as a whole. The next section will cover the overall structure of the thesis, how it is laid out, how the various chapters are linked, and how they meet the objectives of the thesis. Subsequent sections will then provide more detailed methodological information on the data used in this thesis and on the individual chapters that follow, in addition to the methodological information provided in the subsequent chapters themselves.

2.2 Structure of the Thesis

The structure of the thesis is based on the overall research aim and objectives set out in the introduction (Chapter 1). Each of chapters 3 through to 7 contribute to achieving these research objectives as illustrated in **Figure 2.1**.

Chapter 3 is a literature review of studies that have used advanced statistical techniques to identify meal patterns. The challenges and opportunities associated with the various techniques that were identified during this review were used to inform the subsequent work in this thesis in developing a meal pattern analysis technique that could be used for dietary intake assessment. It contributes to achieving objective 1 of the thesis: identify existing meal pattern research, research gaps, challenges, and opportunities.

The findings from the review of statistical approaches in Chapter 3 were used to inform Chapter 4 which involved applying a new approach for the identification of generic, or characteristic, meals to the data from the National Adult Nutrition Survey (NANS) in Ireland. The generic meals that were identified as part of that work were then used to inform subsequent work (Chapter 7) on the development of a meal-based method of dietary intake assessment.

Chapter 5 involves identifying content meal patterns (common combinations of generic meals consumed over a day) in a single dataset (the National Health and Nutrition Examination Survey (NHANES)) using the multiple statistical techniques identified in Chapter 3, in order to determine the relative comparability of those techniques when used on the same data, and to identify their relative strengths and weaknesses. Both Chapters 4 and 5 achieved objective 2 of the thesis: refine the framework for identifying generic meals and meal patterns for use in dietary intake assessment.

A single approach from those used in Chapter 5 was applied to the NHANES data in Chapter 6 to identify and describe the meal patterns that exist in the USA and to further

explore any relationships that exist between meal patterns and health. This chapter achieved objective 3 of the thesis: identify relationships between adherence to identified meal patterns and health parameters.

While the work in the previous chapters involved using data that had previously been collected at the food level and analysing it in a meal-based way, Chapter 7 describes the development of a new method of dietary intake assessment that allows these data to be collected at the meal level in the first instance. The meals used in this chapter were those identified in Chapter 4. Chapter 7 achieved objective 4: develop new meal-based method of dietary intake assessment and compare with established method.

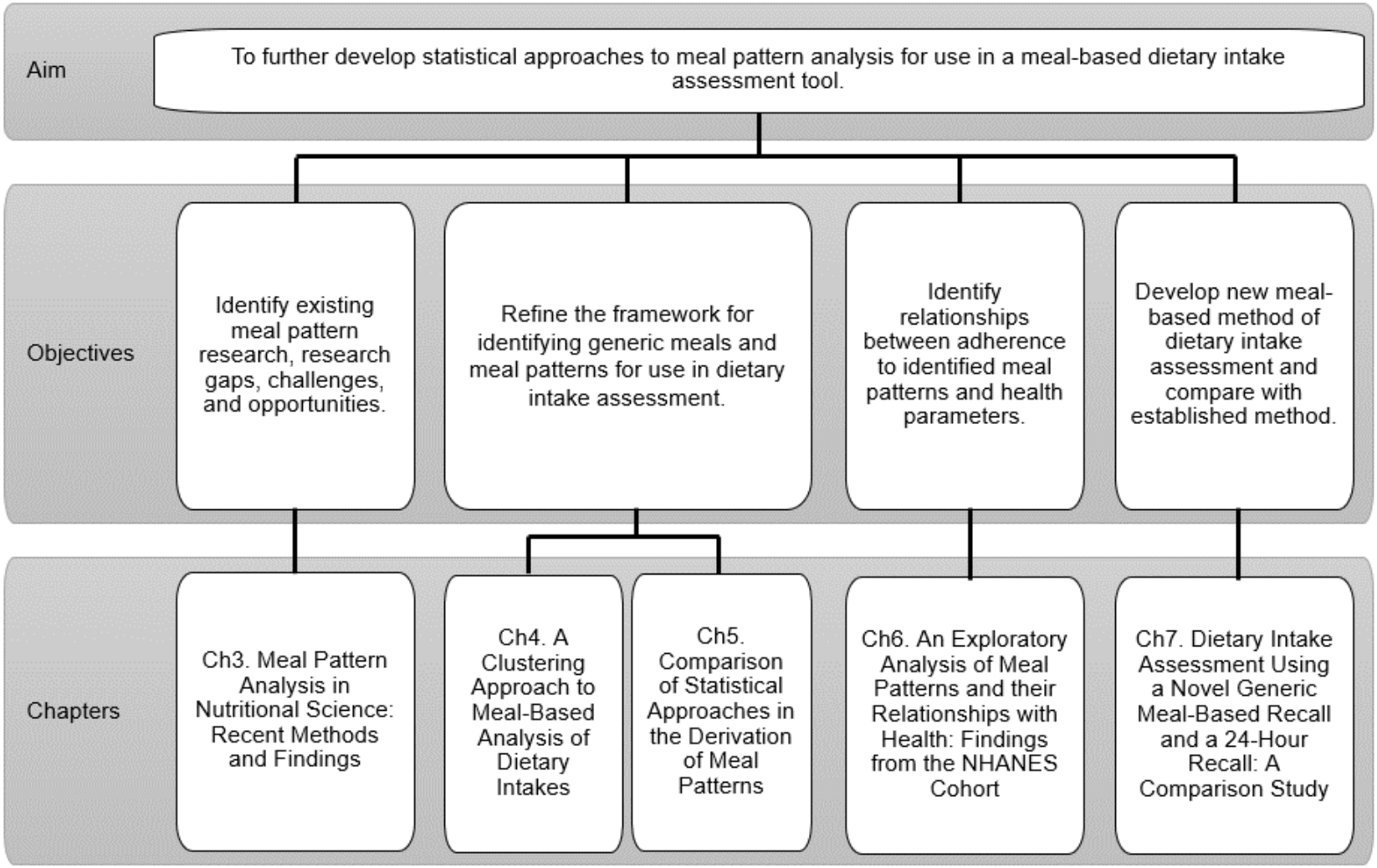


Figure 2.1: Overview of relationships between thesis aim, objectives, and chapters.

2.3 Datasets Used within the Thesis

Three chapters in this thesis describe the secondary analysis of datasets that had previously been created by other research groups. Data from the National Adult Nutrition Survey (NANS) [1] were used in Chapter 4 and data from the National Health and Nutrition Examination Survey (NHANES) were used in Chapters 5 and 6 [2]. Each of these are described below.

2.3.1 National Adult Nutrition Survey (NANS)

Data for NANS were collected between October 2008 and April 2010. Participants in NANS were aged between 18 and 90 years, were free-living, and not pregnant or breast feeding. The sample was representative of Irish adults with respect to age, gender, social class, and geographical location. Ethical approval for NANS was originally provided by University College Cork Clinical Research Ethics Committee of the Cork Teaching Hospitals [1]. The Irish Universities Nutrition Alliance have previously published detailed procedures for all methods used in NANS [1]. Food and nutrient intake data along with anthropometry and demographic data were used in this thesis and are described in brief in the sections that follow.

2.3.1.1 Dietary Data

Each participant completed a four-day weighed food diary. Prior to starting to record food intakes, a researcher visited each participant to provide training on how to complete the food diary and use the food weighing scales provided. Participants were also visited 24–36 hours into the recording period to review the diary and clarify any queries they might have. One or two days after the recording period the participant was visited for the collection of the diary and a final review of clarity and completeness [1]. Reported food intake data were converted to nutrient data using WISP© (Tinuviel Software, Anglesey, UK). The dataset used in this thesis consisted of a row for each food consumed by each participant over the course of the recording period.

2.3.1.2 Anthropometry

Height was measured using a stadiometer with a fixed vertical backboard and an adjustable head piece. Measurements were taken to the nearest 0.1cm with the participant's head positioned in the Frankfurt Plane. Participants' weight was measured in kilograms using a digital scale. Measurements were taken in duplicate to the nearest 0.1kg after participants had voided and were wearing light clothing without shoes.

2.3.1.3 Demographic Data

Demographic data were collected using a questionnaire. Data on age and sex were used in this thesis.

2.3.2 National Health and Nutrition Examination Survey (NHANES)

Data for NHANES was collected during 2017 and 2018. Non-institutionalised civilian residents of the USA of all ages were included in the dataset, but only those aged greater than 18 years were included in the current thesis. NHANES was approved by the National Center for Health Statistics Ethics Review Board [(2)]. Detailed procedures for the methods used to collect data and take measurements in NHANES are given by the Centers for Disease Control and Prevention (CDC) [(2)]. Food and nutrient intake data, along with anthropometry and body composition data, demographic and health data were used in this thesis and are described in brief in the sections below.

2.3.2.1 Dietary Data

Each participant completed two 24-hour recalls (24HR); these were conducted by trained interviewers in person for the first recalls and over telephone for the second recalls using a standardised process. The dataset used in this thesis consisted of a row for each food consumed by each participant over the course of the recording period. The nutrient composition of the foods reported in the 24HRs are from the food and nutrient database for dietary studies (FNDDS) 2017-2018 [(3)].

2.3.2.2 Anthropometry and Body Composition

Height was measured using a stadiometer with a fixed vertical backboard and an adjustable head piece. The participant was instructed to stand straight against the backboard with the head, shoulder blades, buttocks, and heals making contact with the backboard. Measurements were taken to the nearest 0.1cm. Participants' weight was measured in kilograms using a calibrated digital scale. Participants wore a standard examination gown consisting of a disposable shirt, trousers, and slippers. Participants were instructed to stand on the centre of the scale, with their hands at their sides, and looking straight ahead. Measurements were taken to the nearest 0.1kg. Waist (abdominal) circumference measurements were taken on participants' bare abdomen with their hands placed on opposite shoulders. The measurements were taken at the point of the uppermost lateral border of the right ilium of the pelvis, and the measuring tape extended horizontally around the waist from that point. Hip (buttocks) circumference measurements were taken around the maximum protuberance of the buttocks when viewed from participants' right side. Both

measurements were taken to the nearest 0.1cm at the end of a normal exhalation from the participant. Percentage body fat was estimated using dual energy x-ray absorptiometry (DXA) using the Hologic Discovery A fan beam x-ray bone densitometer. Participants were positioned lying flat on the table in alignment with the centre lines markings, with a straight chin, legs positioned together, feet relaxed, and toes pointed upwards; arms were positioned straight by participants' sides with palms facing down and separated from the torso [(2)].

2.3.2.3 Clinical Measurements and Biochemistry

Prior to blood pressure measurement, participants rested quietly in a seated position for five minutes. Three blood pressure measurements were then taken with one minute gap between each one. The mean blood pressure value of the three measurements was used in the current thesis. Measurements were taken on the bare upper right arm unless there were any specific conditions that did not allow for this, in which case the left arm was used. The Omron IntelliSense Blood Pressure Monitor (Model: HEM-907XL) was used to take the blood pressure measurements.

For biochemical measurements, a fasting questionnaire was administered prior to venipuncture to confirm fasting status. For fasting samples, participants are asked to fast for nine hours. Venipuncture was performed on the median cubital, cephalic, or basilic veins of the left arm unless unsuitable, in which case the right arm was used. If that was unsuitable, then veins in the left forearm or dorsal side of the left hand were considered, and if that was unsuitable, the right forearm/hand was considered. The vacutainer blood collection device and blood tubes at room temperature were used to draw the blood. The samples used for plasma glucose were collected in sodium fluoride tubes and measured with a UV in vitro test using the Roche/Hitachi Cobas C System (c311). The samples used for serum total cholesterol and HDL were collected in EDTA or heparin tubes and measured photometrically using the Roche/Hitachi Cobas 6000 Chemistry Analyser. LDL Cholesterol was not measured directly but calculated using the Friedewald equation [(2)].

2.3.2.4 Demographics

Demographic data were collected using a questionnaire. Data on age and sex were used in this thesis.

2.4 Chapter 3. Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings

This chapter encompasses a review which examines the advanced statistical methods that have been used to date in meal pattern research. Given the limited research in this area, the chapter takes the form of a narrative review, allowing a broader focus and discussion rather than defining a narrow research question in this area of limited research. It also allowed consideration to be given to both qualitative and quantitative aspects of meal pattern research regarding the types of statistical methods, meals, the meal patterns themselves and the relationships between meal patterns and diet quality or health outcomes.

While the overall approach taken was that of a narrative review, a systematic approach was applied to searching the literature for meal pattern research. Pubmed, Web of Science, and Google Scholar databases were searched using the following terms: meal pattern, meal assessment, meal intake, meal coding, and food combination. Studies were limited to those in the English language and conducted in humans. Pubmed and Web of Science searches were conducted on 28th of February 2020 and the Google Scholar search was conducted on 2nd March 2020. All searches were limited to publications published in the preceding ten years. A total of 1731 publications were identified (340 from Pubmed, 791 from Web of Science, and 600 from Google Scholar), reducing to 1328 when duplicates were removed. Only six publications remained after titles, abstracts, and full texts were reviewed in a stepwise process to limit the results to publications in which meal patterns were identified and defined, in which data-driven advanced statistical techniques were used, and in which participants were aged 18 years or older. Four further publications were identified from the references of the initial six, and a further three publications were included that were follow-on studies of some of the initial 10 publications, e.g., publications investigating relationships between the meal patterns identified in the initial study and diet quality or health.

Within the chapter, the statistical methods used for meal pattern research were described and compared, the meal patterns and any associations with diet quality or health were reported, and the overall strengths, limitations, and gaps in the literature were also discussed. Subsequent chapters aimed to address some of the gaps in the literature and research questions identified and raised within this review.

2.5 Chapter 4. A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data

A new approach to meal-based analysis was developed for the purpose of this thesis. This method was based on the research gaps identified in the literature review in Chapter 3, and previously applied methods by Woolhead *et al.* [4].

In general, content meal patterns are identified in three broad steps. Firstly, foods are grouped into food groups; secondly, foods within meals are replaced with the food groups to which they belong, and these meals are grouped into generic meals (i.e., typical or characteristic meals); thirdly, participants are grouped based on their meal patterns (the combinations of generic meals that they consume) (**Figure 2.2**).

This chapter describes details of the first two steps in that process, i.e., determination of food groups and generic meals. This work was carried out on the dietary intake data from NANS, as outlined in section 2.3.1.1 above.

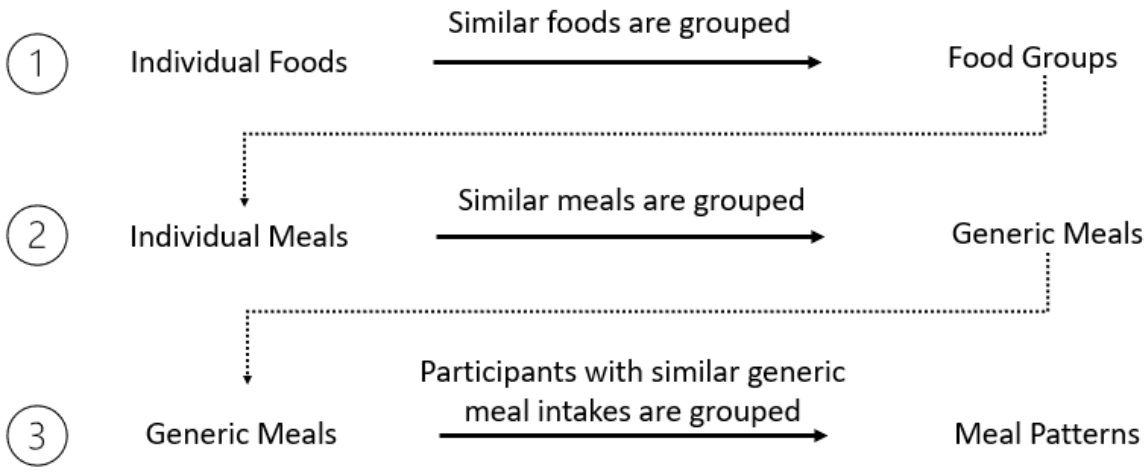


Figure 2.2: Conceptual overview of the three-step process for identifying meal patterns.

2.5.1 Step 1. Food Groups

In this step, foods reported in the dietary survey were grouped into one of 6 groups within the Irish food pyramid (namely vegetables, salad, and fruit; cereal, bread, pasta, potato, and rice; milk, yogurt, and cheese; meat, poultry, fish, eggs, beans, and nuts; fats, spreads, and oils; and foods high in fat, sugar and/or salt) [5]. Individual foods within these broad food groups were then further grouped using *k*-means clustering, grouping foods based on how

similar they were regarding their nutrient density for 12 different nutrients [6], as described below Section 2.5.1.1.

To describe the approach, it is important to consider two aspects of the development of the clustering approach in this and subsequent chapters. Firstly, the input variables and secondly, the number of clusters to select in the analysis.

2.5.1.1 Input Variables

Clustering is a data analysis technique that groups observations (in this case food items) into groups, or clusters, in such a way that the features (or variables; in this case the nutrients of interest within the food items) of the foods within a given cluster are more similar to each other than to the food items in the other clusters [7]. Within the analysis for this chapter, the features (or input variables) used were 12 nutrients in the previously validated nutrition quality index, the Nutrient Rich Foods (NRF) Index [8]. Thus, the variables used for clustering were protein, fibre, vitamin A, vitamin C, vitamin E, calcium, iron, magnesium, potassium, saturated fat, added sugar, and sodium (**Table 2.1**).

Different variations of input variables were tested (**Table 2.2**) in the current thesis before concluding to use the variables listed in **Table 2.1**. The different variations were based on the different combinations of nutrients tested for use in the NRF index by Fulgoni *et al.* [8]. This was based on clustering foods in the "vegetables, salad, and fruit" section of the food pyramid four separate times, once for each of the different input variables based on either the NRF6.3, NRF9.3, NRF11.3, or NRF15.3. An analysis of variation (ANOVA) was carried out on each clustering solution comparing the mean nutrient content of the different clusters. One of the nine input nutrients differed among all clusters when nutrients from the NRF6.3 were used, seven of the 12 input nutrients differed among all clusters when nutrients from the NRF9.3 were used, eight of the 14 input nutrients differed among all clusters when nutrients from the NRF11.3 were used, and three of the 18 input nutrients differed among all clusters when nutrients from the NRF15.3 were used. The use of the nutrients from NRF9.3 gave rise to the greatest variation among clusters (food groups) relative to the number of input variables used, and therefore it was those 12 nutrients that were chosen for the purpose of grouping foods. The original NRF index study also considered expressing nutrients in units of either per 100g or per 100kcal [8]. Both options were also tested using the foods from the "vegetables, salad, and fruit" section of the food pyramid in the current thesis. The use of units per 100g gave rise to outlier foods whereby some food groups only contained a single food or two foods. Therefore, units per 100kcal, was chosen to achieve more balanced food groups, and variables were z-standardised prior to clustering [6].

Table 2.1: An overview of the statistical methods and input variables used at each step of the process for identifying meal patterns.

	Description and Statistical Method	Input Variables
Step 1 Food Groups	Similar foods are grouped together into food groups using *k*-means clustering.	Each food had a value for the following variables used in the clustering. All variables were z-standardised. Their original units are given below. • Protein (g/100kcal) • Fibre (g/100kcal) • Vitamin A (µg/100kcal) • Vitamin C (mg/100kcal) • Vitamin E (mg/100kcal) • Calcium (mg/100kcal) • Iron (mg/100kcal) • Magnesium (mg/100kcal) • Potassium (mg/100kcal) • Saturated fat (g/100kcal) • Added sugar (NANS) or total sugar (NHANES) (g/100kcal) • Sodium (mg/100kcal)

Step 2
Generic Meals
	Similar individual meals characterised using food groups (Step 1) are grouped together into generic meals using partitioning around the medoids clustering.	Each meal had a value for the following variables used in the clustering.
		• NRF9.3 score
		• Binary variables:
		  ○ NANS: 15 binary variables indicating for each group whether it was present/absent in the meal.
		  ○ NHANES: 23 binary variables indicating for each food group whether it was present/absent in the meal.

Step 3 Meal Patterns	The following three statistical approaches were used and their outputs compared. • PAM clustering • Principal component analysis • Latent class analysis	Clustering and PCA: Each person had a numerical variable for each generic meal indicating %TEI and a binary variable for each meal type indicating consumption/non-consumption. LCA: Each person had a categorical variable for each meal type indicating which generic meal was consumed and binary for consumption/non-consumption of each meal type.

LCA, latent class analysis; NANS, National Adult Nutrition Survey; NHANES, National Health and Nutrition Examination Survey; PAM, partitioning around the medoids; PCA, principal components analysis.

Table 2.2: The various combinations of nutrients trialled for inclusion as input variables to the clustering when creating food groups.

NRF Model	Macronutrients	Vitamins	Minerals	Nutrients to Limit
NRF6.3	Protein, Fibre	A, C	Ca, Fe	Saturated Fat, added sugar or total sugar, Na. (Same for all models)
NRF9.3	Protein, Fibre	A, C, E	Ca, Fe, Mg, K	
NRF11.3	Protein, Fibre	A, C, E, B12	Ca, Fe, Mg, Zn, K	
NRF15.3	Protein, Fibre, monounsaturated fat	A, C, D, E, B1, B2, B12, folate	Ca, Fe, Zn, K	

Adapted from Fulgoni *et al.* [8]. NRF, nutrient rich foods.

2.5.1.2 Number of Clusters

When carrying out *k*-means clustering it is necessary to determine the number of clusters in advance of the analysis. Within this thesis the number of clusters was chosen by applying a statistically driven approach whereby 24 different validity indices (methods of choosing clusters) were used to determine the number of clusters to be applied to the data. Using this approach, the number of clusters that was most frequently proposed among the indices was chosen as the number of clusters for the analysis [9]. For two of the food pyramid groups (cereal, bread, potato, pasta and rice and milk, yogurt, and cheese), the clustering solution arising from the most frequently proposed number of clusters led to some clusters containing only two or three foods. In these instances, the next most frequently chosen value for cluster number was used. The range of possible values assessed for cluster number was from 2 to 10 inclusive. A list of the indices used are given in Supplementary Table 2.1.

Foods that did not have a specific group in the food pyramid were not clustered but considered as individual food groups. These were alcoholic beverages, non-alcoholic beverages, and miscellaneous foods such as sauces, dips, and dressings. The fats, spreads, and oils group was not further split using clustering given the small number of foods in this group (n = 36). In total there were 12 food groups identified through clustering and the 3 remaining groups above that do not appear in the food pyramid. These 15 food groups were then used to derive generic meals as described below.

In summary, within this step of the process, foods were grouped firstly according to their grouping in the Irish food pyramid [5]. Within these groupings, clusters were created using *k*-means clustering to group foods that were similar with regard to their nutrient content. It

was these clusters that defined the food groups that were used in the next step to identify generic meals.

2.5.2 Step 2. Generic Meals

Following the *k*-means clustering to determine the 15 food groups, individual foods listed within each meal, reported by participants, were replaced with the food group to which they belonged. For example, cornflakes reported in a breakfast meal would be replaced with "cereals", and low fat milk replaced with "milk and yogurt (and non-dairy alternatives)". Individual meals were then defined by the food groups that now made up each meal, meaning that for each food reported in a given meal, the food group to which it was assigned was used to represent that food, instead of the individual food item.

Following this substitution of food groups in place of foods in the meal, an NRF9.3 score [8] was calculated for each individual meal as an indicator of nutritional quality. The score was calculated based on 9 nutrients to encourage (protein, dietary fibre, calcium, iron, potassium, magnesium, and vitamins A, C, and E) and 3 nutrients to limit (saturated fat, added sugar, and sodium), as described previously. For each of these nutrients the quantity of the nutrient per 100kcal present in each meal as a percentage of reference intakes (RIs) was calculated. RIs were derived from Regulation (EU) No 1169/2011 [10] where available. Those that were not available in that regulation (dietary fibre and added sugar) were instead derived from EFSA (European Food Safety Authority) [11] (**Table 2.3**). The percentage of RIs were limited to 100% to avoid over valuing meals, as has been applied previously [8, 12-14]. The percentage RI scores for the nutrients to encourage were summed, as were the scores for the nutrients to limit. The nutrients to limit score was subtracted from the nutrients to encourage score to give an overall NRF9.3 score for each meal. The added sugar contents of foods were estimated using a 10-step process developed by Louie *et al.* [15]. Thus, at the end of this part, the meals were represented by combinations of food groups and a meal NRF9.3 score.

Table 2.3: The nutrient values used as reference intakes when calculating the nutrient rich foods index (NRF9.3) for generic meals in the NANS and NHANES datasets.

Nutrient	Reference Intakes		Comments
	NANS	NHANES	
Protein	50g	50g	
Fibre	25g	28g	
Vitamin A	800µg	900µg	
Vitamin C	80mg	90mg	
Vitamin E	12mg	15mg	
Calcium	800mg	1300mg	
Iron	14mg	18mg	
Magnesium	375mg	420mg	
Potassium	2000mg	4700mg	
Saturated fat	20g	20g	
Added sugar	50g	N/A	Based on 10% of energy from a 2000kcal diet with an energy to mass conversion factor of 4.
Total sugar	N/A	90g	
Sodium	2400mg	2300mg	NANS values based on value for salt of 6g converted to sodium using a factor of 2.5.

Values used in NANS were from Regulation (EU) No 1169/2011 [10] with the exception of added sugar which was from EFSA (European Food Safety Authority) [11]. Values used in NHANES were from United States Food and Drug Administration [16] with the exception of total sugar which was from Regulation (EU) No 1169/2011 [10]. NANS, National Adult Nutrition Survey; NHANES, National Health and Nutrition Examination Survey.

Partitioning around the medoids (PAM) clustering was then used to group meals based on how similar they were with respect to the food groups they contained and their nutrient profiling score (NRF9.3) [8]. Within each meal type (breakfast, lunch, dinner, snacks, and beverages), and separately for meals consumed on weekdays and weekends, the individual meals were grouped/clustered using PAM clustering which allows for clustering based on both numerical and categorical variables [17, 18]. In the NANS dataset the meal types were condensed to five meal types for this analysis: breakfast (comprising of breakfast), light meals ("light meal (lunch)" and "light meal (evening meal)"), main meals ("main meal (lunch)" and "main meal (evening meal)"), snacks ("morning snack", "afternoon snack", "evening snack", and "night snack"), and beverages ("alcoholic beverage" and "non-alcoholic beverage"). The input variables used to cluster were the NRF9.3 score for each meal and 15 binary variables indicating for each of the 15 food groups whether it was present or absent in the meal (**Table 2.1**). As with k-means clustering, PAM clustering also requires that the number of clusters is specified before clustering. The same approach described in Section 2.5.1.2 was taken. The range of values assessed for the number of clusters in the data was between 4 and 15 inclusive.

After PAM clustering, groups of similar meals were grouped together and each group of meals (cluster) was given a descriptive food-based name and a nutrient composition as follows: 1) the name was based on the foods and food groups that commonly occurred in the grouped meals (e.g., potatoes, beef/ham, and vegetables) and 2) the nutrient composition was the mean composition in units per 100g of the grouped meals. This meal is referred to as a generic meal. A list of the generic meals with their food-based descriptions and examples of portion sizes and nutrient content are given in Supplementary Table 2.2.

Before calculating mean nutrient content and portion size of these generic meals (described below), meals contained within these clusters but considered to be outliers (based on energy and/or micronutrient content) were removed in a two-step process as described previously by Woolhead *et al.* [4]. Firstly, meals that contained greater than 1.5 times the interquartile range for total energy content were removed. For example, one participant consumed a meal containing roast potatoes, mixed vegetables, breaded fish, cheese, four glasses of wine, and two pints of stout which provided 2188kcal. This was greater than the upper value for inclusion of 1988kcal, therefore this meal was not included when calculating the mean nutrient content or portion sizes for that generic meal. Secondly, meals that contained greater than 10 times the mean for any micronutrient were removed. For example, one participant consumed a meal containing 568g of milk, 104g of reduced fat cheddar cheese, and 208g of white bread which provided 2273mg of calcium. This was greater than the upper value for inclusion of 1966mg of calcium, therefore this meal was not included when calculating the mean nutrient content or portion sizes for that generic meal. Once outliers were removed, the weight in grams for the individual meals in each generic meal, were used to determine generic portion sizes. Each generic meal was assigned 7 generic portion sizes. These portion sizes were determined for a given generic meal by ordering each of the individual meals by weight and dividing the meals into sevenths. The median weights of each seventh were assigned as the generic portion sizes for that meal [6]. Examples of portion sizes for different generic meals are given in Supplementary Table 2.2.

2.5.3 Comparison of Generic Meal Intakes with Food-Based Intakes

To determine if the process of identifying and using generic meals to estimate reported intake is accurate, estimates of mean daily nutrient intakes were calculated using both the original food-based data and generic meal-based data and were compared with each other.

All analysis was carried out using R version 4.0.3 [19] in the RStudio integrated development environment (version 1.3.1093) [20]. Mean nutrient intakes arising from the food-based and generic data were compared using a paired *t*-test. To account for the multiple testing of energy and 30 nutrients, a Bonferroni-adjusted α of 0.05/31 was considered for statistical significance, i.e., *P* values of < 0.002 were considered statistically significant. The nutrients compared between the two datasets included fat, saturated fat, monounsaturated fat, polyunsaturated fat, protein, carbohydrate, total sugars, added sugars (all of which were compared using both gram amounts and in terms of % total energy intake (TEI)), dietary fibre (g), calcium (mg), iron (mg), potassium (mg), phosphorous (mg), sodium (mg), total vitamin A (µg), retinol (µg), carotene (µg), vitamins C (mg), D (µg), E (mg), B12 (µg), and folate (µg).

The correlation of nutrient intakes between the two datasets was examined using Spearman correlation coefficients. Nutrient intakes in both datasets were divided into quartiles. This allowed for the calculation of the proportion of individuals who remained in the same quartile in both datasets (exact agreement), the proportion who were classified in the same or adjacent quartiles (exact agreement + adjacent), the proportion who were classified two quartiles apart (disagreement), and the proportion who were classified three quartiles apart (extreme disagreement). Bland-Altman analysis was carried out whereby the mean difference between the two datasets and the limits of agreement (mean difference ± 1.96SD) for each nutrient were calculated. It is expected that ≥ 95% of observations will fall within the limits of agreement for comparable methods [21].

Finally, participants were classified, separately for both datasets, according to nutrient-based dietary guidelines [22-24]. For example, they were classified whether their nutrient intakes were low, adequate, or high according to those guidelines. The nutrients assessed included protein, carbohydrate, fat, monounsaturated fat, polyunsaturated fat, saturated fat, salt, dietary fibre, calcium, iron, folate, thiamine, riboflavin, vitamins A, B12, and C. The proportion of individuals who were classified into the same category in both datasets was calculated for each nutrient.

2.6 Chapter 5. A Comparison of Statistical Approaches in the Derivation of Meal Patterns

Once the process to identify food groups and generic meals in the NANS dataset was finalised in Chapter 4, the next step, as described in Chapter 5, was to identify meal patterns, i.e., the common combinations of generic meals consumed over a day. Given the data-driven nature of the approach taken, a different dataset was used to determine the

applicability of the approach to a dataset other than the one in which it had been developed. The dataset used was the dietary intake dataset from NHANES.

As a new dataset was being used, the two steps (food groups and generic meals) described in Chapter 4 were repeated on NHANES data to identify generic meals in that cohort. Some small adjustments to the process were required to accommodate differences between the NANS and NHANES datasets, as described below.

After having identified generic meals in NHANES, the third step was to identify meal patterns, i.e., combinations of generic meals being consumed. Three different statistical methods were used to identify meal patterns and their outputs compared: clustering, principal component analysis (PCA), and latent class analysis (LCA), as described in Section 2.6.3 below.

2.6.1 Step 1. Food Groups

Foods listed within the reported dietary recalls were grouped based on the five groups used in the USA dietary guidelines (fruit, vegetables, grains, protein foods, and dairy) [25]. Foods that did not belong in any of those groups were assigned to one of the following five groups: foods high in fat, sugar, or salt; fats and oils; non-alcoholic beverages; alcoholic beverages; and other. This gave rise to 10 food groups in total before clustering was applied. *K*-means clustering was then applied to each of the 10 food groups separately to create subgroups containing foods that were most similar to each other with regard to nutrients listed within the Nutrient Rich Foods (NRF9.3) (**Table 2.1**).

As with the analysis of the NANS dataset, when carrying out *k*-means clustering it was necessary to determine the number of clusters in advance. In NANS the number of clusters was chosen by applying 24 different indices used to determine the number of clusters to the data and choosing the number of clusters that was most frequently proposed among the indices [9]. The range of possible values assessed for cluster number was from 2 to 10 inclusive in NANS; however, given that the highest number of clusters identified in NANS food groups was 3, the range of possible values assessed for cluster number in the NHANES dataset was from 2 to 8 inclusive. The reduced the computing power and length of time required to run the analysis on the larger NHANES dataset. The other aspects of the clustering were the same as those carried out in the NANS dataset.

The clustering of the non-alcoholic beverages group resulted in unbalanced clusters with very small numbers of foods in some clusters (i.e., 151 foods in one cluster and 10 in the other), so this group was split into energy containing and energy free non-alcoholic beverages instead of clustering. This was also the case for the “other” food group in which

clustering also resulted in unbalanced clusters (i.e., 300 in one cluster and 10 in the other). This group was, therefore, not split further. Given that the alcoholic beverages group only had 76 beverages and that alcohol was not one of the substances used as an input variable for the clustering, it was decided that this group would not be further split.

Fats and oils were not clustered, instead their original three subgroups from the FNDDS were kept: animal fats; margarines; and dressings and vegetable oils [3]. This was due to the low number of foods in the group and the unbalanced clusters that arose (i.e., 42 in one cluster, 13 in another cluster, 2 in the other cluster). It was deemed appropriate to keep the manual groupings in this case to allow for some distinction in the polyunsaturated fat content of meals that contained these groups. Poor agreement was found for polyunsaturated fat when comparing nutrient intakes from generic meal intakes with those from the food-based data in previous research [4] and in the previous chapter using NANS data. In the NHANES dataset there was a final number of 23 food groups after the clustering was complete (**Table 2.1**).

2.6.2 Step 2. Generic Meals

Like the analysis in the NANS dataset, in the NHANES dataset, the meals were next defined based on the food groups they contained, i.e., the individual foods reported by the participants were replaced by the food group to which that food belonged. Unlike the NANS dataset, intakes were not split based on whether they occurred on a weekday or a weekend. This was because in the NHANES dataset, each participant completed two 24HRs as opposed to the four-day diet records completed in NANS. As a result, some participants had dietary records for two weekdays, some had two weekend days, and others had a mix. For that reason, it was decided not to split meal intakes based on whether they had been consumed on a weekday or weekend day. Instead, meals were split based on whether male or female participants had consumed them. This would also create the potential to examine differences in meal patterns based on sex as differences in dietary intake based on sex have been identified in relation to nutrient intakes [26] and dietary patterns [27].

As previously, meals were also split according to meal type. In the case of the NHANES dataset these were condensed to five meal types for the purpose of this analysis: breakfast (comprising breakfast, desayuno, and amuerzo), lunch (lunch, comida, and brunch), dinner (dinner, supper, and cena), snack (snack, merienda, entre comida, botana, bocadillo, tentempie, and extended consumption), and beverages (drink and bebida). These categorisations were based on previous methods from the literature [28-31].

Partitioning around the medoids (PAM) clustering was applied separately to each meal type, allowing similar meals in each meal type to be grouped. Similarity was based on 24 variables: 23 binary variables, each one representing one of the food groups and indicating whether that food group was present or not in the meal, and a single numeric value, namely the NRF9.3 index score [8] for each meal, which was used as an indication of the nutritional quality of the meal. The NRF9.3 was calculated in the same way as it was for the NANS dataset described previously in Section 2.5.2; however different reference intakes were used for the NHANES calculations. These were from the United States Food and Drug Administration [16] with the exception of total sugar which was from Regulation (EU) No 1169/2011 [10] because a value for total sugar was not provided by the United States Food and Drug Administration [16] (**Table 2.3**). In the NHANES dataset values for total sugar were used instead of added sugar because time constraints meant it was not possible to carry out the systematic process used in NANS to estimate the added sugar content of foods. The studies describing original development of the NRF9.3 score indicated that either added sugar or total sugar could be used [8, 32].

As with k-means clustering, PAM clustering also requires that the number of clusters is specified before clustering. The same approach described in the Section 2.6.1 was taken. The range of values assessed for the number of clusters in the data was between 4 and 15 inclusive. Descriptions, nutrient content, and portion sizes of generic meals were defined in the same way as in the NANS dataset as described above (Section 2.5.2).

2.6.3 Step 3. Meal Patterns

Meal patterns refer to the combination of generic meals consumed over the course of the day. This section describes how the generic meals identified using the methods described in the previous section (2.6.2) were analysed to determine the meal patterns that were present in the NHANES dataset. The purpose of Chapter 5 was to use three different methods of identifying meal patterns and comparing the patterns that arise from the different methods. The three methods used were PAM clustering, PCA, and LCA. Therefore, it was important to ensure that the structure and content of the input datasets were as similar as possible among the methods. However, given the different ways in which the methods work, it was not possible to have an identical structure for the input datasets for all methods, and some of the variables were adapted to suit the statistical method being used. The methods and variables used are described in the sections that follow and in **Table 2.1**.

2.6.3.1 Step 3a. Clustering

Five different variations of clustering were considered and compared, prior to selecting the approach to be used in this chapter. The approaches considered differed from each other in one or more of three ways: the type of data used as input variables to the clustering, the clustering method applied, and the distance measure used in the clustering (which determines how different individual observations are from each other and how different individual clusters are from other clusters). A summary of the different variations of clustering trialled are given in **Table 2.4**. The chosen method corresponds to method 3 in **Table 2.4** and was chosen as the approach for use in Chapter 5. This PAM clustering method was used to identify groups of participants with similar intakes of generic meals, i.e., meal patterns. The input variables were the mean percentage of total energy intake (%TEI) that each participant consumed from each generic meal over the two days' recall data, and five binary variables indicating whether or not the participants consumed each of the five meal types (breakfast, lunch, dinner, snacks, beverages) (**Table 2.1**). This method was chosen because it showed greatest variation among the clusters (meal patterns) and therefore best serves the purpose of identifying meal patterns that are distinct from one another. This variation was based on this method having the largest range of median energy intakes across the clusters and the largest number of non-negligible effect sizes for differences among clusters across 40 nutrients.

Table 2.4: The different variations of clustering trialled to identify meal patterns.

	Input Data Values	Clustering Method	Clustering Distance Metric
1	Each participant had a %TEI value from each of the generic meals.	*K*-means	Euclidean
2	Each participant had a %TEI value from each of the generic meals.	PAM	Spearman correlation
3	Each participant had a series of numerical values and a series of binary values. • The numerical values were %TEI from each of the generic meals. • The binary values indicated whether or not each of the 5 meal types were consumed.	PAM	Gower
4	Binary indicating the consumption, or not, of each of the generic meals.	PAM	Gower
5	Categories based on participants' %TEI values from each of the generic meals: • Category 1: 0% • Category 2: >0% and <15% • Category 3: ≥15% and < 30% • Category 4: ≥30%	PAM	Gower

PAM, partitioning around the medoids; %TEI, % total energy intake.

2.6.3.2 Step 3b. PCA

While the PAM clustering algorithm described in the previous section (2.6.3.1) is suitable for analysing datasets containing both numeric and binary variables, PCA is only suitable for numeric data. To allow for the inclusion of binary data in the PCA dataset, the use of 0 to define non-consumption of a meal type and 1 to define consumption of a meal type was considered. In this scenario, however, the %TEI values (scaled to between 0 and 1) from the generic meals had no influence on the resulting meal pattern solution which was instead dominated by the binary variables. The reason for this was that while the proportion of total energy intake arising from each of the generic meals could in theory could range from 0 to 1, in reality, they ranged from 0 to 0.89. Even when only assessing meals that had a value greater than 0, their median was 0.13. Therefore, the use of 1 in the binary variable to indicate consumption of a given meal type far outweighed any of values representing the proportion of TEI arising from any individual generic meal. To avoid the situation where only the binary variables were contributing to the meal pattern solution, 0 was used to indicate non-consumption and the median value of 0.13 was used to indicate consumption which ensured that both the numeric and the binary variables contributed to the final solution.

2.6.3.3 Step 3c. LCA

LCA is suitable for use with categorical data, so the inclusion of binary variables was not an issue. The change that was required for LCA was that it could not include the numeric variables that were used in clustering and PCA, i.e., the %TEI each participant consumed from each of the generic meals. Instead, the dataset used for LCA included both a categorical variable and a binary variable for each meal type. The categorical variables contained the data on which generic meal the participant consumed for that meal type, e.g., for the categorical breakfast variable a participant may have consumed generic breakfast 1, generic breakfast 2, etc. The binary variables were the same as those used in the previous methods, i.e., an indication of consumption or non-consumption of each meal type. The other difference that arose with LCA was that there are two observations for each participant representing the two days of dietary recall. For the other methods, it was possible to calculate the mean %TEI arising across the two days of dietary recall giving rise to only one observation per participant, but this was not possible with the categorical data required for LCA.

2.6.4 Comparison of Methods

A number of approaches were taken to compare the meal patterns that were identified by the statistical methods described above. Firstly, the number of meal patterns identified by each of the methods were compared. The number of meals per day, the presence of meal skipping, and the different meal types consumed in the different meal patterns were also compared. Each meal pattern was defined in terms of the top two most commonly consumed generic meal at each meal type. The number of meal types that the meal patterns from different methods had in common was calculated.

2.7 Chapter 6. An Exploratory Analysis of Meal Patterns and their Relationships with Health: Findings from the NHANES Cohort

Given its ability to handle both numeric and binary data and identify meal patterns in which certain meal types were not consumed, the meal patterns identified in the previous chapter (Chapter 5) using the clustering approach were used in Chapter 6. The clustering used in the previous chapter had assigned each individual to a meal pattern. In Chapter 6 the relationships between adherence to certain meal patterns and nutrient intakes, demographics, and health parameters were explored.

To compare the variables, they were plotted on parallel plots with a different line on the plot for each meal pattern. This allowed for the visual representation and exploration of trends that exist in the variables of interest. The variables were z-standardised prior to plotting to allow multiple variables with different units to be plotted on the same axes. Specifically, the variables that were examined were participants' overall dietary quality as indicated by the Nutrient Rich Foods index (NRF9.3); intakes of energy (kcal), sugar (g), fat (g), MUFA (g), PUFA (g), SFA (g), dietary fibre (g), calcium (mg), iron (mg), sodium (mg), vitamin D (µg); age (years); BMI (kg/m^2); waist-to-hip ratio; body fat (%); systolic blood pressure (mmHg); diastolic blood pressure (mmHg); fasting glucose (mmol/l); total cholesterol (mmol/l); HDL Cholesterol (mmol/l); and LDL Cholesterol (mmol/l). This involved comparing these variables among the different meal patterns and was completed separately for males and females.

Given the exploratory approach of this chapter, confirmatory statistics and *p*-values were not used. In the data from the male participants, for example, the 13 meal patterns would give rise to 78 pairwise comparisons and comparing multiple variables among those 13 meal patterns would increase the total number of comparisons further. While corrections could be made for multiple comparisons, given the large number of comparisons involved, it would result in a very small significance level and therefore provide little meaningful information to aid interpretation. In the context of comparing health outcomes across different meal patterns, the use of *p*-values could give rise to false confidence in the interpretation of the results. In order to quantify the uncertainty around the medians, confidence intervals were calculated.

2.8 Chapter 7. Dietary Intake Assessment Using a Novel Generic Meal-Based Recall and a 24-Hour Recall: A Comparison Study

In this chapter a new generic meal-based method of dietary recall was compared with the established 24HR. The meal-based recall involved participants being presented with a series of meal images, separately for each meal type, and asked to choose the meal image that was most similar to the meal that they consumed for that meal type, as described below. The meal images were based on the generic meals identified from the NANS data described in Chapter 4. The Human Research Ethics Committee, University College Dublin granted ethical approval to conduct this study (LS-21-64-OHara-Gibney).

2.8.1 Participant Recruitment and Eligibility

The target sample size was 160 participants based on a previous review of comparisons between digital and paper-based 24HRs which found a range of sample sizes from 53 to 167 [33]. There were no studies on meal-based methods of dietary assessment on which to base the sample size. Recruitment was carried out using local radio, local newspaper, posters, social media, and word of mouth. While participation in the study was restricted to those living in Ireland, the remote web-based nature of the study allowed participants to take part from any part of the country. Two interviews were completed on local radio to assist with the recruitment of participants, one took place on the 7th of February 2022 on Ocean FM and the other took place on the 18th of February on Midwest Radio. An article also appeared in the newspaper The Sligo Weekender on the 7th of April under the heading "Volunteers sought for online study", describing the study and providing information as to how participants can find out more information. Posters (Appendix 1) were distributed throughout UCD's Belfield campus and were also used on social media. An article was also published on RTÉ Brainstorm [34] providing information more generally about dietary intake assessment and also specific details about the study to encourage participation. People were eligible to take part if they were over 18 years of age, were fluent in English, had regular access to the internet, and were not current or former students of a degree in nutrition or dietetics.

2.8.2 Overview of Study Design

All those interested in taking part were directed to the study webpage which was created on Qualtrics® XM. Participants could read the participant information sheet including contact details for the researchers and complete the eligibility screening questionnaire (Appendix 2) to take part in the study via the webpage. Those who were eligible were then automatically directed to complete an online consent form to provide electronic informed consent.

Participants who were deemed eligible to take part and provided consent, were contacted by email providing details on the next steps of the study and the links to the two web-based dietary intake assessment tools (the 24HR and the generic meal-based recall). A crossover design was used with regard to the order in which participants completed the recalls. Participants were randomised to complete one or other of the two methods first, and then complete the second method at least 3 hours later on the same day. Participants were also randomised as to whether they would recall a weekday or a weekend day. Two weeks after having completed the first set of recalls, participants completed the recalls again in the

reverse order to which they were completed on the first occasion, followed by the completion of the evaluation questionnaire (**Figure 2.3**).

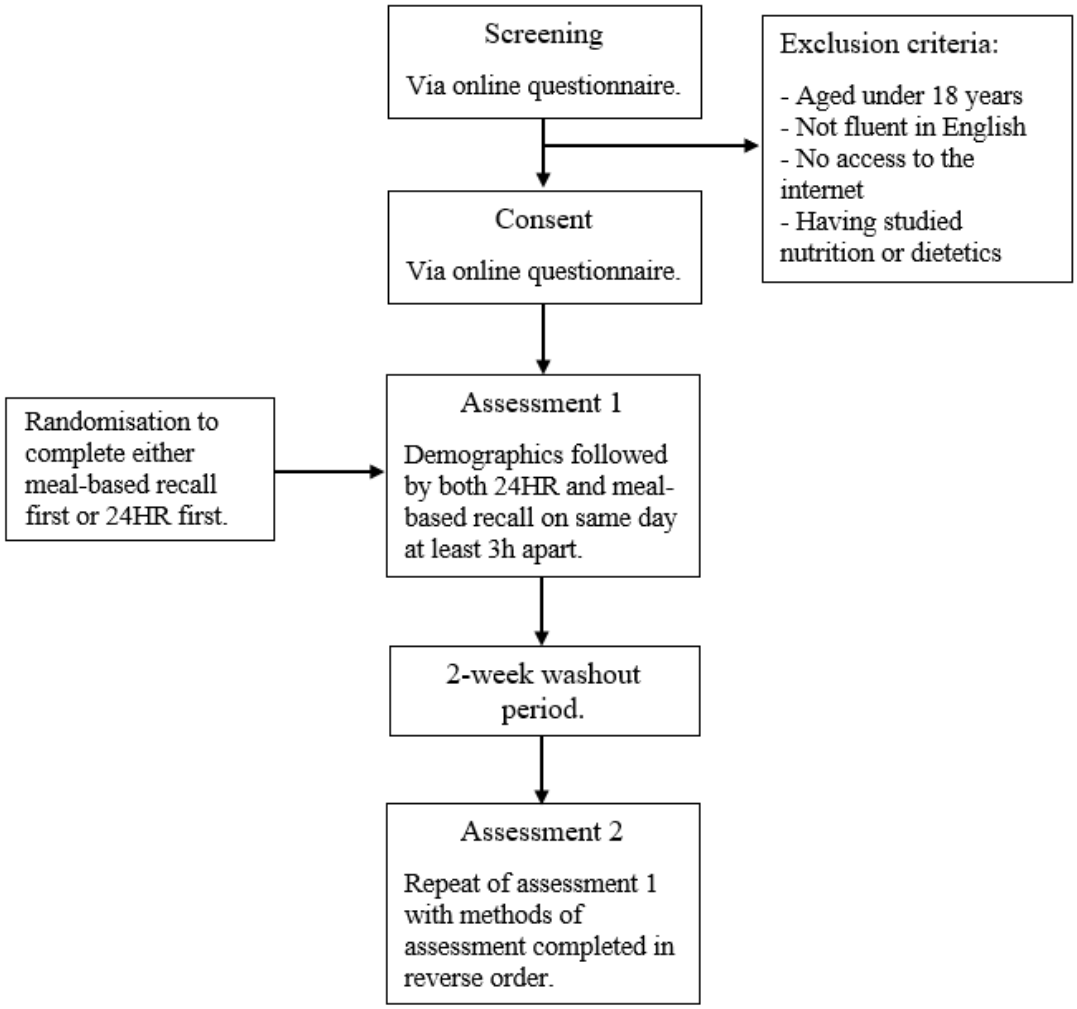


Figure 2.3: Overview of study design used for comparison between the generic recall and the 24-hour recall (24HR).

2.8.3 Overview of the Generic Meal-Based Dietary Recall

The meal-based dietary intake assessment method was administered using Qualtrics® XM, a web-based platform for questionnaires. Prior to completing the dietary intake assessment, participants provided information on their sex, age, weight, and height via an online questionnaire. Participants were then asked to select the meal types they consumed on the previous day from the following list: breakfast, morning snack, lunch, afternoon snack, evening meal, evening snack, and beverage (for beverages consumed alone without food). For each meal type selected, participants were presented with a series of images of generic meals that are associated with that meal type and asked to choose the meal image that

was most similar to the meal that they consumed. For each selected meal, participants were asked to choose from three different images of that meal, each representing a different portion size, and then asked to select whether the chosen image was smaller than the chosen portion, the same size as the portion chosen, or larger than the portion chosen. For beverages, participants were asked to choose from three different images of that beverage representing three different portion sizes and then asked to select the number of those portion sizes that they consumed. For each meal type, participants could choose the option that none of the images presented were representative of their intake for that meal type, e.g., "none of the above options are similar to what I ate for breakfast". If participants selected this option, a box appeared to enter a free text description of what they consumed for that meal. This allowed the researchers to determine whether a suitable generic meal could have been chosen or whether there was in fact no matching generic meal.

The images of generic meals presented to participants were based on the generic meals identified from the NANS data in Chapter 4, the methods for which are described above (Section 2.5). Of the 63 generic meals identified in that chapter, 31 were consumed during the week and 32 during the weekend [6]; there was, however, overlap between weekday and weekend meals. That is, some weekday meals were the same as the weekend meals. When these duplicates were removed, participants were presented with 43 meal images: 5 breakfasts, 5 snacks (repeated for morning, afternoon, and evening snacks), 10 lunches, 19 dinners, and 4 beverages.

The generic meal images were created by preparing the meals and taking photographs of them. To determine what foods to include in the photographs, the food groups that appeared in more than 50% of the individual meals in a given generic meal were examined. The most frequent foods appearing in each of those food groups were included in the photograph of that generic meal. For example, one of the breakfast generic meals contained one food group that appeared in more than 50% of the individual meals in that generic meal. That food group was the breads, oats, pasta, and rice group, and the most frequently occurring foods were bread and porridge which were included as the photograph representing that meal. Three photographs, representing three different portion sizes were taken of each generic meal. As described in Section 2.5.2 above about Chapter 4, each generic meal had seven portion sizes assigned by ordering each of the individual meals by weight and dividing the meals into sevenths based on septile values for meal weight. The median weights for each seventh were assigned as the generic portion size for that meal. The nutrient composition for each of the portion sizes was calculated using the meal weight for that portion and the generic meal nutrient composition [6]. Within the generic meal recall presented in Chapter 7, the second, fourth, and sixth portion size were used as the three

portion size images shown to participants, with the options asking whether the image chosen was smaller, the same, or larger than that consumed allowing participants to be categorised as consuming the first, third, fifth, or seventh portion size for a given meal.

2.8.4 24-Hour Recall

Participants completed their 24HRs using a validated web-based self-administered 24HR tool called Foodbook24 which follows the multi-pass recall method [35-37]. Participants first chose the meal types that they consumed from the following list: breakfast, morning snack, lunch, afternoon snack, evening meal, and evening snack with the option to add additional snacks. For each of the selected meal types, participants added the foods and beverages they consumed at those meal types by text-searching from the food list using a search bar. Portion size was then reported based on the number of the food or beverage items consumed and/or from portion size photographs as appropriate. Participants were then presented with the list of foods they had recorded for review, before being presented with a list of commonly forgotten foods. The food list contained food composition data from McCance and Widdowson's The Composition of Foods Integrated Dataset 2021 with some additions relevant to dietary intakes in Ireland. The development of Foodbook24 and its food list is described in detail elsewhere [35, 37].

2.8.5 Statistical Analysis

All analysis was carried out using R version 4.2.2 [19] in the RStudio integrated development environment (version 2022.07.2+576) [20]. Data from the 24HRs were used to identify participants likely to be mis-reporters of energy intake, based on a ratio of estimated energy intake to basal metabolic rate (EI:BMR) using the BMR equations from Henry [38]. Based on the Goldberg equations [39], an EI:BMR less than 0.96 was deemed indicative of under-reporting, with a ratio greater than 2.49 indicative of over-reporting. The analysis presented in Chapter 7 includes all participants, given the negligible differences observed when mis-reporters were removed.

The Shapiro-Wilk test was used to determine whether the differences in the variables between methods were normally distributed and confirmed with visual inspection of histograms. Wilcoxon's signed rank test was carried out to compare nutrient intake estimates arising from the web-based 24HR with estimates arising from the generic meal-based recall. Wilcoxon's effect size *r* was calculated. An effect size of ≥ 0.1 and < 0.3 was considered small, ≥ 0.3 and <0.5 was considered moderate, and ≥ 0.5 was considered large [40]. Bland–Altman analysis was also carried out whereby the mean difference between the

two datasets and the limits of agreement (mean difference ± 1.96 SD) for each nutrient were calculated. The correlation of nutrient intakes between the two methods was assessed using Spearman rank correlation coefficients. A correlation coefficient of <0.20 was considered poor correlation, a coefficient of ≥0.20 and <0.50 was considered acceptable correlation, and a coefficient of ≥0.5 was considered good correlation [41].

Cross-classification of quartiles was performed for all nutrients assessed. That is, nutrient intakes from both methods were divided into quartiles to determine the proportion of participants who remained in the same quartile for both methods (exact agreement), the proportion who were classified in the same or adjacent quartiles (exact + adjacent), the proportion who were classified 2 quartiles apart (disagreement), and the proportion who were classified 3 quartiles apart (extreme disagreement). Participants were also classified, separately for both methods, according to nutrient-based dietary guidelines [22-24]. For example, they were classified as to whether their nutrient intakes were low, adequate, or high according to those guidelines. The nutrients assessed were those deemed to be of public health relevance and included protein, carbohydrate, fat, monounsaturated fat, polyunsaturated fat, saturated fat, salt, dietary fiber, calcium, iron, folate, thiamin, riboflavin, and vitamin C. The proportion of individuals who were classified into the same category based on both methods was calculated for each nutrient.

2.8.6 Evaluation Questionnaire

On completion of the study, participants were asked to complete an evaluation questionnaire giving their opinions on whether the meals presented in the meal-based dietary intake assessment tool were representative of what they consume, to rate the clarity and ease of use of the tool, and whether they preferred that tool to the food-based tool (Appendix 3).

2.9 References

1. Irish Universities Nutrition Alliance (2011) National Adult Nutrition Survey (2008–2010). Available at: https://www.iuna.net/surveyreports (accessed 14 December 2023).

2. Centers for Disease Control and Prevention (CDC) (2018) National Health and Nutrition Examination Survey (NHANES) Data 2017-2018. Available at: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 (accessed 14 December 2023).

3. U.S. Department of Agriculture (2020) USDA Food and Nutrient Database for Dietary Studies 2017-2018. Available at: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds-download-databases/ (accessed 14 December 2023).

4. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

5. Food Safety Authority of Ireland (2019) Healthy Eating, Food Safety and Food Legislation: A Guide Supporting the Healthy Ireland Food Pyramid. Dublin, Ireland: *Food Safety Authority of Ireland*.

6. O'Hara C, O'Sullivan A, and Gibney ER (2022) A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data. *J Nutr*, **152**(10), 2297-2308.

7. James G, Witten D, Hastie T*, et al.* (2013) An Introduction to Statistical Learning with Applications in R. New York, USA: Springer.

8. Fulgoni VL, Keast DR, and Drewnowski A (2009) Development and Validation of the Nutrient-Rich Foods Index: A Tool to Measure Nutritional Quality of Foods. *J Nutr*, **139**(8), 1549-1554.

9. Charrad M, Ghazzali N, Boiteau V*, et al.* (2014) Nbclust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, **61**(6), 1-36.

10. Regulation (EU) No 1169/2011 (2011) Regulation of the European Parliament and of the Council of 25 October 2011 on the Provision of Food Information to Consumers. European Union.

11. EFSA (European Food Safety Authority) (2019) Dietary Reference Values for Nutrients. Summary Report. no. e15121.

12. Berendsen AAM, van Lieshout L, van den Heuvel E*, et al.* (2016) Conventional Foods, Followed by Dietary Supplements and Fortified Foods, Are the Key Sources of Vitamin D, Vitamin B6, and Selenium Intake in Dutch Participants of the Nu-Age Study. *Nutr Res*, **36**(10), 1171-1181.

13. Sluik D, Streppel MT, van Lee L*, et al.* (2015) Evaluation of a Nutrient-Rich Food Index Score in the Netherlands. *J Nutr Sci*, **4**, e14.

14. Streppel MT, Sluik D, van Yperen JF*, et al.* (2014) Nutrient-Rich Foods, Cardiovascular Diseases and All-Cause Mortality: The Rotterdam Study. *Eur J Clin Nutr*, **68**(6), 741-747.

15. Louie JC, Moshtaghian H, Boylan S*, et al.* (2015) A Systematic Methodology to Estimate Added Sugar Content of Foods. *Eur J Clin Nutr*, **69**(2), 154-161.

16. United States Food and Drug Administration (2020) Nutrition Labeling of Food. In *Code of Federal Regulations 21CFR1019*.

17. Kaufman L and Rousseeuw PJ (1990) Partitioning around Medoids (Program PAM). In *Finding Groups in Data: An Introduction to Cluster Analysis*, pp. 68-125. Hoboken, NJ, USA: John Wiley & Sons Inc.

18. Maechler M, Rousseeuw PJ, Struyf A*, et al.* (2022) Cluster: Cluster Analysis Basics and Extensions. R Package Version 2.1.3. Available at: https://CRAN.R-project.org/package=cluster (accessed 30 May 2022).

19. R Core Team (2022) R: A Language and Environment for Statistical Computing, 4.2.2 ed. Vienna, Austria: R Foundation for Statistical Computing.

20. RStudio Team (2022) Rstudio: Integrated Development for R., 2022.07.2+576 ed. Boston, MA, USA: RStudio PBC.

21. Bland MJ and Altman DG (1999) Measuring Agreement in Method Comparison Studies. *Stat Methods Med Res*, **8**(2), 135-160.

22. Institute of Medicine (2005) Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids. Washington, DC, USA: The National Academies Press.

23. Institute of Medicine (2000) Dietary Reference Intakes: Applications in Dietary Assessment. Washington, DC, USA: The National Academies Press.

24. World Health Organization (2007) Protein and Amino Acid Requirements in Human Nutrition, *WHO Technical Report Series.* Geneva, Switzerland: WHO Press.

25. U.S. Department of Agriculture (2023) Learn How to Eat Healthy with Myplate. Available at: https://www.myplate.gov/ (accessed 12th November 2023).

26. Bennett E, Peters SAE, and Woodward M (2018) Sex Differences in Macronutrient Intake and Adherence to Dietary Recommendations: Findings from the UK Biobank. *BMJ Open*, **8**(4), e020017.

27. Imamura F, Micha R, Khatibzadeh S*, et al.* (2015) Dietary Quality among Men and Women in 187 Countries in 1990 and 2010: A Systematic Assessment. *Lancet Glob Health*, **3**(3), e132-142.

28. Kant AK and Graubard BI (2015) 40-Year Trends in Meal and Snack Eating Behaviors of American Adults. *J Acad Nutr Diet*, **115**(1), 50-63.

29. Mathias KC, Almoosawi S, and Karagounis LG (2017) Protein and Energy Intakes Are Skewed toward the Evening among Children and Adolescents in the United States: NHANES 2013–2014. *J Nutr*, **147**(6), 1160-1166.

30. Barrera CM, Moore LV, Perrine CG*, et al.* (2019) Number of Eating Occasions and Source of Foods and Drinks among Young Children in the United States: NHANES, 2009-2014. *Nutrients*, **11**(4), 897.

31. Sullivan VK, Na M, Proctor DN*, et al.* (2021) Consumption of Dried Fruits Is Associated with Greater Intakes of Underconsumed Nutrients, Higher Total Energy Intakes, and Better Diet Quality in US Adults: A Cross-Sectional Analysis of the National Health and Nutrition Examination Survey, 2007-2016. *J Acad Nutr Diet*, **121**(7), 1258-1272.

32. Drewnowski A (2009) Defining Nutrient Density: Development and Validation of the Nutrient Rich Foods Index. *J Am Coll Nutr*, **28**(4), 421S-426S.

33. Timon CM, van den Barg R, Blain RJ*, et al.* (2016) A Review of the Design and Validation of Web- and Computer-Based 24-H Dietary Recall Tools. *Nutr Res Rev*, **29**(2), 268-280.

34. O'Hara C (2022) We Know Nutrition Apps Are Popular, but Are They Useful? Available at: https://www.rte.ie/brainstorm/2022/0303/1284230-nutrition-apps-food-behaviour-research/ (accessed 19th October 2023).

35. Evans K, Hennessy A, Walton J*, et al.* (2017) Development and Evaluation of a Concise Food List for Use in a Web-Based 24-H Dietary Recall Tool. *J Nutr Sci*, **6**, e46.

36. Timon C, Evans K, Kehoe L*, et al.* (2017) Comparison of a Web-Based 24-H Dietary Recall Tool (Foodbook24) to an Interviewer-Led 24-H Dietary Recall. *Nutrients*, **9**(5), 425.

37. Timon CM, Blain RJ, McNulty B*, et al.* (2017) The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. *J Med Internet Res*, **19**(5), e158.

38. Henry CJ (2005) Basal Metabolic Rate Studies in Humans: Measurement and Development of New Equations. *Public Health Nutr*, **8**(7A), 1133-1152.

39. Black AE (2000) Critical Evaluation of Energy Intake Using the Goldberg Cut-Off for Energy Intake:Basal Metabolic Rate. A Practical Guide to Its Calculation, Use and Limitations. *Int J Obes*, **24**, 1119-1130.

40. Cohen J (1992) A Power Primer. *Psychol Bull*, **112**(1), 115–159.

41. Lombard MJ, Steyn NP, Charlton KE*, et al.* (2015) Application and Interpretation of Multiple Statistical Tests to Evaluate Validity of Dietary Intake Assessment Methods. *Nutr J*, **14**, 40.

2.10 Supplementary Material

Supplementary Table 2.1: The 24 indices used in determining the number of clusters in the clustering steps of identifying meal patterns.

Index	Interpretation
KL	Max value
CH	Max value
Hartigan	Max difference between hierarchy levels of index
CCC	Max value
Scott	Max difference between hierarchy levels of index
Marriot	Max value of second differences between levels of the index
TrCovW	Max difference between hierarchy levels of index
TraceW	Max value of second differences between levels of the index
Friedman	Max difference between hierarchy levels of index
Rubin	Min value of second differences between levels
Cindex	Min value
DB	Min value
Silhouette	Max value
Duda	Smallest number of clusters such that index > critical value
Beale	Number of clusters such that critical value ≥ alpha
Ratkowsky	Max value
Ball	Max difference between hierarchy levels of index
Ptbiserial	Max value
McClain	Min value
Dunn	Max value
Hubert	Graphical method: cluster number where a peak occurred on the second differences plot.
SDindex	Min value
Dindex	Graphical method: cluster number where a peak occurred on the second differences plot.
SDbw	Min value

Information on the interpretation of the various indices was adapted from Charrad *et al.* [9].

Supplementary Table 2.2: The generic meals identified in the National Adult Nutrition Survey dataset.

Time of Week	Meal Type	Food-Based Description	Portion Sizes in Grams			Nutrient content per 100g			
			P2	P4	P6	Energy (kcal)	Total Fat (g)	Protein (g)	Carbohydrate (g)
Weekend	Breakfast	Bread/toast with fat spread, tea with milk, eggs, sausages, rashers, and pudding	371.5	515.0	747.0	89.2	4.0	3.8	9.1
Weekend	Breakfast	Bread/toast, with fat spread and marmalade/jam. Tea with milk and sugar.	338.0	480.0	678.0	84.2	3.0	2.3	12.3
Weekend	Breakfast	Bread/porridge, tea with milk, and fruit juice/whole fruit	379.0	567.5	845.0	61.1	1.8	2.2	9.5
Weekend	Breakfast	Cereal with tea and milk	251.5	425.0	593.5	74.9	1.4	2.8	13.4
Weekday	Breakfast	Toast or porridge	136.5	259.0	462.0	112.9	4.1	3.8	15.7
Weekday	Breakfast	Bread/toast with fat spread and jam/marmalade, tea with milk and sugar.	340.0	458.0	649.0	74.1	2.7	2.2	10.5
Weekday	Breakfast	Bread or porridge, tea with milk, and fruit juice/whole fruit	457.0	619.0	881.5	49.7	1.3	1.8	8.1
Weekday	Breakfast	Cereal with milk and tea	275.0	443.0	630.0	72.2	1.3	2.7	12.9
Weekend	Beverage	Spirits or lager with cola or red bull	250.0	467.5	915.0	57.0	0.5	0.0	8.0
Weekend	Beverage	Lager/stout	378.0	662.5	1665.0	41.5	0.0	0.3	1.2
Weekend	Beverage	Milk and tea/coffee	245.0	275.0	332.0	11.2	0.4	0.6	1.4
Weekend	Beverage	Water	250.0	366.0	628.0	0.0	0.0	0.0	0.0

Time of Week	Meal Type	Food-Based Description	Portion Sizes in Grams			Nutrient content per 100g			
			P2	P4	P6	Energy (kcal)	Total Fat (g)	Protein (g)	Carbohydrate (g)
Weekday	Beverage	Coffee/tea with milk and sugar	255.0	284.0	330.0	23.0	0.4	0.4	4.7
Weekday	Beverage	Water	225.0	330.0	568.0	0.0	0.0	0.0	0.0
Weekday	Beverage	Tea and milk	232.0	262.0	315.0	7.2	0.3	0.5	0.6
Weekday	Beverage	Lager/stout	375.0	637.5	1148.0	48.0	0.0	0.2	1.7
Weekend	Light meal	Sandwich with bread, mayonnaise/dressing, meat with veg and tea with milk	509.0	631.0	846.0	69.0	3.0	3.5	7.3
Weekend	Light meal	Sandwich with bread, fat spread, meat/eggs, tea and milk	378.0	500.0	741.0	98.2	4.7	4.3	9.4
Weekend	Light meal	Bread and soup and fruit	235.0	383.5	648.0	126.5	5.3	5.2	14.1
Weekend	Light meal	Sandwich with bread, cheese, ham, and tea	295.5	440.5	715.0	122.8	5.8	5.8	11.9
Weekend	Light meal	Sandwich with fat spread and meat, with biscuit/chocolate, tea, and milk	406.0	528.0	744.5	95.3	3.6	3.6	12.4
Weekend	Light meal	Biscuits/chocolate/cake with tea and milk	318.0	486.0	713.0	84.8	3.8	2.6	10.5
Weekday	Light meal	Sandwich with bread, fat spread, eggs, sausages, and bacon, with tea and milk and sugar	403.5	527.0	679.0	106.4	4.9	4.0	11.6
Weekday	Light meal	Bread, fat spread, with banana, and tea with milk	458.0	635.0	853.5	69.5	2.6	2.5	9.4

Time of Week	Meal Type	Food-Based Description	Portion Sizes in Grams			Nutrient content per 100g			
			P2	P4	P6	Energy (kcal)	Total Fat (g)	Protein (g)	Carbohydrate (g)
Weekday	Light meal	Bread with chicken and water	350.0	544.0	757.0	110.8	5.4	5.5	9.9
Weekday	Light meal	Ham sandwich with tea and milk	365.0	479.5	709.0	88.3	3.1	4.9	10.6
Weekday	Light meal	Bread with soup and/or whole fruit	243.0	380.0	620.0	109.0	4.2	4.0	13.6
Weekday	Light meal	Sandwich with bread, mayonnaise/dressing, meat with veg and water	417.0	588.0	853.0	89.3	3.9	4.8	9.0
Weekday	Light meal	Sandwich with bread, cheese, fat spread, ham, with tea and milk	400.0	515.0	710.0	110.1	5.1	4.6	11.9
Weekday	Light meal	Tea and milk with biscuit/chocolate/cake	321.0	469.0	676.0	83.3	3.5	2.6	10.5
Weekend	Main meal	Potatoes with beef, veg, tea, milk, and sweet dessert	770.5	1062.0	1268.5	82.1	3.3	4.8	8.3
Weekend	Main meal	Bread/rice with chicken curry /stir fry	518.0	794.5	1123.0	90.4	2.9	5.4	10.1
Weekend	Main meal	Chips/potatoes with meat and tea	507.5	678.0	943.0	97.8	4.8	4.7	8.8
Weekend	Main meal	Potatoes, ham/beef, green veg, and water	595.0	825.0	1109.0	71.8	3.1	5.3	5.8
Weekend	Main meal	Potatoes, beef, gravy, carrots, and green veg	404.0	621.0	861.0	101.9	4.6	6.8	8.5

Time of Week	Meal Type	Food-Based Description	Portion Sizes in Grams			Nutrient content per 100g			
			P2	P4	P6	Energy (kcal)	Total Fat (g)	Protein (g)	Carbohydrate (g)
Weekend	Main meal	Potatoes, beef, veg, gravy, water	637.0	909.0	1241.0	79.5	3.6	4.7	6.8
Weekend	Main meal	Potatoes with chicken, veg, and water	651.0	867.0	1175.5	62.5	2.1	4.9	6.6
Weekend	Main meal	Bread with fat spread, ham, egg, onions, coleslaw, tea and milk	641.0	834.0	1007.0	99.6	4.8	4.5	8.8
Weekend	Main meal	Potatoes/chips, beef, veg, sugar-sweetened beverage	498.0	715.5	935.5	121.1	5.2	6.3	11.9
Weekend	Main meal	Bread, rice/spaghetti, sauce, chicken, pulses, veg, and water	577.5	817.0	1082.0	84.3	3.5	4.9	8.3
Weekend	Main meal	Potatoes, beef/bacon/ham, vegetables, biscuit, tea, and milk	691.0	893.5	1159.5	77.5	3.2	4.4	7.4
Weekend	Main meal	Bread with meat, tea, milk, and biscuits	596.0	771.0	1108.0	107.5	4.9	4.7	11.2
Weekend	Main meal	Rice/bread/pizza with beef/poultry dish	345.5	546.5	853.5	150.4	6.3	6.7	14.6
Weekend	Main meal	Potatoes with peas and meat	348.0	571.0	825.0	119.8	5.0	7.6	10.6
Weekday	Main meal	Bread, beef dishes, tea, milk, and sugar	533.0	740.0	1041.0	98.6	4.2	4.7	10.3
Weekday	Main meal	Potatoes, meat/fish, and veg	430.0	604.5	877.5	101.1	4.4	6.5	8.7
Weekday	Main meal	Potatoes, chicken, peas	382.0	570.5	790.0	106.1	3.9	7.7	10.7

Time of Week	Meal Type	Food-Based Description	Portion Sizes in Grams			Nutrient content per 100g			
			P2	P4	P6	Energy (kcal)	Total Fat (g)	Protein (g)	Carbohydrate (g)
Weekday	Main meal	Rice/pasta/bread, meat products, beg, sauce, and water	594.0	822.5	1083.0	88.4	3.9	4.7	8.3
Weekday	Main meal	Rice/pasta, chicken dishes, water	541.5	764.0	1025.0	80.7	2.5	5.4	9.3
Weekday	Main meal	Rice/pasta/bread, chicken/fish, veg, tea, milk, sugar, chocolate	682.0	835.0	1036.0	86.8	3.2	4.4	10.1
Weekday	Main meal	Potatoes, chicken/fish, veg, water, tea, milk	663.0	897.0	1097.0	62.9	2.1	4.9	6.5
Weekday	Main meal	Rice/pasta/bread, beef/chicken	347.0	525.0	776.5	138.6	5.7	7.1	13.7
Weekday	Main meal	Potatoes, beef/ham, veg	702.0	909.0	1163.0	90.1	4.0	4.7	9.1
Weekday	Main meal	Chips/potatoes, fish/beef/bacon/chicken, water	527.0	716.0	979.0	99.3	4.7	5.2	9.3
Weekday	Main meal	Potatoes, beef/fish/chicken	565.0	782.0	1043.0	76.4	3.3	5.1	6.7
Weekend	Snack	Bread/scone with fat spread and tea with milk	295.5	398.0	546.0	96.2	3.9	3.3	12.4
Weekend	Snack	Biscuit/chocolate with tea and milk	252.0	312.0	403.0	52.1	2.1	1.3	7.4
Weekend	Snack	Biscuit/chocolate with water	66.0	267.0	522.0	154.5	6.8	2.1	21.8
Weekend	Snack	Piece of fruit	112.5	250.0	480.0	86.3	3.2	1.8	12.5
Weekday	Snack	Biscuits/chocolate	40.0	128.5	363.0	233.9	10.8	4.1	31.1

Time of Week	Meal Type	Food-Based Description	Portion Sizes in Grams			Nutrient content per 100g			
			P2	P4	P6	Energy (kcal)	Total Fat (g)	Protein (g)	Carbohydrate (g)
Weekday	Snack	Bread with fat spread and tea and milk	301.0	383.5	527.0	92.0	3.7	2.9	12.0
Weekday	Snack	Biscuits/chocolate and tea and milk and sugar	247.0	307.0	393.0	50.6	2.1	1.3	7.2
Weekday	Snack	Fruit and water	140.0	274.5	519.0	32.9	0.4	0.5	7.3

Seven portion sizes were calculated for each generic meal; three are given in the table as examples, i.e., the second (P2), fourth (P4), and sixth (P6) portion sizes are given.

CHAPTER 3

Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings

A version of this chapter has been published as:

O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

3.1 Abstract

There is a scarcity of dietary intake research focusing on the intake of whole meals rather than focusing on the nutrients and foods of which those meals are composed. This growing area of research has recently begun to utilise advanced statistical techniques to manage the large number of variables and permutations associated with these complex meal patterns. The aim of this narrative review was to evaluate those techniques and the meal patterns they detect. The ten observational studies identified used techniques such as principal component analysis, clustering, latent class analysis, and decision trees. They examined meal patterns under three categories: temporal patterns—relating to the timing and distribution of meals; content patterns—relating to combinations of foods within a meal and combinations of those meals over a day; and context patterns—relating to external elements of the meal such as location, activities while eating, and the presence or absence of others. The most common temporal meal patterns were the three meals per day pattern, the skipped breakfast pattern, and a grazing pattern consisting of smaller but more frequent meals. The three meals per day pattern was associated with increased diet quality compared to the other two patterns. Studies identified between seven and twelve content patterns with limited similarities between studies and no clear associations between the patterns and diet quality or health. One study has simultaneously examined temporal and context meal patterns, finding limited associations with diet quality. No study has simultaneously examined other combinations of meal patterns. Future research that further develops the statistical techniques required for meal pattern analysis is necessary to clarify the relationships between meal patterns and diet quality and health.

3.2 Introduction

Suboptimal dietary intake is known to contribute to the global burden of chronic non-communicable diseases, including cardiovascular disease, type 2 diabetes, and certain cancers [1]. The accurate assessment of food intake is an essential element of identifying such diet–disease associations and developing advice to support appropriate change of dietary behaviours [2]. Traditionally, diet-disease epidemiological studies have focused on linking disease or risk of disease with individual nutrients or foods, with a focus in more recent decades to examining the links with dietary patterns (i.e., the combinations of foods consumed habitually that represents the diet as a whole) [3-7]. These approaches reveal valuable insights and have contributed to the development of both nutrient and food-based dietary guidelines [8]. Yet, unlike a meal-based approach, they do not consider foods consumed in combination as part of a meal, the distribution or number of those meals over a day, or the context in which those meals are consumed [9, 10].

A focus on meal level information in the examination of dietary links to disease risk may improve the provision of meal-based dietary advice and enhance existing food-based dietary guidelines, with practical benefits for individuals relating to meal planning and preparation [10]. Furthermore, it has been proposed that a meal-based approach may be more appropriate for personalised nutrition using internet and mobile technology by reducing user burden with regard to data input and supporting meal-based personalised dietary feedback and advice [11].

Research of meal patterns to date has primarily focused on relatively simple statistical approaches based on frequencies; for example, the total number of meals per day [12, 13], the percentage of meals consumed in certain locations with the presence or absence of other people [14, 15], and the percentage of meals that contained certain food groups or combinations of food groups without considering daily patterns [16, 17]. However, these approaches cannot simultaneously capture the multiple variables that define meal patterns such as those relating to the timing, distribution, content, and context of meals [18-20].

More recently studies have applied advanced statistical data mining techniques to the concept of meal patterns. These techniques can better capture the complexity that is inherent within dietary intake datasets, while incorporating the many variables that are required for meal pattern analysis [10, 21]. Despite the recent growth in their use, no review of these techniques as they apply to meal patterns has yet been published.

The objectives of this review are to firstly provide a brief conceptual and theoretical overview of advanced statistical techniques, as they apply to meal pattern analysis; and secondly to

critically review the research that has employed these techniques, examining the meal patterns identified and their associations with diet quality and health, while noting gaps in the literature that warrant further research.

3.3 Overview of Statistical Techniques

Meal pattern analysis is the identification of patterns that emerge from measured dietary intake variables such as the temporal aspects of meals, their content, and the context in which they are consumed. Individuals are then grouped with those who have similar patterns [10, 17].

The statistical approaches used in meal pattern analysis reduce an interminably vast number of possible patterns—arising from various combinations of foods, times, or contexts related to a given meal—to a smaller number that are representative of those found in the sample population [22]. This smaller number of patterns can be investigated for associations with diet quality or health [18]. The statistical approaches used in meal pattern analysis to date include principal component analysis (PCA), clustering, latent class analysis (LCA), and decision trees [9, 18-20, 23-28]. This section provides an overview of the underlying principles of these approaches as they relate to meal pattern analysis, summarised also in **Table 3.1**. There are a number of statistical techniques that have been used in preliminary steps to the above-mentioned techniques; however, as these preliminary steps were not intended to identify the meal patterns themselves, they are not discussed in this section.

Before discussing each individual method, it is important to note that there are considerable differences in the datasets used for the meal pattern analyses discussed in this review, and the methods chosen to collect the data are typically driven by the research question at hand and the feasibility of collecting the required data. The structure of the data used for such analysis can influence the approach taken and as such it is worth considering here before detailing the statistical techniques themselves. Several methods can be used to collect food intake data, for example diet records, recalls, and questionnaires. Diet records and recalls (e.g., 4-day food diaries or 24-hour recalls (24HRs)) collect information about each individual food or ingredient consumed over a specified time period [29, 30]. Portion sizes may be determined by estimates based on common household measures or photographs. Alternatively, in the case of diet records, foods may be weighed before consumption [29, 30]. Each observation in the dataset represents an individual participant and an individual food consumed by that participant with associated information for the quantity of food consumed, its nutrient content, and the time and meal type at which it was consumed. There will be multiple observations for each participant representing the multiple foods consumed over

the course of the recording period. The nutrient content of these foods can then be summed to give the total nutrient intake per day for each individual or as a mean intake for the sample population. When multiple days of food intake are recorded the mean daily nutrient intake for individuals or for the sample population can be calculated [29, 30].

On the other hand, questionnaires used in meal pattern analysis do not focus on individual foods consumed by participants, but rather on the meals themselves. The exact approach varies depending on the research question. For example Englund-Ögge *et al.* [23] asked participants to report the number of times they consume different meal types (e.g. breakfast, lunch, dinner) per week, Wilson *et al.* [24] asked participants to report whether they consumed certain meal types (nothing, snack, large meal, small meal) within certain time periods in the previous 24 hours, and Riou *et al.* [28] asked participants to report the number of meals they consume and the time at which they are consumed. While these approaches allow for the investigation of the temporal aspects of meal patterns, they do not allow for the investigation of the content of those meals which would require data arising from approaches such as diet records or recalls.

3.3.1 Principal Component Analysis

Principal component analysis (PCA) is a common statistical technique used to determine variation and uncover patterns in any dataset [22]. It is specifically a dimension reduction method, whereby the number of dimensions in a dataset is equal to the number of variables in that dataset [22]. The aim of the analysis is to reduce the number of dimensions by creating indices (i.e., weighted summations) of correlated variables. This reduction allows us to keep the variables which are most important in explaining the variance in the dataset [22] (**Table 3.1**).

Let us consider its use in meal pattern analysis. Using PCA to investigate population-based meal patterns, Woolhead *et al.* [9] examined the percentage energy contribution to overall energy intake from 63 meals (the variables) consumed in a national food consumption survey conducted in Ireland. These meals were defined by the food groups of which they were composed, thus allowing for the examination of meal patterns based on the content of meals rather than solely the timing or distribution of intakes over a given time period. Each participant had an observed percentage energy value for each of the variables (meals). Hypothetically, the percentage energy values for two of the meals could be plotted in a two-dimensional space such as a scatterplot to examine the relationships between them. However, to examine the full dataset and assess all possible combinations of just two meals would require $\frac{63!}{2!(63-2)!}$ comparisons, resulting in 1953 separate plots. This is clearly

not a feasible solution, and each plot would only capture a fraction of the variance in the total data. Woolhead *et al.* [9] thus used PCA to address this issue by assessing all 63 meals together i.e. examining the datapoints for percentage energy from each of the 63 meals in a 63-dimensional space [22].

PCA identifies components that are linear functions of all variables. In the example above [9], each component is a linear function of the 63 meal variables. However, not all variables will be equally important to all components, and this distinguishes the components from each other. The relative importance of a given variable on a particular component is quantified by the variable loading value. A small selection of variables with high absolute loading values are typically selected to characterise each component, for example, one component may have high loadings for fruit-based breakfast, sandwich-based light meal, and meat/fish with pasta/rice/potato and vegetables main meal; while another component may have high loadings for cooked breakfast and meat/fish with rice/potato/pasta and soups/sauces main meal [9].

3.3.2 Clustering

Clustering describes several different techniques used to identify subgroups, or clusters, within a given dataset (**Table 3.1**). The aim of clustering analysis is to group observations that are least dissimilar among themselves, but most dissimilar from observations in other clusters [22]. These approaches have been applied to temporal meal patterns [18, 19] and to a combined assessment of temporal and context meal patterns [28]. All three studies used different methods of clustering. No study has been identified that has applied these techniques to content meal patterns. Clustering methods are not robust in the presence of missing data. While none of the studies reviewed here reported missing values in the variables used for clustering, a variety of methods have been proposed to deal with this issue and are discussed at length elsewhere [31-33].

There are several clustering techniques, some of which have been used in meal pattern analysis and will be discussed here. Hierarchical clustering is an iterative process that starts with each cluster containing just a single observation and ends with only one cluster composed of all observations grouped together [22]. For example, each observation assessed by Chau *et al.* [18], who examined meal patterns using this method in 4,508 adults in Taiwan, can be represented by a vector containing six elements, each corresponding to the energy intake of a single participant during one of six 4-hour periods in a day. At each step in hierarchical clustering the two of these vectors or groups of these vectors that are

least dissimilar from each other are joined, until the desired number of clusters are achieved.

Partitioning around the medoids (PAM) clustering is another iterative process. Unlike hierarchical clustering, the number of clusters sought is pre-specified. Initially, observations are randomly assigned to the chosen number of clusters. The medoid of each cluster is then determined; this is the observation in the cluster that is closest to the centre of the cluster [34]. Observations are then re-assigned to the cluster with the nearest medoid. The medoids for the newly formed clusters are then re-determined and observations re-assigned. This process is repeated until there are no further changes to which cluster each observation belongs i.e. the dissimilarity within clusters is minimised [34].

Finally, *k*-means clustering is a similar approach to PAM, with the centroid (based on the mean of the variables in the cluster) being used in place of the medoid to minimise dissimilarity within clusters. Both approaches are limited to identifying clusters that can be separated by a straight line.

One of the decisions required when clustering is choosing the number of clusters to represent the data. Different approaches were taken in the studies reviewed here. Chau *et al.* [18] selected five clusters to represent the data as they explained a considerable proportion of the variance (55%), and to choose six clusters would only explain a small fraction more (0.5%) of the variance and would render the clusters more difficult to interpret. Riou *et al.* [28], who used clustering to assess meal patterns in 2,994 adults in France, selected five clusters based on resampling and a cluster-robustness approach called consensus clustering, which identifies the number of clusters that provide the most stable results across multiple samplings [35]. There are numerous other procedures that can be carried out that aim to estimate the optimal number of clusters for a dataset as examined in detail elsewhere [36, 37].

3.3.3 Latent Class Analysis

Latent class analysis (LCA) aims to identify an unobserved, or latent, variable that represents some number of observed categorical variables [38]. It assumes that this latent variable is composed of a number of mutually exclusive and exhaustive classes, by this we mean that each participant can only belong to one class and that all participants are assigned to a class. This allows for the identification of subgroups within a sample, based on patterns of multiple observed variables [39]. LCA has been applied to both temporal meal patterns [20] and content meal patterns [26]. In the context of temporal meal patterns [20] the observed variables applied were binary, denoting the presence or absence of an eating

occasion during each hour of the day. This gives 24 time periods with two possible observations for each participant, amounting to 2^{24} possible unique patterns of intake. In the case of content meal patterns [26] the data were reduced to the three meal types (breakfast, light meal, and main meal) as observed variables, within which each participant was categorised to have consumed one of 5, 7, or 5 meals, respectively. This amounts to $5 \times 7 \times 5 = 175$ possible patterns. LCA allows these large numbers of possible patterns to be reduced to a smaller number of latent classes representing the patterns that exist in the sample population [38] (**Table 3.1**).

The number of latent classes must be specified before fitting a latent class model to the data. The two parameters that must be estimated in order to fit the model to the specified number of classes are the prevalence of each latent class and the probabilities of observing each of the variable categories within each class [38]. In the temporal example above, this is the probability of the presence of an eating occasion during each hour of the day [20]. For example, of the three latent classes reported, 43% of participants belonged to the "Conventional" latent class that had a high probability of consuming a meal at midday and 18:00h. In the content meal pattern example this is the probability of consumption of one of the meals at each meal type [26]. For example, one of the four classes reported comprised of 9% of the participants who had relatively high probabilities of consuming a cooked breakfast, of skipping a light meal and consuming a protein- and carbohydrate-based main meal.

3.3.4 Decision Trees

Only one of the identified studies applied a supervised statistical approach to meal pattern analysis [27] (**Table 3.1**). Supervised approaches aim to use the input variables to predict some outcome variable [40]. This is unlike the unsupervised approaches described in the preceding sections where the outcome variable is absent; instead, the aim is to determine associations and patterns among the input variables [40]. While different types of decision tree methods are available, Hearty and Gibney [27] applied a C5 decision tree approach using meal intake at either breakfast or main meal to predict whether an individual's diet scored in either the first or fifth quintile of the healthy eating index.

It is possible to use decision trees with both continuous and categorical data, they are generally easy to interpret, and can be represented graphically [41]. The decision tree can be represented in a format similar to a hierarchy or tree diagram, where the top of the diagram represents the full dataset, and it is split into specific subsets at each branch [22]. In the case of the study by Hearty and Gibney [27], who applied decision trees separately to

the breakfast meal type and the main meal type, the top of the diagram represented all participants in the study each with an associated variable for various food combinations (meals). The values of these variables are either 1 or 0 defining, respectively, whether the given meal was consumed or not at each meal type. Participants were then split into two subgroups based on a rule applied to the dataset. The rules applied by Hearty and Gibney [27] were based on the presence or absence of various meals at either breakfast or main meal for each participant. For example, the first rule split the participants by assigning those who consumed the "Bread and Confect/Snack" meal at breakfast to one subgroup and those who did not to another subgroup. This process is repeated by applying additional rules to each new subgroup to create further subgroups until no further subgroups can be created. Hearty and Gibney [27] applied a stopping rule which only allowed subgroups to be further split if they contained at least 75 records.

The number of rules applied and the order in which those rules are applied will impact on the overall outcome of the final tree. However, given the vast number of combinations involved, it is not computationally possible to compare all trees that could arise from a given dataset. Instead, a "nonbacktracking" or "greedy" approach is used; at each split in the tree, the best rule is chosen based on that split alone and not on the potential impact it may have at subsequent splits in the tree [41]. Several methods are available for choosing the best rule at each step in the decision tree e.g. statistical significance, information gain, and error reduction [41]. The method chosen by Hearty and Gibney [27] was based on information theory; this uses the gain ratio which expresses the proportion of information that appears to aid prediction that is generated by the different possible rules. The rule with the highest gain ratio at each step is used to split the participants into subgroups [41].

3.4 Meal Patterns

Studies are reviewed here under the headings of temporal patterns, content patterns, and combined patterns. There is no section for context patterns as no study was identified that applied advanced statistical techniques to these patterns alone; however, one study has investigated the combined patterns of both the temporal and context aspects of meal consumption [28]. No other analyses of combinations of different meal pattern types were identified.

3.4.1 Temporal Patterns

Temporal meal patterns refer to those accounting for the distribution of dietary intake over a given time, typically 24 hours. In published papers to date, statistical methods used to

identify temporal meal patterns have included PCA [23, 24], LCA [20], and clustering [18, 19]. The three studies using either PCA or LCA all identified three patterns, while those using clustering identified four to five patterns (**Table 3.2**). Most studies divided the day into time periods of varying durations from one hour [19, 20] to four hours [18], or considered periods of different durations throughout the day (i.e., five 3-hour periods, one 2-hour period, and one 7-hour period) [24].

The variables used for pattern identification at the various time periods also differed between studies and was influenced by the method of dietary data collection employed. The 24HR method was used by three studies [18-20] (**Table 3.2**). As mentioned above, 24HRs produce a detailed food file which lists each individual food or drink consumed within the preceding 24 hours as well as a portion (gram amount) for each food and the associated nutrients for each food. Each of these foods are reported within a specific meal/time context (e.g. breakfast, lunch, dinner, snack) which allows derivation of nutrient and energy intake at each meal; these data can then be used as input variables for meal pattern analysis. Using data from 24HRs, Chau *et al.* [18] used the energy content of each meal, while Khanna *et al.* [19] used energy content of each meal relative to total energy intake (% contribution to total energy). A binary variable was used by Leech *et al.* [20] denoting whether or not an eating event had occurred during each hour of the day; only eating events with greater than or equal to 210kJ were considered. The two studies using energy intake as an indicator can allow comparisons between groups in relation to the quantity (in terms of energy) consumed during the various time periods of the day [18, 19]. This is not captured, however, when only an indicator as to whether or not an eating event occurred during the various time periods [20]. The use of % contribution to total energy accounts for the fact that while individuals may differ in their total energy intake, they may have similar temporal patterns in relation to how that energy is distributed throughout the day [19].

The remaining two studies used questionnaire-based methods of dietary data collection which did not allow for the derivation of nutrient or energy intake at each meal because information regarding individual foods and portion sizes were not gathered using these methods. Englund-Ögge *et al.* [23] used the data associated with 8 different meal types (breakfast, morning snack, lunch, afternoon snack, dinner, evening snack, supper, night meal), where participants reported the frequency at which they consumed each meal, within the week, from 0 to 7 times per week. Finally, Wilson *et al.* [24] asked participants to report, for each time period, whether they ate nothing, a snack, a small meal, or a large meal and whether the drank nothing, alcohol, water, or something else. Points were assigned to the responses as follows: one point for a snack, three for a small meal, and five for a large meal; water was assigned no points with other beverages assigned one point. The number

of points during each time period was calculated relative to the total number of points in the day [24].

Although differing approaches to meal pattern analysis were applied across the studies, there were a number of similarities identified. One pattern that was similar across all studies was that which consisted of three meals per day with few or no snacks. The three meals typically happened at times that are culturally associated with breakfast, lunch, and dinner [18, 19, 23, 24]. A similar pattern was found by Leech *et al.* [20], with respect to lunch and dinner meals, but without reference to breakfast. This three meal per day type of pattern is associated with higher intakes of protein, polyunsaturated fat, calcium, phosphorous, vitamin D, and vitamin E and lower carbohydrate intake (adjusted for age, sex, education level, employment status, chronic disease, geography, and day of dietary recall) [18], and better overall diet quality (adjusted for sex, ethnicity, age group, BMI, survey year, and household poverty–income ratio) [42] compared to other identified patterns. Eicher-Miller *et al.* [42] also found that a greater proportion of those following this pattern had normal BMIs and a lower proportion had raised BMIs compared with other patterns. When applied to intakes up to 22 weeks gestation of a group of pregnant women in Norway (n = 65,487), those in the highest two quartiles for adherence to this type of pattern were found to have lower risk of preterm delivery compared to those in the lowest quartile for adherence to that pattern. This analysis adjusted for maternal age, pre-pregnancy BMI, height, parity, total energy intake, maternal education, marital status, smoking status, income, previous preterm delivery, fibre intake (as an indicator of overall healthy eating), alcohol intake, first trimester nausea, irregular work hours, and physical activity level. No associations with preterm delivery were identified within other meal patterns [23] (**Table 3.2**).

Some of the other patterns identified in these studies could be considered variations of this pattern. These typically also included three meals per day, but with self-reported consumption of “snack meals” rather than “main meals” at these times [23], with one to two additional intakes in the afternoon and/or night [18], or with timing of lunch intake later in the day than considered traditional. The later lunch pattern has been positively associated with systolic blood pressure and diastolic blood pressure among females compared to those following the more traditional timing of this pattern. The association remained after the adjustment for education level, country of birth, smoking status, physical activity, sleeping habits, overall diet quality, and BMI [43] (**Table 3.2**).

Another common pattern identified was that with little or no intake in the morning, typically known as breakfast skipping. Within those following such a pattern, peaks in intake typically happen later in the day, from noon onwards [18, 23, 24]. Wilson *et al.* [24] (adjusting for sex,

age, marital status, social support, education, work schedule, BMI, and smoking), examined the impact of meal patterns on mood in 1,304 adults in Australia, and found a higher prevalence of mood disorders (i.e. the lifetime prevalence of depressive symptoms) after 5 years in those whose adherence to this pattern had either increased or been consistently high compared to those with low adherence. No significant associations with mood disorders were identified among the other patterns either at baseline or at follow up (**Table 3.2**).

The final pattern that was common among multiple studies related to a pattern characterised by many smaller intakes spread over the day rather than a low number of larger distinct intakes. This was typically referred to as a grazing pattern [19, 24], and has been associated with the lowest diet quality compared with other patterns [42]. In the case of the study by Leech *et al.* [20] this pattern was one in the same with the pattern having little or no intake in the morning and found to be associated with the lowest diet quality compared to other patterns [44]. No associations were found however between that pattern and adiposity, adjusting for education level, country of birth, smoking, physical activity, and sleep duration. [44] (**Table 3.2**).

Khanna *et al.* [19] additionally identified two other patterns that were more likely to have largest intake in late afternoon and early evening and second largest in the morning and afternoon, or vice versa. No differences were identified with regard to diet quality between these patterns [42] (**Table 3.2**).

3.4.2 Content Patterns

Content patterns require a different approach to those outlined in the previous section on temporal patterns [10]. As the aim of analysing content meal patterns is to summarise the content of meals (in terms of food groups), it is not sufficient to only examine the variables used for temporal meal patterns. For example, only assessing the energy content of meal types or the energy consumed during various time periods of a day provides information about how intake is distributed between the meals or throughout the day but not about the types of foods that provided that energy [9]. The approaches taken, therefore, to the collection of data for the assessment of content meal patterns must gather the information describing actual food intakes. In addition, the statistical approaches employed must be able to reduce to huge number of possible food combinations that make up meals into a smaller number of interpretable groups.

Of the four studies identified that assessed content patterns, dietary intake data were collected using 4-day [9, 26], 7-day [27], and 4 × 4-day [25] food diaries (**Table 3.2**). As

discussed above, similar to the data collected by 24HR, these data are stored in a comprehensive food file, detailing each individual food consumed by each participant with the location, day, the time at which those foods were consumed, and whether they formed part of a meal in combination with other foods. The mass in grams of each consumed food is also recorded, allowing for the determination of the energy or nutrients consumed from a given food, meal, or during a given time period.

To reduce the huge number of possible combinations of unique foods eaten by a study population and allow for meaningful pattern analysis, all four studies first condensed the unique foods consumed into more aggregated food groups, resulting in 20 [9, 25, 26] and 62 [27] food groups in these studies, which were developed based on nutrient profile and culinary use of specific foods. Three of the studies then applied the frequent item sets data mining method to categorise the most commonly consumed food group combinations at each different meal type identifying 63 [9] and 80 [25] food group combinations at breakfasts, light meals, main meals, snacks, and beverages. Uzhova *et al.* [26] further aggregated these categories to 14 generic meals (breakfast, light meals, and main meals only). Hearty and Gibney [27] categorised generic meals based on the main food groups in each meal identifying 134 food group combinations, but do not appear to have used the frequent item sets method. The nutrient composition of these commonly consumed meals was then determined, and they were denoted as generic meals [27]. Variables from these generic meals were then used as the input to the statistical techniques (decision trees, PCA, and LCA) that identified meal patterns.

There are other methods which could conceivably be used to derive generic meals in place of the frequent item sets method including topic modelling, gaussian copula graphical models, and principal component analysis [45-47]. Like previous methods, the application of these methods also required foods first to first be condensed into food groups. White *et al.* [45] applied topic modelling to 60 food groups to identify generic meals. Each generic meal was based on the probability of each food group appearing within a given meal. Labels were assigned by the authors to the meals to describe the meal type in question (e.g. breakfast, lunch etc.) based on the top ten most probable food groups in a given generic meal. Fifteen generic meals were identified. Another approach used 39 food groups with semiparametric Gaussian copula graphical models to identify food groups that are correlated, and therefore likely to be eaten together as part of a meal [46]. Finally, Murakami *et al.* [47] used PCA to identify meal-specific dietary patterns, which could possibly be used to generate generic meals. Individual foods were condensed into 22 food groups, and PCA was carried out based on the amount of each food group consumed at each meal type (breakfast, lunch, or dinner). While these studies applied statistical techniques to identify

generic meals in their sample populations—a preliminary step in the analysis of content meal patterns—they did not go on to use these generic meals to assess meal patterns themselves (i.e. either how combinations of these generic meals are consumed over time or how they are related to health or diet quality).

Hearty and Gibney [27] applied artificial neural networks and decision trees to determine if the generic meals identified could predict whether participants belonged to the first or fifth quintile for Healthy Eating Index score; however, only the findings from the decision tree approach were described at a meal level (**Table 3.2**). For example, generic main meals such as "meat/fish and chips", "pizza", or "chips and fruit/veg/salad" were reported as predictive of quintile 1, whereas generic main meals such as "rice/pasta and fruit/veg/salad", "potatoes, veg/meat and yogurt", "fruit/veg/salad", or "potatoes and veg/fish" were reported as predictive of quintile 5. Different combinations of these meals over time, however, were not reported [27].

Of the three studies that did examine different combinations of meals within a given time (e.g. day), two were carried out in the same Irish cohort using PCA [9] and LCA [26], while one was carried out in a Japanese cohort using PCA [47]. The food groups used to define generic meals differed between the two studies that used PCA. The approach taken by Uzhova *et al.* [26] differed not only in their use of LCA, but also through the use of an additional aggregation step in defining generic meals, excluding the analysis of snacks and beverages, accounting for skipped meals, and distinguishing between weekday and weekend meal patterns.

Seven of the eleven meal patterns identified in the Japanese cohort were likely to include vegetables and/or rice as part of a breakfast meal [47], whereas none of the twelve [9] or seven [26] meal patterns identified in the Irish cohort contained breakfasts that were likely to consist of either vegetables or the rice/pasta/potatoes food group. All three studies identified patterns that included bread-based breakfasts and other patterns where vegetable consumption was unlikely. While Uzhova *et al.* [26] did not include beverages in their analysis, both Murakami *et al.* [47] and Woolhead *et al.* [9] identified meal patterns that were likely to include consumption of alcoholic beverages.

The meal patterns characterised by bread-based breakfasts with rice, vegetables, and meat/fish at both light meal and main meal identified by Murakami *et al.* [47], had comparable patterns identified by Woolhead *et al.* [9] and Uzhova *et al.* [26] but with a sandwich-based light meal in place of rice, vegetables, and meat/fish. Further similarities can be identified between the two studies in the same Irish cohort with common patterns including those based on cereal/toast breakfast, sandwich light meal and protein- and carbohydrate-based

main meal and others based on a higher likelihood of fruit consumption at breakfast, a light meal that does not contain bread, and likely to have lower overall meat intake. Uniquely, Woolhead *et al.* [9] identified a pattern characterised by consumption of confectionary at multiple meals. While consumption of confectionary was a feature in some of the patterns identified by Murakami *et al.* [47], it was not the defining feature in any of the patterns. As the confectionary food group was further aggregated into other generic meals in the approach taken by Uzhova *et al.* [26] it is not likely that such a pattern could have been identified in that study.

Uzhova *et al.* [26] were the only authors to distinguish between weekday and weekend meal patterns and to investigate the relationships between patterns and clinical variables. Four dominant weekday patterns and three dominant weekend patterns were identified by Uzhova *et al.* [26]. One meal pattern was found to be common to both weekdays and weekends which consisted of cooked breakfast, skipped light meal, and protein-carbohydrate-based main meal. However, those who consumed this pattern at the weekend tended to consume greater quantities of potatoes/potato dishes and have a greater overall energy intake than those who consumed the pattern primarily on weekdays. While those consuming certain meal patterns were found to have higher or lower intakes of certain nutrients, no general conclusions could be drawn regarding relationships between certain meal patterns and overall diet quality. Clinical variables were assessed after participants were grouped based on their dominant meal patterns for both weekends and weekdays. Significant differences were identified between those with the same weekday pattern (cereal and/or toast breakfast, skipped light meal or sandwich, and protein- and carbohydrate-based main meal), but differing weekend patterns. Those with a combination of the above weekday pattern and a weekend pattern consisting of cooked breakfast, skipped light meal, and protein carbohydrate main meal were more likely to have a higher diastolic blood pressure compared to those with a weekend pattern consisting of cereal and/or toast breakfast, sandwich light meal, and protein carbohydrate or just protein main meal; and a higher serum ferritin compared to those with a weekend pattern consisting of cereal and/or toast for breakfast, skipped light meal, and protein carbohydrate main meal. Despite this there was no clear relationship between the meal patterns and multiple clinical variables. Those consuming different meal patterns were not found to be different with regard to anthropometry, blood lipids, glucose, or CRP [26]. The analyses carried out by Uzhova *et al.* [26] adjusted for age, sex, social class, and energy intake.

3.4.3 Combined Patterns

One study was identified that investigated different types of meal patterns in a single population. Riou *et al.* [28] investigated combinations of both temporal and context patterns in 2,994 adults in France. Dietary intake data were collected by a questionnaire regarding the number of meals consumed and the time at which those meals were consumed. To assess temporal patterns, days were split into six time periods ranging from two to five hours in duration and participants were categorised has having consumed a meal or not during these periods, with the number of meals consumed also being counted. Context patterns in this study related to observations external to the meals. Specifically, the contextual variables examined by Riou *et al.* [28] included location (home, workplace, restaurant), who the meal was consumed with (alone, family members, colleagues or friends) and activities during the meal (television, radio, computer, reading, chatting). Patterns were identified based on these variables using the partitioning around the medoids clustering method (**Table 3.2**).

Five meal patterns were identified. The temporal aspects of these meal patterns hold similarities with the patterns described in the temporal patterns section above. Three of the patterns identified were likely to have three meals per day at times culturally associated with breakfast, lunch, and dinner. These patterns differed in their contextual aspects. One of the patterns represented those likely to have meals at work or a restaurant with colleagues or friends while chatting; another represented those likely to have meals at home, mostly alone, and therefore unlikely to chat but likely to watch television or listen to the radio during meals; the third of the three meals per day patterns was characterised by eating at home with family while chatting [28].

Two of the patterns identified were composed of those likely to consume one to two meals per day and not consume breakfast. One of these patterns was primarily characterised by consumption of meals at home with family while watching television, while the other pattern represented those who were likely to consume meals at work or in a restaurant with friends or colleagues while chatting [28].

The authors considered differences in food group intake across the identified patterns, adjusting for gender, age, education, occupation, income, underprivileged neighbourhoods, household type, and loneliness. When compared with the group characterised by consumption of three meals per day at home with family while chatting all other patterns had poor adherence to the 5-a-day consumption guideline of fruit and vegetables; this was particularly pronounced in those following the two patterns that typically did not consume breakfast, who were also less likely to adhere to 3 dairy products per day guideline [28].

3.5 Discussion

Meal-based methods of dietary assessment are a departure from the more familiar epidemiological methods that require detailed and accurate reporting of individual food intakes [11]. While meal-based methods may not offer the same degree of detail and accuracy, they can complement existing food-based dietary guidelines and may be superior for use in personalised nutrition delivered via internet and mobile technology due to the potential for reduced burden of data collection [11]. The use of advanced statistical techniques that inform these meal-based methods is still an emerging area, with only ten published studies identified [9, 18-20, 23-28].

Despite the methodological differences among studies, some common patterns prevailed in the temporal patterns of meal consumption: the 3 meals per day, skipped breakfast, and grazing patterns [18-20, 23, 24]. The patterns relating to the content of meals, however, were more heterogenous than the temporal patterns, with fewer consistent findings between studies. This may reflect differences arising from studies of populations with known differences in the types of foods consumed; that is, foods consumed as part of a typical Japanese diet differ from those consumed as part of a typical Western diet [48]. These differences were also observed in this review comparing the meal patterns among these two study populations. For example, breakfasts consumed by a Japanese cohort were likely to include rice and/or vegetables [25], whereas none of the breakfasts identified in an Irish cohort contained these food groups [9, 26].

Another source of differences between content meal patterns is the varying ways in which foods are grouped [9, 25-27]. In this regard the study of content meal patterns shares similarities in approach with the study of dietary patterns insofar as both condense the unique foods consumed by the study population into a pre-specified number of food groups [4]. All studies reviewed here used pre-existing food groups from previous research; no attempt has yet been made to create groups specifically for use in meal pattern analysis. It has been suggested by Newby *et al.* [49] that, in general, all studies need not use the same food groups, but instead the choice of groupings should be driven by the research question at hand. However, it is important to note that as the use of food groups introduces a degree of subjectivity and prior knowledge into what are otherwise data-driven approaches; the choice of groupings will likely impact on patterns identified [4, 49]. This in turn has likely given rise to some of the differences observed in this review between the content meal patterns from studies using different food groupings and highlights the need for careful consideration of the food groups used.

Given the range of statistical approaches (PCA, clustering, LCA, decision trees) applied to meal pattern analysis, comparisons between studies should be interpreted with caution. The extent to which the use of different approaches impact on the outcome is unclear as, to the authors' knowledge, no studies have compared the use of different statistical techniques to identify meal patterns in the same study population. Future research comparing approaches to meal pattern analysis could provide important methodological insights such as those reported for the more frequently researched area of dietary patterns that also employs techniques such as PCA and clustering [50-55].

Much of the research carried out in meal pattern analysis has been exploratory in nature identifying patterns of meal consumption that exist in the sample population [9, 18-20, 23-25, 28]. While exploratory research forms and essential part of the scientific process it is not without limitations [56, 57]. Results from these data-driven approaches may not be generalisable to samples from other populations [9]. While common dietary patterns have been identified in different populations [49], this has yet to be confirmed for meal patterns. Exploratory analyses identify groups within the sample population. Groups identified by these methods are typically assigned descriptive names by the researchers. These names introduce some subjectivity and should be interpreted with caution as there is no way of quantifying the variability within each group with regard to how well each member of the group is represented by the name assigned [40]. It is not possible to determine the likelihood of that these groups truly exist in the whole population rather than merely existing in the sample data [40]. However, it may be possible to determine whether the patterns are biologically meaningful if there are associations with health/disease status [49].

The meal pattern research reviewed here has primarily used unsupervised statistical techniques to identify groups of individuals with similar meal patterns. The research examining relationships between these meal patterns and diet quality or health outcomes remains sparse and warrants further investigation. Only six studies have examined these relationships with regard to temporal patterns [18, 23, 24, 42-44], one study with regard to content patterns [26], and one study with regard to combined temporal and context patterns [28]. In brief, those following the more traditional 3 meal per day pattern tended to have a higher diet quality than those following a skipped breakfast or grazing pattern [18, 23, 24, 42-44]. Unlike temporal meal patterns no individual content meal patterns have been identified as having notably strong relationships than other patterns with either diet quality or health outcomes [26]. The single study of combined temporal and context patterns by Riou *et al.* [28] identified those following a pattern characterised by 3 meals per day consumed with family while chatting as having greater adherence to the 5-a-day guideline for fruit and vegetable consumption.

Given their observational nature, the results from these studies may be impacted by confounding. Different variables were chosen in different studies as covariates to adjust for confounding. These choices can also impact on results and should be justified based on existing evidence or theoretical knowledge of their impact on confounding [58]. While the covariates used were listed in all the studies reviewed here [18, 23, 24, 26, 28, 42-44], not all provided a clear justification for their choice [26, 28, 42]. The decision tree approach taken by Hearty and Gibney [27] did not account for covariates. Future work in this area should consider approaches to account for covariates, for example, the use of adjusted residuals from a regression model as the input for the decision tree [59]. It should also be noted that these observational studies do not establish a cause-and-effect relationship but may provide data for causal inference and potentially inform future intervention studies [60].

This review examined studies in three main categories of meal patterns, namely temporal, content, and context patterns. These classifications were initially put forward by Mäkelä *et al.* [17] in the context of social and cultural aspects of meals using the terms eating patterns, meal format, and social organisation of eating. They were further adapted to the nutrition context by Leech *et al.* [10] who used the terms patterning, format, and context. While no consensus yet exists regarding the terminology, the current three-category approach accounts for the fact that people do not perceive dietary intake purely as a collection of nutrients, foods, or indeed meals [61], by capturing information regarding timing, social, and behavioural aspects of eating occasions [10]. Despite this, other aspects of meal patterns have not yet been examined using the statistical approaches reviewed here. For example, no studies were identified that examined sensory, psychological, or physical aspects such as emotions, satisfaction, appetite, fatigue etc. alongside those other aspects of meal patterns mentioned above. Furthermore, only limited aspects of temporal meal patterns have yet been examined. The research to date has primarily focused on the variation in meal intakes across a 24-hour period. Only three studies examined the variation between weekdays and weekends with the same temporal patterns being identified on both weekdays and weekends [18, 24]. With regard to content meal patterns, however, participants were found to adhere to different patterns on weekdays compared to the weekend [26]. No seasonal differences were identified in temporal meal patterns by Chau *et al.* [18], but this has not been examined with regard to content or context meal patterns. Only one study was identified that traced meal patterns across a number of years; Wilson *et al.* [24] found that the same temporal patterns existed after five years in a cohort of Australian adults and that participants were likely to fall into the same meal pattern category at follow up.

Expanding meal patterns to include these aspects would increase the complexity and require a multidisciplinary approach [62]; however, this may give rise to further useful insights

about meal patterns. Furthermore, the mobile technology exists to allow for the inclusion of such additional variables through ecological momentary assessment, i.e., the assessment of people's experiences of their environment in real time [21, 63]. Further development of the statistical approaches to meal pattern analysis will allow for the investigation of combinations of these variables and how they change over time [21]. In particular, supervised statistical approaches have the capacity to identify associations between meal patterns and overall diet quality or health in individuals for whom these data have been collected, and then used to predict diet quality or health outcomes for other individuals (without diet quality or health data) based on their meal patterns [27]. This in turn may have applications in personalised nutrition using internet and mobile technology [11].

3.6 Conclusion

A range of statistical techniques provide feasible solutions to interpreting complex dietary intake data and detecting insightful patterns of meal consumption relating to the timing, content, and context of meals. The observational studies reviewed here suggest that meal patterns consisting of three meals per day are associated with increased diet quality compared with the skipped breakfast or grazing meal patterns; however, further research is required to validate these findings. No clear associations with diet quality or health have been identified for meal patterns defined by the content of those meals or context in which they are consumed. To greater elucidate the role of meal patterns in diet quality and health, future research should aim to further develop the statistical approaches that are applied. Research is lacking on the simultaneous analysis of multiple meal pattern categories, how meal patterns vary over time, and the extent to which the grouping of foods and different types of statistical techniques impact on overall outcome. These advances will be important if meal pattern research is to be applied to internet- and mobile-based dietary assessment and feedback.

3.7 Tables

Table 3.1 Statistical approaches to meal pattern analysis in nutritional science.

Statistical Approach	Primary objective	Application to meal pattern analysis	References
Principal component analysis	• Variables (dimensions) that are correlated are grouped. • The total number of dimensions is reduced by only retaining a selection of grouped variables (components). • The retained components are those that are most important for explaining the variance in the data.	• The various possible combinations of foods, meals, timings of intake etc. could lead to millions of possible unique meal patterns. • Principal component analysis can reduce this large numbers of combinations to a smaller number of patterns that can be assessed for relationships with diet quality or health.	Englund-Ögge *et al.* [23]; Wilson *et al.* [24]; Woolhead *et al.* [9]; Murakami *et al.* [25]
Clustering	• Observations are grouped in a way that minimises within-cluster dissimilarity and maximises between-cluster dissimilarity. • Dissimilarity is typically measured using mathematical formulae for distance between points.	• Clustering can identify groups of individuals that eat meals at similar times over the course of a day and in a similar context.	Chau *et al.* [18]; Khanna *et al.* [19]; Riou *et al.* [28]
Latent class analysis	• Groups of observations are identified that have similar probabilities of belonging to the same categories in the variables of interest.	• Study participants can be grouped based on having high probabilities for eating during the same time periods of the day or consuming the same combinations of meals over a day.	Leech *et al.* [20]; Uzhova *et al.* [26]

Statistical Approach	Primary objective	Application to meal pattern analysis	References
Decision trees	• Observations are split into groups based on rules that are applied to the data. • Further rules are applied that continue to split the resulting groups until they cannot be further split or reach a stopping rule set by the researcher.	• Groups can be split based on the presence or absence of certain food combinations (meals) at various meal types while accounting for some outcome variable of interest. • This can allow for the use of meal intake for the prediction of an outcome variable such diet quality or a health biomarker.	Hearty and Gibney [27]

Table 3.2: Summarised findings of studies using advanced statistical approaches for meal pattern analysis.

Study	Statistical Approach	Data Collection Method	Meal Pattern Construct	Meal Patterns	Selected Diet Quality Findings	Selected Health Findings
4,508 adults in Taiwan, Chau *et al.* (18)	Clustering	24-hour Recall	Temporal	5 patterns based on dietary intake during 6 4-hour time periods.	Traditional timing pattern (3 meals/d) had the highest nutrient density. Delayed lunch with little or no morning intake pattern had the lowest nutrient density.	N/A
65,487 pregnant women in Norway, Englund-Ögge *et al.* (23)	PCA	Meal Frequency Questionnaire	Temporal	3 patterns based weekly frequency of 8 meal types e.g. breakfast, morning snack, lunch etc.	N/A	The lowest risk of preterm delivery relative to the 1st quartile was seen among those in 3rd (HR of 0.89; 95% CI: 0.79, 0.99) and 4th (HR of 0.88; 95% CI: 0.78, 0.99) quartiles for main meal pattern (3 meals/d) with $p = 0.046$.
1,304 adults in Australia, Wilson *et al.* (24)	PCA	Food Habits Questionnaire	Temporal	3 patterns based on dietary intake during 7 time periods.	N/A	No relationship between meal patterns and mood disorders at baseline. After 5 years, there was higher prevalence in those with increased (PR of 1.85;

Study	Statistical Approach	Data Collection Method	Meal Pattern Construct	Meal Patterns	Selected Diet Quality Findings	Selected Health Findings
						95% CI: 1.11 to 3.09) or consistently high (PR of 2.04; 95% CI: 1.20 to 3.28) adherence to the late pattern (lower intakes in morning and higher at night) relative to those with low adherence with $p <$ 0.001.
5,242 adults in Australia, Leech *et al.* (20)	LCA	24-hour Recall	Temporal	3 patterns based on dietary intakes in each hour of the day.	N/A	N/A
4,544 adults in Australia, Leech *et al.* (44)	As described by Leech *et al.* (20) above				Grazing pattern (frequent eating starting and continuing later in the day) had the lowest diet quality score on the Dietary Guidelines Index.	No association between patterns and BMI, BMI category, or waist circumference in the model adjusting for the highest number of covariates.
4,482 adults in Australia, Leech *et al.* (43)	As described by Leech *et al.* (20) above				N/A	Among women, the later lunch pattern (later lunch and evening meal than conventional) compared to

Study	Statistical Approach	Data Collection Method	Meal Pattern Construct	Meal Patterns	Selected Diet Quality Findings	Selected Health Findings
						conventional pattern (3 meals/d at conventional times) was positively associated with systolic and diastolic blood pressure and hypertension prevalence when adjusting for the highest number of covariates.
7,565 adults in the United States of America, Khanna *et al.* [19]	Clustering	24-hour Recall	Temporal	4 patterns based on dietary intake at each hour of the day.	N/A	N/A
9, 326 adults in the United Staes of America, Eicher-Miller *et al.* [42]	As described by Khanna *et al.* [19] above.				Cluster 1 (evenly spaced 3 meals/d with similar energy content) had highest healthy eating index score. Cluster 4 (5 meals/d, frequent intake at midday and midnight) had lowest.	Greater proportion of those in cluster 1 had normal BMI and lower proportion of those in cluster 1 had overweight/obese BMI than those in other clusters.

Study	Statistical Approach	Data Collection Method	Meal Pattern Construct	Meal Patterns	Selected Diet Quality Findings	Selected Health Findings
1,500 adults in Ireland, Woolhead *et al.* [9]	PCA	Food Diary	Content	12 patterns based on combinations of foods consumed at each meal.	N/A	N/A
242 adults in Japan, Murakami *et al.* [25]	PCA	Food Diary	Content	11 patterns based on combinations of foods consumed at each meal.	N/A	N/A
1,500 adults in Ireland, Uzhova *et al.* [26]	LCA	Food Diary	Content	4 weekday and 3 weekend patterns based on combinations of foods consumed at each meal.	Differences between different patterns in nutrient intake were reported, but no clear differences in overall diet quality.	Those with meal pattern "cooked breakfast, skipped light meal and protein carbohydrate main meal" had higher diastolic blood pressure ($p < 0.05$) compared to "cereal and/or toast at breakfast, sandwich for light meal, and protein carbohydrate or just protein main meal", and higher serum ferritin compared to "cereal and/or toast for breakfast, skipped light meal, and protein carbohydrate main meal"

Study	Statistical Approach	Data Collection Method	Meal Pattern Construct	Meal Patterns	Selected Diet Quality Findings	Selected Health Findings
						(OR of 3.14; 95% CI: 1.63 to 6.03).
1,379 adults in Ireland and the United Kingdom, Hearty and Gibney [27]	Decision Tree	Food Diary	Content	Daily patterns not given. 10 food combinations identified that were likely to predict whether an individual was in the first or fifth quintile for the Healthy Eating Index.	Meals likely to predict a lower Healthy Eating Index score were "bread and confectionary/snack", "breakfast cereal", "meat/fish products and chip", "pizza", and "chips and fruit/veg/salad". Those more likely to predict a higher score were combinations of breakfast cereal, fruit juice, and bread, "Rice/pasta dish and fruit/veg/salad", "potatoes, veg/meat and yogurt", "fruit/veg/salad", and "potatoes and veg/fish".	N/A

Study	Statistical Approach	Data Collection Method	Meal Pattern Construct	Meal Patterns	Selected Diet Quality Findings	Selected Health Findings
2,994, adults in France, Riou *et al.* (28)	Clustering	Questionnaire	Temporal and Context	5 patterns based on number and location of meals, and activities and others present during meals.	Those following the type 3 pattern (3 meals/d eaten at home with family) were most likely to consume fruit and vegetables daily.	N/A

BMI, body mass index; LCA, latent class analysis; N/A, not applicable; OR, odds ratio; PCA, principal component analysis; PR, prevalence ratio.

3.8 References

1. Afshin A, Sur PJ, Fay KA*, et al.* (2019) Health Effects of Dietary Risks in 195 Countries, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *The Lancet*, **393**(10184), 1958-1972.

2. Welch A (2019) Dietary Assessment. In *Manual of Dietetic Practice*, 6th ed., pp. 64-69 [J Gandy, editor]. Hoboken, NJ, USA: John Wiley & Sons, Inc.

3. Hu F (2002) Dietary Pattern Analysis: A New Direction in Nutritional Epidemiology. *Curr Opin Lipidol*, **13**(1), 3-9.

4. Michels KB and Schulze MB (2005) Can Dietary Patterns Help Us Detect Diet-Disease Associations? *Nutr Res Rev*, **18**(2), 241-248.

5. Edefonti V, De Vito R, Dalmartello M*, et al.* (2020) Reproducibility and Validity of a Posteriori Dietary Patterns: A Systematic Review. *Adv Nutr*, **11**(2), 293-326.

6. Jannasch F, Riordan F, Andersen LF*, et al.* (2018) Exploratory Dietary Patterns: A Systematic Review of Methods Applied in Pan-European Studies and of Validation Studies. *Br J Nutr*, **120**(6), 601-611.

7. Tapsell LC, Neale EP, Satija A*, et al.* (2016) Foods, Nutrients, and Dietary Patterns: Interconnections and Implications for Dietary Guidelines. *Adv Nutr*, **7**(3), 445-454.

8. Herforth A, Arimond M, Álvarez-Sánchez C*, et al.* (2019) A Global Review of Food-Based Dietary Guidelines. *Adv Nutr*, **10**(4), 590-605.

9. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

10. Leech RM, Worsley A, Timperio A*, et al.* (2015) Understanding Meal Patterns: Definitions, Methodology and Impact on Nutrient Intake and Diet Quality. *Nutr Res Rev*, **28**(1), 1-21.

11. Gibney MJ and Walsh MC (2013) The Future Direction of Personalised Nutrition: My Diet, My Phenotype, My Genes. *Proc Nutr Soc*, **72**(2), 219-225.

12. Popkin BM and Duffey KJ (2010) Does Hunger and Satiety Drive Eating Anymore? Increasing Eating Occasions and Decreasing Time between Eating Occasions in the United States. *Am J Clin Nutr*, **91**(5), 1342-1347.

13. Mekary RA, Giovannucci E, Willett WC*, et al.* (2012) Eating Patterns and Type 2 Diabetes Risk in Men: Breakfast Omission, Eating Frequency, and Snacking. *Am J Clin Nutr*, **95**(5), 1182-1189.

14. Laska MN, Graham D, Moe SG*, et al.* (2011) Situational Characteristics of Young Adults' Eating Occasions: A Real-Time Data Collection Using Personal Digital Assistants. *Public Health Nutr*, **14**(3), 472-479.

15. Mak TN, Prynne CJ, Cole D*, et al.* (2012) Assessing Eating Context and Fruit and Vegetable Consumption in Children: New Methods Using Food Diaries in the UK National Diet and Nutrition Survey Rolling Programme. *Int J Behav Nutr Phys Act*, **9**(126).

16. Lennernäs M and Andersson I (1999) Food-Based Classification of Eating Episodes. *Appetite*, **32**(1), 53-65.

17. Mäkelä J, Kjærnes U, Pipping Ekström M*, et al.* (1999) Nordic Meals: Methodological Notes on a Comparative Survey. *Appetite*, **32**(1), 73 - 79.

18. Chau CA, Pan WH, and Chen HJ (2017) Employment Status and Temporal Patterns of Energy Intake: Nutrition and Health Survey in Taiwan, 2005–2008. *Public Health Nutr*, **20**(18), 3295-3303.

19. Khanna N, Eicher-Miller HA, Boushey CJ*, et al.* (2011) Temporal Dietary Patterns Using Kernel K-Means Clustering. *IEEE International Symposium on Multimedia*, **2011**, Dana Point, CA, USA, 375-380.

20. Leech RM, Worsley A, Timperio A*, et al.* (2017) Temporal Eating Patterns: A Latent Class Analysis Approach. *Int J of Behav Nutr Phys Act* **14**(1), 1-9.

21. Pendergast FJ, Leech RM, and McNaughton SA (2017) Novel Online or Mobile Methods to Assess Eating Patterns. *Curr Nutr Rep*, **6**(3), 212-227.

22. James G, Witten D, Hastie T*, et al.* (2013) An Introduction to Statistical Learning with Applications in R. New York, USA: Springer.

23. Englund-Ögge L, Birgisdottir BE, Sengpiel V*, et al.* (2017) Meal Frequency Patterns and Glycemic Properties of Maternal Diet in Relation to Preterm Delivery: Results from a Large Prospective Cohort Study. *PLoS One*, **12**(3), 1-19.

24. Wilson JE, Blizzard L, Gall SL*, et al.* (2019) An Eating Pattern Characterised by Skipped or Delayed Breakfast Is Associated with Mood Disorders among an Australian Adult Cohort. *Psychol Med*, **50**(16), 2711-2721.

25. Murakami K, Livingstone MBE, Sasaki S*, et al.* (2018) Applying a Meal Coding System to 16-D Weighed Dietary Record Data in the Japanese Context: Towards the Development of Simple Meal-Based Dietary Assessment Tools. *J Nutr Sci*, **7**(e29), 1-10.

26. Uzhova I, Woolhead C, Timon CM*, et al.* (2018) Generic Meal Patterns Identified by Latent Class Analysis: Insights from NANS (National Adult Nutrition Survey). *Nutrients*, **10**(3), 310.

27. Hearty ÁP and Gibney MJ (2008) Analysis of Meal Patterns with the Use of Supervised Data Mining Techniques—Artificial Neural Networks and Decision Trees. *Am J Clin Nutr*, **88**(6), 1632-1642.

28. Riou J, Lefèvre T, Parizot I*, et al.* (2015) Is There Still a French Eating Model? A Taxonomy of Eating Behaviors in Adults Living in the Paris Metropolitan Area in 2010. *PLoS One*, **10**(3), 1-18.

29. Shim JS, Oh K, and Kim HC (2014) Dietary Assessment Methods in Epidemiologic Studies. *Epidemiol Health*, **36**, e2014009.

30. Buttriss J, Welch A, Kearney JM*, et al.* (2018) Public Health Nutrition. 2nd ed*.* Chichester, W. Sussex, England: John Wiley & Sons, Inc.

31. Troyanskaya O, Cantor M, Sherlock G*, et al.* (2001) Missing Value Estimation Methods for DNA Microarrays. *Bioinformatics*, **17**(6), 520-525.

32. Idri A, Abnane I, and Abran A (2016) Missing Data Techniques in Analogy-Based Software Development Effort Estimation. *Journal of Systems and Software*, **117**, 595-611.

33. Cismondi F, Fialho AS, Vieira SM*, et al.* (2013) Missing Data in Medical Databases: Impute, Delete or Classify? *Artif Intell Med*, **58**(1), 63-72.

34. Abu-Jamous B, Fa R, and Nandi AK (2015) Integrative Cluster Analysis in Bioinformatics*.* New York, NY, USA: John Wiley & Sons.

35. Monti S, Tamayo P, Mesirov J*, et al.* (2003) Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Machine Learning*, **52**(1-2), 91-118.

36. Milligan GW and Cooper MC (1985) An Examination of Procedures for Determining the Number of Clusters in a Data Set. *Psychometrika*, **50**(2), 159-179.

37. Charrad M, Ghazzali N, Boiteau V*, et al.* (2014) Nbclust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, **61**(6), 1-36.

38. Collins LM and Lanza ST (2010) Latent Class and Latent Transition Analysis with Applications in the Social, Behavioral, and Health Sciences, *Wiley Series in Probability and Statistics.* New Jersey, USA: John Wiley & Sons.

39. Lanza ST and Rhoades BL (2013) Latent Class Analysis: An Alternative Perspective on Subgroup Analysis in Prevention and Treatment. *Prev Sci*, **14**(2), 157-168.

40. Hastie T, Tibshirani R, and Friedman J (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, *Springer Series in Statistics.* New York, USA: Springer Science+Business Media.

41. Quinlan JR (1993) C4.5: Programs for Machine Learning, *The Morgan Kaufmann Series in Machine Learning.* San Mateo, CA, USA: Morgan Kaufmann Publishers.

42. Eicher-Miller HA, Khanna N, Boushey CJ*, et al.* (2016) Temporal Dietary Patterns Derived among the Adult Participants of NHANES 1999-2004 Are Associated with Diet Quality. *J Acad Nutr Diet*, **116**(2), 283-291.

43. Leech RM, Timperio A, Worsley A*, et al.* (2019) Eating Patterns of Australian Adults: Associations with Blood Pressure and Hypertension Prevalence. *Eur J Nutr*, **58**(5), 1899-1909.

44. Leech RM, Timperio A, Livingstone KM*, et al.* (2017) Temporal Eating Patterns: Associations with Nutrient Intakes, Diet Quality, and Measures of Adiposity. *Am J Clin Nutr*, **106**(4), 1121-1130.

45. White R, Harwin WS, Holderbaum W*, et al.* (2016) Investigating Eating Behaviours Using Topic Models. *IEEE 14th International Conference on Machine Learning and Applications, 2015*, Miami, FL, USA, 265-270.

46. Schwedhelm C, Knüppel S, Schwingshackl L*, et al.* (2018) Meal and Habitual Dietary Networks Identified through Semiparametric Gaussian Copula Graphical Models in a German Adult Population. *PLoS One*, **13**(8), 1-16.

47. Murakami K, Livingstone MBE, and Sasaki S (2019) Meal-Specific Dietary Patterns and Their Contribution to Overall Dietary Patterns in the Japanese Context: Findings from the 2012 National Health and Nutrition Survey, Japan. *Nutrition*, **59**, 108-115.

48. Murakami K, Livingstone MBE, Fujiwara A*, et al.* (2020) Application of the Healthy Eating Index-2015 and the Nutrient-Rich Food Index 9.3 for Assessing Overall Diet Quality in the Japanese Context: Different Nutritional Concerns from the US. *PLoS One*, **15**(1), e0228318.

49. Newby PK, Sc D, Tucker KL*, et al.* (2004) Empirically Derived Eating Patterns Using Factor or Cluster Analysis: A Review. *Nutr Rev*, **62**(5), 177-203.

50. Hearty AP and Gibney MJ (2009) Comparison of Cluster and Principal Component Analysis Techniques to Derive Dietary Patterns in Irish Adults. *Br J Nutr*, **101**(4), 598-608.

51. Hearty AP and Gibney MJ (2013) Dietary Patterns in Irish Adolescents: A Comparison of Cluster and Principal Component Analyses. *Public Health Nutr*, **16**(5), 848-857.

52. Kant AK, Graubard BI, and Schatzkin A (2004) Dietary Patterns Predict Mortality in a National Cohort: The National Health Interview Surveys, 1987 and 1992. *J Nutr*, **134**(7), 1793-1799.

53. Newby PK, Muller D, and Tucker KL (2004) Associations of Empirically Derived Eating Patterns with Plasma Lipid Biomarkers: A Comparison of Factor and Cluster Analysis Methods. *Am J Clin Nutr*, **80**(3), 759 - 767.

54. Bamia C, Orfanos P, Ferrari P*, et al.* (2005) Dietary Patterns among Older Europeans: The EPIC-Elderly Study. *Br J Nutr*, **94**(1), 100-113.

55. Crozier SR, Robinson SM, Borland SE*, et al.* (2006) Dietary Patterns in the Southampton Women's Survey. *Eur J Clin Nutr*, **60**(12), 1391-1399.

56. Tukey JW (1980) We Need Both Exploratory and Confirmatory. *The American Statistician*, **34**(1), 23-25.

57. Jebb AT, Parrigon S, and Woo SE (2017) Exploratory Data Analysis as a Foundation of Inductive Research. *Human Resource Management Review*, **27**(2), 265-276.

58. Zeraatkar D, Cheung K, Milio K*, et al.* (2019) Methods for the Selection of Covariates in Nutritional Epidemiology Studies: A Meta-Epidemiological Review. *Current Developments in Nutrition*, **3**(10), nzz104.

59. Venkatasubramaniam A, Wolfson J, Mitchell N*, et al.* (2017) Decision Trees in Epidemiological Research. *Emerg Themes Epidemiol*, **14**, 11.

60. Potischman N and Weed DL (1999) Causal Criteria in Nutritional Epidemiology. *Am J Clin Nutr*, **69**(6), 1309S-1314S.

61. Bisogni CA, Falk LW, Madore E*, et al.* (2007) Dimensions of Everyday Eating and Drinking Episodes. *Appetite*, **48**(2), 218-231.

62. Meiselman HL (2009) Dimensions of the Meal: A Summary. In *Meals in Science and Practice: Interdisciplinary Research and Business Applications* [HL Meiselman, editor]. Cambridge, England: Woodhead Publishing Limited.

63. Stone AA and Shiffman S (2002) Capturing Momentary, Self-Report Data: A Proposal for Reporting Guidelines. *Ann Behav Med*, **24**(3), 236-243.

CHAPTER 4

A Clustering Approach to Meal-Based Analysis of Dietary Intake Applied to Population and Individual Data

A version of this chapter has been published as:

O'Hara C, O'Sullivan A, and Gibney ER (2022) A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data. *J Nutr*, **152**(10), 2297-2308

4.1 Abstract

Examination of meal intakes can elucidate the role of individual meals or meal patterns in health not evident by only examining nutrient and food intakes. To date, meal-based research has been limited to a focus on population rather than individual level intakes, without considering portion size or nutrient content when characterising meals. The objective of this study was to characterise meals commonly consumed, incorporating portion size and nutritional content, and to determine the accuracy of nutrient intake estimates using these meals at both population and individual levels. The 2008–2010 Irish National Adult Nutrition Survey (NANS) data were used. 1500 participants, with an age (mean ± SD) of 44.5 ± 17.0 and BMI of 27.1 ± 5.0, recorded their intake using a 4-day weighed food diary. Food groups were identified using *k*-means clustering. Partitioning around the medoids clustering was used to categorise similar meals into groups (generic meals) based on their Nutrient Rich Foods Index (NRF9.3) score and the food groups that they contained. The nutrient content for each generic meal was defined as the mean content of the grouped meals. Seven standard portion sizes were defined for each generic meal. Mean daily nutrient intakes were estimated using the original and the generic data. The 27,336 meals consumed were aggregated to 63 generic meals. Effect sizes from the comparisons of mean daily nutrient intakes (from the original vs generic meals) were negligible or small, with *P* values ranging from <0.001 to 0.941. When participants were classified according to nutrient-based guidelines (high, adequate, or low), the proportion of individuals who were classified into the same category ranged from 55.3% to 91.5%. In conclusion, a generic meal-based method can estimate nutrient intakes based on meal rather than food intake at the sample population and individual levels. Future work will focus on incorporating this concept to a meal-based dietary intake assessment tool.

4.2 Introduction

As nutrients and foods are consumed as part of meals, meal-based intake information may complement existing dietary guidelines with regard to understanding and meal planning by consumers [1, 2]. Knowledge derived from meal-based research would allow dietary guidelines to provide advice on the basis of different meal types and their combinations, the timing of meals, and the types of foods consumed together as part of those meals, rather than focusing only on the types of foods that should be consumed over the course of a day [2, 3]. This advice may be more intuitively understood as people tend to conceptualise their dietary intakes as meals rather than daily intakes of foods [4, 5]. A meal-based approach may also be superior in technology-based intake assessments and the provision of personalised nutrition recommendations, as meal intakes can be recorded in a way that is less burdensome than recording of individual food intakes [6].

Previous meal-based research has used national dietary survey data to identify a small list of generic, or commonly consumed, meals in a given population that are representative of the much larger list of actual meals consumed by that population [7-11]. These generic meals typify the actual meals that are consumed and are defined by their nutritional content and the food groups that they contain [7-11]. Research on generic meals to date has focused on their potential for estimating dietary intakes at a population level [7-9] or characterising the types of foods that individuals commonly combine to make meals [7-11]. No studies have assessed whether generic meals can be used to classify individuals' dietary intakes into categories according to nutrient-based dietary guidelines. This method of categorising individual intakes has been used in other forms of technology-based personalised nutrition, whereby the dietary advice provided to individuals is determined by whether they consume low, moderate, or high quantities of nutrients according to nutrient-based guidelines [12]. Furthermore, the use of technology in dietary intake assessment is known to reduce respondent burden [13], and meal-based methods have been proposed as a lower burden way in which dietary intakes could be assessed with regard to nutrient-based guidelines and dietary feedback provided using internet and mobile technology [6]. Determining the ability of meal-based intakes to correctly estimate nutrient intakes at an individual level is important for the development of such meal-based intake assessment and feedback tools.

When comparing intakes estimated using a meal-based method with those from food-based methods (such as food diaries), good agreement has been obtained using generic meals to determine mean daily intakes at a sample population level, with median correlation coefficients ranging from 0.44 [7] to 0.61 [9] for correlations between estimated nutrient intakes derived from the food data and the generic data. However, previous methods of

identifying generic meals to date have failed to incorporate portions sizes and also have limited emphasis on nutritional content when identifying generic meals. This may limit their ability to capture the variability of nutrient intakes between individuals and may not be sufficient to identify generic meals and meal patterns that are relevant for the promotion of health [7, 8, 14]. The approaches used to date have relied on frequent itemsets data mining methods [7-9], latent variable mixture modelling [10], and principal components analysis [11]. While these approaches are effective in identifying common combinations of food groups within meals, they cannot distinguish between those meals that contain similar or identical food groups but are considerably different in their nutritional contents, for example, the foods in the meat/fish group can vary considerably with regard to their energy content and fat composition [7, 9-11, 15].

The objectives of this study are to determine generic meals using existing national dietary survey data incorporating both portion size and the nutritional content of meals, and to assess the degree of agreement between nutrient intakes derived using the generic meals and the original meals at both the sample population and individual level.

4.3 Methods

4.3.1 Study Sample

The analysis was carried out on previously collected data from the National Adult Nutrition Survey (NANS) in Ireland [16]. NANS recruited participants aged between 18 and 90 years who were free-living and not pregnant or breast feeding. Data were collected between October 2008 and April 2010. The sample was representative of Irish adults with respect to age, gender, social class, and geographical location. Each participant completed a four-day weighed food diary. Reported food intake data were converted to nutrient data using WISP© (Tinuviel Software, Anglesey, UK). In this original dataset each participant has nutrient intake values for each food consumed during the recording period. Ethical approval was provided by University College Cork Clinical Research Ethics Committee of the Cork Teaching Hospitals. Further details on the study sample are available elsewhere [16].

4.3.2 Food Groups

Clustering is a data analysis technique that groups observations (in this case foods) into groups, or clusters, in such a way that the features (or variables; in this case the nutrients of interest within the foods) of the observations within a given cluster are more similar to each other than to the observations (foods) in other clusters [17]. Within the current analysis,

foods reported in the dietary survey were firstly grouped into one of 6 groups within the Irish Food pyramid (vegetables, salad, and fruit; cereal, bread, pasta, potato, and rice; milk, yogurt, and cheese; meat, poultry, fish, eggs, beans, and nuts; fats, spreads, and oils; and foods high in fat, sugar and/or salt) [18], and then further grouped using *k*-means clustering (**Figure 4.1**). A seventh group (other) was used for the remaining foods that did not belong to any of the aforementioned groups. This approach thus identified subgroups of foods within each of the existing food pyramid food groups, with similar compositions of the key nutrients of concern, outlined below. The Irish food pyramid was used as the starting point to allow potential future applications of this work (dietary assessment and dietary advice) to be easily understood in the context of existing dietary guidelines. The purpose of using these newly derived subgroups rather than existing food groupings, within the NANS dataset, was to allow the food groups to be created in a data-driven manner that accounts for nutrients of public health importance and demonstrates a pathway for any future researcher to adopt to future datasets. Specifically, the features (or input variables) chosen for the clustering were the 12 nutrients used in the previously validated nutrition quality index, the Nutrient Rich Foods Index [19]: protein, fibre, vitamin A, vitamin C, vitamin E, calcium, iron, magnesium, potassium, saturated fat, added sugar, and sodium; that is, foods were clustered based on their similarity in relation to those nutrients. Food supplements and energy-free foods were not included in this analysis, as they had not been included in the original validation studies of the Nutrient Rich Foods Index [19]. Variables were in units per 100kcal and were z-standardised prior to clustering. *K*-means clustering requires that the number of clusters is chosen prior to clustering. Within this study the number of clusters was chosen by applying 24 different indices used to determine the number of clusters to the data and choosing the number of clusters that was most frequently proposed among the indices [20]. For two of the groups (cereal, bread, potato, pasta and rice and milk, yogurt, and cheese), the clustering solution arising from the most frequently proposed number of clusters led to some clusters containing only 2–3 foods. In these instances, the next most frequently chosen value for cluster number was used. The range of possible values assessed for cluster number was from 2 to 10 inclusive. Details on the indices used are provided by Charrad *et al.* [20]. The values of the indices obtained for the vegetables, salad, and fruit group of the food pyramid have been provided as an example in Supplementary Table 4.1. Foods that did not have a specific group in the food pyramid were not clustered, but considered as individual food groups: alcoholic beverages, non-alcoholic beverages, and miscellaneous foods such as sauces, dips, and dressings. The fats, spreads, and oils group was not further split using clustering given the small number of foods in this group (n = 36). The original groups in the food pyramid and the number of clusters identified within each group are presented in **Table 4.1**. In total there were 12 food groups identified through

clustering and the 3 remaining groups above that do not appear in the food pyramid (**Table 4.1**). These 15 food groups were then used to derive generic meals as described below.

4.3.3 Generic Meals

Following the *k*-means clustering to determine the 15 food groups, individual foods listed within each meal, reported by participants, were replaced with the food group to which they belonged. Individual meals were then defined by the individual food groups that made up each meal, meaning that for each food reported in a given meal, the food group to which it was assigned was used to represent that food, instead of the individual item. Following this substitution of food groups in place of foods in the meal, an NRF9.3 [19] score was calculated for each individual meal as an indicator of nutritional quality. The score, developed by Fulgoni *et al.* [19], was calculated based on 9 nutrients to encourage (protein, dietary fibre, calcium, iron, potassium, magnesium, and vitamins A, C, and E) and 3 nutrients to limit (saturated fat, added sugar, and sodium). For each of these nutrients the quantity of the nutrient per 100kcal present in each meal as a percentage of European reference intakes (RIs) was calculated. RIs were derived from Regulation (EU) No 1169/2011 [21] where available. Those that were not available in that regulation (added sugar and dietary fibre) were instead derived from the European Food Safety Authority (EFSA) [22]. The percentage of RIs were limited to 100% to avoid over-valuing meals. The percentage RI scores for the nutrients to encourage were summed, as were the scores for the nutrients to limit. The nutrients to limit score was subtracted from the nutrients to encourage score to give an overall NRF9.3 score for each meal.

Within the NANS study participants self-selected the meal type to which each recorded intake belonged from the following list: breakfast, light meal (lunch), light meal (evening meal), main meal (lunch), main meal (evening meal), morning snack, afternoon snack, evening snack, night snack, alcoholic beverage, and non-alcoholic beverage. These meal types were condensed into the following 5 types: breakfasts, light meals, main meals, snacks, and beverages. Meals were also categorised by whether they were consumed on a weekday or a weekend day, giving rise to meals being divided into 10 categories (e.g., weekend breakfasts, weekday breakfasts etc.). Within each meal category, the individual meals (defined by the food groups that they contain) were then grouped/clustered using partitioning around the medoids (PAM) clustering allowing for clustering based on both numerical and categorical variables [23, 24]. The variables used to cluster were the NRF9.3 score for each meal and 15 binary variables indicating for each food group whether it was present or absent in the meal. PAM clustering requires that the number of clusters is chosen prior to clustering. Similar to the first clustering step, the number of clusters was chosen by

applying 24 different cluster number indices to the data and choosing the number of clusters that was most frequently proposed among the indices [20]. Values obtained from the various indices for the clustering of weekend breakfast meals have been provided in Supplementary Table 4.2 as an example. The range of possible values assessed for cluster number was from 4 to 15 inclusive. This clustering process gave rise to a total of 63 clusters, i.e., generic meals.

The nutrient content of a given generic meal was then calculated as the mean nutrient content per 100g of the individual meals that make up that generic meal. Before calculating mean nutrient content and portion size (described below), meals considered to be outliers (based on energy and/or micronutrient content) within each generic meal were removed in a two-step process. Firstly, meals that contained greater than 1.5 times the interquartile range for energy were removed. Secondly, meals that contained greater than 10 times the mean for any micronutrient were removed.

The weight in grams for the individual meals in each generic meal, were used to determine generic portion sizes. Each generic meal was assigned 7 generic portion sizes. These portion sizes were determined for a given generic meal by ordering each of the individual meals by weight and dividing the meals into sevenths based on septile values. The median weights of each seventh were assigned as the generic portion sizes for that meal (**Figure 4.1**).

4.3.4 Generic Intakes

To create the generic dataset, the nutrient intake values of meals consumed in the original dataset were replaced with the appropriate generic meal composition based on the portion size consumed (**Figure 4.1**). Mean daily intakes of the nutrients of interest were calculated using both the original and generic datasets.

4.3.5 Statistical Analysis

All analysis was carried out using R version 4.0.3 [25] in the RStudio integrated development environment (version 1.3.1093) [26]. Mean nutrient intakes arising from the original and generic datasets were compared using a paired *t*-test. To account for the multiple testing of energy and 30 nutrients, a Bonferroni-adjusted α of 0.05/31 was considered for statistical significance, i.e., *P* values of < 0.002 were considered statistically significant. The nutrients compared between the two datasets included fat, saturated fat, monounsaturated fat, polyunsaturated fat, protein, carbohydrate, total sugars, added sugars (all of which were compared using both gram amounts and in terms of % total energy intake (TEI)), dietary

fibre (g), calcium (mg), iron (mg), potassium (mg), phosphorous (mg), sodium (mg), total vitamin A (µg), retinol (µg), carotene (µg), vitamins C (mg), D (µg), E (mg), B12 (µg), and folate (µg).

The correlation of nutrient intakes between the two datasets was examined using Spearman correlation coefficients. Nutrient intakes in both datasets were divided into quartiles. This allowed for the calculation of the proportion of individuals who remained in the same quartile in both datasets (exact agreement), the proportion who were classified in the same or adjacent quartiles (exact agreement + adjacent), the proportion who were classified two quartiles apart (disagreement), and the proportion who were classified three quartiles apart (extreme disagreement). Bland-Altman analysis was carried out whereby the mean difference between the two datasets and the limits of agreement (mean difference ± 1.96SD) for each nutrient were calculated. It is expected that ≥ 95% of observations will fall within the limits of agreement for comparable methods [27].

Finally, participants were classified, separately for both datasets, according to nutrient-based dietary guidelines [28-30]. For example, they were classified whether their nutrient intakes were low, adequate, or high according to those guidelines. The nutrients assessed included protein, carbohydrate, fat, monounsaturated fat, polyunsaturated fat, saturated fat, salt, dietary fibre, calcium, iron, folate, thiamine, riboflavin, vitamins A, B12, and C. The proportion of individuals who were classified into the same category in both datasets was calculated for each nutrient.

4.4 Results

Of the 1500 participants who completed the NANS study 740 (49.3%) were male and 760 (50.7%) were female. The overall age was 44.5 ± 17y (mean ± SD). The overall BMI was 27.1 ± 5.0 kg/m^2 (**Table 4.2**). Implausible estimates of energy intake were deemed to be those giving rise to an EI:BMR ratio of less than 1.10 [31]. The results of the analysis were not impacted when participants with implausible estimates of energy intake were removed (data not shown). The results presented here, therefore, are those based on the data from all 1500 participants.

A total of 27336 individual meals were consumed by the 1500 participants during their 4-day recording period. Of these meals, 17848 were consumed on a weekday and 9488 were consumed during the weekend. Overall, the most consumed meal type was snacks (n = 7710), followed by main meals (n = 6025), breakfast (n = 5698), light meals (n = 4455), and beverages (n = 3448). These individual meals were aggregated to 63 generic meals. These

were comprised of 4 breakfasts, 8 light meals, 11 main meals, 4 snacks, 4 beverages during the weekdays, and 4 breakfasts, 6 light meals, 14 main meals, 4 snacks, and 4 beverages during the weekend (**Figure 4.1**). Generic meal portion sizes and food-based descriptions based on the most frequently occurring foods in each food group of each meal are presented in Supplementary Table 4.3. The nutritional content of each of the generic meals were determined from the mean nutritional content of the individual meals making up those generic meals excluding outlier meals. 3825 meals were excluded on the basis of being outliers with regard to energy content and 2279 were excluded on the basis of being outliers with regard to micronutrient content. Therefore, a total of 21232 meals were used to determine the nutrient content of the generic meals.

4.4.1 Comparison of Mean Intakes

When comparing the sample population mean intakes of the original versus the generic meal data (original dataset vs dataset with generic meal substituted in), the mean percentage difference between estimated mean daily nutrient intakes for the original and generic datasets was 5.6%. The percentage difference between the datasets for all of the macronutrients and energy was <5%, ranging from 0.0% for total sugars as % total energy intake (TEI) to 4.2% for added sugars as %TEI. The percentage differences for the micronutrients ranged from 3.0% for sodium (mg) to 25.3% for retinol (µg) (**Table 4.3**).

Despite the small percentage differences for most nutrients, statistically significant differences were observed in mean intakes of 20 of the 30 nutrients assessed and for energy (adj. $P < 0.002$). Nutrients that were not significantly different were fat in grams, saturated fat in grams, monounsaturated fat as %TEI, polyunsaturated fat as %TEI, protein as %TEI, carbohydrate in grams, total sugars in grams, total sugars as %TEI, added sugars in grams, added sugars as %TEI, and dietary fibre in grams (**Table 4.3**). All of the micronutrients assessed were found to differ significantly between the two datasets. However, of the 20 significant differences, 15 had a negligible effect size (Cohen's $d < 0.2$) and the remainder had a small effect size (Cohen's $d \geq 0.2$ and < 0.5). Those with a small effect size were potassium (Cohen's $d = 0.262$), phosphorous (Cohen's $d = 0.263$), vitamin D (Cohen's $d = 0.229$), folate (Cohen's $d = 0.255$), and vitamin B12 (Cohen's $d = 0.265$) (**Table 4.3**).

Spearman correlation coefficients between the original and generic datasets ranged from 0.23 for polyunsaturated fat (%TEI) to 0.75 for potassium (mg), with a mean correlation across all nutrients of 0.53 (**Table 4.4**). Bland-Altman analysis identified 13 nutrients for which ≥ 95% of participants fell within the limits of agreement. The proportion of individuals

that fell within the limits of agreement ranged from 93.7% for sodium (mg) to 98.4% for retinol (µg) (**Table 4.4**). The Bland-Altman plots for the macronutrients are shown in **Figure 4.2**.

4.4.2 Individual Intake Classification

The proportion of participants remaining in the same quartile in both the original and generic datasets ranged from 31% for polyunsaturated fat (%TEI) to 53% for potassium (mg). Extreme disagreement of 5% of participants or greater was observed for 6 nutrients: polyunsaturated fat as %TEI (8%), monounsaturated fat as %TEI (6%), total fat as %TEI, (6%), vitamin D (6%), saturated fat as %TEI (5%), and vitamin B12 (5%). The mean proportion of individuals in extreme disagreement across all nutrients was 3.3% (**Table 4.4**).

When participants were classified (high, adequate, or low) according to nutrient-based guidelines, the proportion of individuals who were classified into the same category ranged from 55.3% for polyunsaturated fat (%TEI) to 91.5% for both protein (g/kg body weight) and salt (g). The mean proportion of exact agreement across all nutrients was 79.8% (**Table 4.5**).

4.5 Discussion

This study examined a novel generic meal-based method of exploratory analysis of dietary intake data. Unlike previous work in this area [7-11], the research described here utilizes data-driven methods of defining food groups based on nutrients from the Nutrient Rich Foods Index [19], and incorporates a range of standard portion sizes, which aimed to improve accuracy of a meal-based intake method to estimate intakes at both a population and individual level, which had not previously been considered. Differences between mean nutrient intakes determined using this generic meal-based method and the standard food-based method were found to be small or negligible. This study also compared nutrient intakes at the individual level, for the first time, demonstrating that participants can be classified according to nutrient-based dietary guidelines (for example, high, adequate, or low intakes) using the generic meal-based method, suggesting a possible role for the use of generic meals in meal-based dietary intake assessment.

Food groupings vary from study to study and are often tailored to the specific research in which they are used [32]. Previous research examining generic meals has used pre-existing food groups as part of the process [7-11]. The current research aimed to develop a generic meals process to assess dietary intakes with regard to nutrients of public health importance,

specifically the 12 nutrients from the Nutrient Rich Foods Index [19]. As it was not feasible to group hundreds of foods manually based on their similarities for 12 different nutrients, clustering was applied instead. Whilst the use of nutrients in the derivation of generic meals appeared to improve the process, the use of the specific 12 nutrients in the generic meal-based approach may have impacted on the agreement between the two methods for intakes of those nutrients, above others. For example, the mean correlation between methods, for nutrients used as input for *k*-means clustering was 0.57, whereas the mean correlation for nutrients that were not included was 0.48. This may provide scope for the inclusion of a different set of nutrients or other food components depending on the research question at hand. Indeed, previous studies that have clustered foods have also incorporated prior knowledge of nutrition in the selection of variables. For example, Pennington and Fisher [33] clustered fruit and vegetables using food components known to be primarily provided by these foods; Burden *et al.* [34] clustered foods based on macronutrient content with the aim of minimising their differences between groups; and in a study concerning the antioxidant content of fruit and vegetables, Patras *et al.* [35] clustered foods based on their content of known antioxidant compounds.

In the present study, the choice of input variables for clustering deliberately focused on nutrients of known public health importance, introducing a degree of prior knowledge to a largely data-driven method. One of the limitations of this approach is that some of the variables included may be irrelevant or redundant for the purpose of clustering and thereby mask the underlying structures in the data [36]. The alternative is to exclude prior knowledge and use a purely data-driven approach to selecting input variables for clustering [36, 37] that would enable foods to be grouped based only on the food components that differentiate those foods. While these data-driven methods are valuable approaches to identify inherent patterns or groupings in the data, they may not perform as well in identifying groupings that are relevant to disease prevention or health promotion [14], which is of importance in this current study. While a purely data-driven approach to clustering was not used in this instance, such approaches warrant further investigation as groupings driven by the underlying data structures of food composition may help researchers develop food groups for food frequency questionnaires [33, 34], be useful in teaching about food composition, and aid health professionals to produce dietary advice about foods with similar compositions when the content of multiple nutrients are of interest [33].

The methods described in this study build on previously published studies that used the frequent itemset data mining method [7-9] to derive the generic meals used in their analysis. These studies did not include portion size in their generic meal framework and the methods used accounted for food group “descriptions” but not nutrient content when identifying

generic meals. The clustering method used in this current work derived generic meals based on nutrients of public health interest from the Nutrient Rich Foods Index [19] while also accounting for the food groups of which those generic meals were comprised. By identifying generic meals separately for weekdays and weekends, this method accounts for the differences that exist between weekday and weekend dietary intakes in relation to energy [38-40] and nutrient intake, food group consumption [38], and meal patterns [8]. This approach was used to enhance the accuracy of the method, but also to ensure that identification of generic meals was guided by nutrients that are of public health importance and are therefore nutrients that are central to dietary guidelines [41-43]. As such, any dietary assessment or feedback tool incorporating this generic meal approach will be of relevance to existing guidelines.

Previous studies on generic meals have not incorporated portion size, and this is evident in the small variance in nutrient intakes seen in those studies [7, 9]. The inclusion of portion size in the current study identified a greater range of nutrient intakes more reflective of the actual range of intakes observed in the original dataset. The approach used here assigned seven standard portion sizes to each generic meal based on actual meal intake weights. Comparable approaches have been used in established methods of dietary intake assessment. Standard portion sizes used in semi-quantitative food frequency questionnaires (FFQs) are based on known population consumption patterns [44, 45], and participants must choose the portion size that they typically consume for each food in the questionnaire [46, 47]. While 24-hour recalls (24HRs) have traditionally asked participants to estimate portion size using weight, household measures, or food models/images [48], web-based versions also provide the option of standard portion sizes when weight/volume is unknown [49-53]. The portion sizes presented for each food are based on various centiles of intake observed in national diet surveys [49-53]. The majority of FFQs are designed to include portion size [54, 55]. Despite this, it has been questioned whether this practice is warranted [54-56] given that improvements in estimates of nutrient intakes due to portion size are small [57, 58]. During the development of the generic meals method described in the present study, both the inclusion and exclusion of portion size was assessed (data not shown). The inclusion of portion size resulted in better agreement with the original data for absolute nutrient intakes but had no impact on agreement for energy-adjusted nutrient intakes.

The percentage differences for most nutrients between the generic and original data for mean daily nutrient intakes (mean difference of 5.6%) were small considering the error that is inherent in dietary intake assessment. For comparison, mean daily estimated energy and protein intake using 4-day diet records, three 24HRs, or a food frequency questionnaire differs from doubly labelled water energy values by 20%–27% and from protein intake

based on urinary nitrogen by 4–10% [59]. On comparing the results of the current study with previous studies of generic meals in the same population [7], improved agreement with the original data was observed for absolute nutrient intakes but similar for energy-adjusted intakes. The results of the current, however, found poorer agreement than those reported for generic meals in a Japanese population [9]. However, that study only reported energy-adjusted nutrient intakes and not absolute intakes. The differences between the current study and others, however, are not limited to portion size, and therefore the varying results may only be partly attributable to the inclusion of portion size in our study.

The current study demonstrated the ability of our generic meals approach to rank individuals within the sample population based on nutrient intakes. When individuals were classified into quartiles based on nutrient intakes using the generic data, the proportions classified in the same or adjacent quartiles as the original data are comparable to the equivalent proportions reported in previous generic meal research (72–88%) [7] and in validation studies comparing FFQs with diet records (65–88%) [60, 61] and 24HRs (55–86%) [62, 63]. This study also shows that participants can be classified according to nutrient-based guidelines using the generic dataset, with the agreement with the original dataset ranging from 55.3% to 91.5%. Several factors appear to have led to the range of agreements observed. Those nutrients with poorer agreement tended to be those expressed in %TEI, those that were not used as input variables in the clustering, and those that had three possible categories in the nutrient-based guidelines as opposed to two. Further work is warranted to determine the best combination of these variables or whether this approach is limited due to being unsuitable for certain nutrients.

The work presented here has significant potential for use in the provision of population and individual dietary advice. Currently, dietary advice is provided at a variety of levels. For example, government dietary guidelines are typically food- and nutrient-based and targeted at national populations or certain population subgroups based on characteristics such as sex and age [64]. Researchers have also shown that dietary advice can be targeted to population subgroups based on their metabolic profiles [65-67] and genetic profiles [67-69], while research has also examined the targeting of advice at an individual level based on existing dietary intakes [12, 70-74]. Those studies providing advice at the individual level use food-based dietary assessment to classify individuals according to nutrient-based guidelines (for example, high, adequate, or low). These classifications determine the advice provided [12, 70-74]. The current study demonstrated for the first time the ability of a generic meals approach to classify individuals according to nutrient-based dietary guidelines. The results show promise for the use of generic meals in dietary assessment and feedback, for example, it may be possible for images of the generic meals described in this study to be

presented to participants who could select the image of the meals and portions that most represent their dietary intakes, which in turn could be used to derive nutrient intakes, rather than recording each individual food consumed. Such a meal-based approach is likely to be more intuitive and less burdensome for end users [1, 6].

However, while the current study has described a method for the exploratory analysis of dietary intake data, this generic meal-based method has not yet been validated for use in a dietary intake assessment tool. The results presented here compare generic meal-based intakes with the food-based intakes from which those generic meals were derived. Given the inherent correlations between these datasets, this may overestimate the agreement between the two methods. The generic meal-based method described here may facilitate the development of a dietary assessment tool that is meal- rather than food-based. Further work is required to examine the generic meal-based method for use in collecting new dietary data from individuals and categorising their intakes according to nutrient-based dietary guidelines. It is also unknown in meal-based research whether certain generic meals better represent their food-based equivalents than others as analysis of accuracy is carried out at the mean daily intake level [7-9]. Analysis at the individual generic meal level could provide additional insights to further develop this process.

The work presented here should be considered in the context of its strengths and limitations. The primarily data-driven nature of the method limits the applicability of the results to other populations. However, the method itself can be applied to datasets from other populations to assess generic meal intakes. The nationally representative nature of the sample used here [16] strengthens the generalisability of the results within the Irish population. Despite this, the data used in this study were collected between 2008 and 2010, so the generic meals described may not necessarily be reflective of present-day meal intakes in Ireland. All self-reported dietary intake data are subject to measurement error [75]. The use of a 4-day weighed dietary intake record to gather the dietary data used in this study will mitigate, but not eliminate, this problem [59, 76].

4.6 Conclusion

The novel generic meals approach described here characterises the meals consumed in a nationally representative Irish population providing a means to estimate nutrient intakes based on meal rather than food intake at the sample population and individual levels. Future work will aim to determine the utility of this exploratory work in other datasets and assess the adequacy of this approach in the development of a meal-based approach to dietary intake assessment. This may enable a meal-based method of dietary assessment that is

less burdensome and more intuitive for individuals to complete compared to food-based methods [1, 6].

4.7 Tables

Table 4.1: The numbers of clusters identified within each of the groups of the Irish Food Pyramid with descriptions of the clusters.

Food Pyramid Groups	Number of Clusters / Groups	Description of Clusters / Groups
Vegetables, salad, and fruit	2	No clear distinction of traditional food groups or descriptions was observed between these clusters, simply called F&V1 and F&V2.
Cereal and breads, potatoes, pasta, and rice	3	1. Cereals 2. Potatoes 3. Breads, oats, pasta, and rice
Milk, yogurt, and cheese	2	1. Milk and Yogurt (and non-dairy alternatives) 2. Cheese
Meat, poultry, fish, eggs, beans, and nuts	2	No clear distinction of traditional food groups or descriptions was observed between these clusters, called proteinfoods1 and proteinfoods2.
Fats, spreads, and oils (not further split)	1	Not applicable: only 1 group.
Foods and drinks high in fat, sugar, or salt	2	No clear distinction of traditional food groups or descriptions was observed between these clusters, simply called HFSS1 and HFSS2.
Other (not clustered, manually grouped)	3	1. Alcoholic beverages 2. Non-alcoholic beverages 3. Soups, sauces, and miscellaneous foods

K-means clustering was carried out on the foods in each of the groups of the Irish food pyramid to identify clusters / subgroups within these groups. Clustering was performed separately for each group. Input variables for the clustering were the 12 nutrients in the Nutrient Rich Foods Index (NRF9.3) as reported by Fulgoni et al. (19). The number of clusters in each group was identified using 24 different indices to determine the optimal number of clusters between 2 and 10 based on the process described by Charrad et al. (20).

Table 4.2: Participant demographics.

	Male n = 740 (49.3%)	**Female** n = 760 (50.7%)	**Total** n = 1500
Age (y)	43.8 ± 17.2	45.2 ± 16.8	44.5 ± 17.0
BMI (kg/m^2)	27.6 ± 4.6	26.6 ± 5.3	27.1 ± 5.0
Weight (kg)	85.7 ± 14.7	69.7 ± 13.4	77.5 ± 16.1

Values are given as median ± standard deviation.

Table 4.3: Mean daily nutrient intakes estimated using the original food-based dataset and the generic meal-based dataset.

	Dataset		Difference	*P* value	Effect Size	Effect Size
	Original	Generic				
	Mean ± SD	Mean ± SD	%		Cohen's d	Magnitude
Energy (kJ)	8431 ± 2747	8088 ± 2394	-4.1	<0.001*	0.166	Negligible
Fat (g)	75.7 ± 29.4	73.6 ± 22.4	-2.8	0.002*	0.081	Negligible
Fat (%TEI)	33.8 ± 6.5	34.5 ± 2.9	2.1	<0.001*	0.108	Negligible
Saturated Fat (g)	29.7 ± 12.9	29.3 ± 9.4	-1.3	0.174	0.035	Negligible
Saturated Fat (%TEI)	13.3 ± 3.6	13.7 ± 1.4	3.0	<0.001*	0.135	Negligible
Monounsaturated Fat (g)	27.7 ± 11.4	26.7 ± 8.1	-3.6	<0.001*	0.094	Negligible
Monounsaturated Fat (%TEI)	12.3 ± 2.7	12.5 ± 1.1	1.6	0.003	0.077	Negligible
Polyunsaturated Fat (g)	13.3 ± 6.5	12.8 ± 3.9	-3.8	0.001*	0.088	Negligible
Polyunsaturated Fat (%TEI)	6.0 ± 2.2	6.0 ± 0.6	0.0	0.941	0.002	Negligible
Protein (g)	83.3 ± 26.9	79.9 ± 22.0	-4.1	<0.001*	0.156	Negligible
Protein (%TEI)	17.0 ± 3.6	16.8 ± 1.9	-1.2	0.020	0.060	Negligible
Carbohydrate (g)	228 ± 78.9	226 ± 70.1	-1.0	0.181	0.035	Negligible
Carbohydrate (%TEI)	42.9 ± 6.9	44.0 ± 3.5	2.6	<0.001*	0.179	Negligible
Total Sugars (g)	90.3 ± 43.1	87.7 ± 34.0	-2.9	0.006	0.072	Negligible
Total Sugars (%TEI)	16.9 ± 5.8	16.9 ± 3.1	0.0	0.765	0.008	Negligible
Added Sugars (g)	39.5 ± 31.1	39.2 ± 20.4	-0.8	0.573	0.015	Negligible
Added Sugars (%TEI)	7.2 ± 4.7	7.5 ± 2.5	4.2	0.002	0.080	Negligible
Dietary Fibre (g)	19.1 ± 7.9	19.1 ± 6.0	0.0	0.916	0.003	Negligible
Calcium (mg)	895 ± 369	844 ± 256	-5.7	<0.001*	0.169	Negligible
Iron (mg)	11.9 ± 5.0	11.3 ± 3.7	-5.0	<0.001*	0.145	Negligible
Potassium (mg)	3035 ± 966	2859 ± 813	-5.8	<0.001*	0.262	Small
Phosphorous (mg)	1378 ± 461	1297 ± 361	-5.9	<0.001*	0.236	Small
Sodium (mg)	2493 ± 901	2418 ± 712	-3.0	<0.001*	0.096	Negligible

	Dataset		Difference	*P* value	Effect Size	Effect Size
	Original	Generic				
	Mean ± SD	Mean ± SD	%		Cohen's d	Magnitude
Total Vitamin A (µg)	1023 ± 831	875 ± 326	-14.5	<0.001*	0.191	Negligible
Retinol (µg)	410 ± 623	307 ± 102	-25.3	<0.001*	0.167	Negligible
Carotene (µg)	3674 ± 3174	3408 ± 1623	-7.3	<0.001*	0.095	Negligible
Vitamin C (mg)	79.4 ± 52.4	71.9 ± 28.6	-9.4	<0.001*	0.178	Negligible
Vitamin D (µg)	3.2 ± 2.6	2.7 ± 0.9	-15.6	<0.001*	0.229	Small
Vitamin E (mg)	9.5 ± 4.9	8.9 ± 2.9	-6.3	<0.001*	0.116	Negligible
Total Folate (µg)	318 ± 152	285 ± 90.4	-10.3	<0.001*	0.255	Small
Vitamin B12 (µg)	4.7 ± 3.5	3.8 ± 1.1	-19.1	<0.001*	0.265	Small

Statistically significant *p* values refer to results of a paired t-test with Bonferroni-adjustment; *p* values < 0.002 are indicated with an asterisk. Effect sizes of < 0.2 were deemed to be negligible, and those ≥ 0.2 and < 0.5 were deemed to be small. TEI, Total Energy Intake.

Table 4.4: Correlation coefficients, cross-classification of quartiles, and percentage of participants within the Bland-Altman limits of agreement between the original food-based dataset and generic meal-based dataset.

	Correlation (Spearman)	Exact Agreement (%)	Exact Agreement + Adjacent (%)	Disagreement (%)	Extreme Disagreement (%)	Proportion Within Bland-Altman LOA (%)
Energy (kJ)	0.69	52	89	10	1	94.5
Fat (g)	0.54	43	82	14	3	94.9
Fat (%TEI)	0.36	34	75	19	6	94.9
Saturated Fat (g)	0.51	41	81	15	4	95.3*
Saturated Fat (%TEI)	0.39	35	76	20	5	95.8*
Monounsaturated Fat (g)	0.52	42	81	15	3	94.8
Monounsaturated Fat (%TEI)	0.31	33	72	21	6	95.1*
Polyunsaturated Fat (g)	0.48	40	79	17	4	95.4*
Polyunsaturated Fat (%TEI)	0.23	31	70	22	8	96.6*
Protein (g)	0.61	46	85	12	2	93.9
Protein (%TEI)	0.46	39	79	17	4	96.0*
Carbohydrate (g)	0.66	47	87	11	2	94.5
Carbohydrate (%TEI)	0.46	39	79	17	4	94.4
Total Sugars (g)	0.64	44	86	12	2	94.5
Total Sugars (%TEI)	0.47	40	80	15	4	95.1*
Added Sugars (g)	0.61	46	84	14	2	94.3
Added Sugars (%TEI)	0.57	45	83	14	3	94.7
Dietary Fibre (g)	0.66	46	87	11	2	94.3
Calcium (mg)	0.61	45	86	11	2	94.3
Iron (mg)	0.62	46	86	12	2	94.4
Potassium (mg)	0.75	53	91	8	1	94.9
Phosphorous (mg)	0.69	48	88	10	1	94.1

	Correlation (Spearman)	Exact Agreement (%)	Exact Agreement + Adjacent (%)	Disagreement (%)	Extreme Disagreement (%)	Proportion Within Bland-Altman LOA (%)
Sodium (mg)	0.56	42	82	15	3	93.7
Total Vitamin A (µg)	0.54	42	82	15	2	97.5*
Retinol (µg)	0.47	37	79	17	4	98.4*
Carotene (µg)	0.51	40	81	16	3	95.0*
Vitamin C (mg)	0.62	43	86	12	2	94.2
Vitamin D (µg)	0.34	33	74	21	6	96.0*
Vitamin E (mg)	0.51	39	82	14	4	95.0*
Total Folate (µg)	0.61	46	85	13	2	94.9
Vitamin B12 (µg)	0.44	38	79	16	5	97.5*

All correlations were statistically significant, $P < 0.001$. Exact agreement is the percentage of participants cross-classified into the same quartile; exact agreement + adjacent is the percentage of participants cross-classified into the same or adjacent quartiles; disagreement is the percentage of participants cross-classified 2 quartiles apart; and extreme disagreement is the percentage of participants cross-classified 3 quartiles apart. Proportion within Bland-Altman LOA of ≥ 95% are marked by an asterisk. The Bland-Altman LOA is the mean difference between the two datasets ± 1.96SD. LOA, limits of agreement; SD, standard deviation; TEI, total energy intake.

Table 4.5: Exact agreement between the original food-based dataset and generic meal-based dataset for categorisation of participants' nutrient intakes according to nutrient-based dietary guidelines.

Nutrient	Possible categories for classification of individual nutrient intakes	Proportion classified to the same category (%)
Protein (g/kg BW)	Low, adequate, and high	91.5
Carbohydrate (%TEI)	Low, adequate, and high	65.5
Fat (%TEI)	Low, adequate, and high	62.4
Monounsaturated Fat (%TEI)	Low, adequate, and high	85.5
Polyunsaturated Fat (%TEI)	Low, adequate, and high	55.3
Saturated Fat (%TEI)	Adequate and high	82.9
Salt (g)	Adequate and high	91.5
Dietary Fibre (g)	Low and adequate	87.1
Calcium (mg)	Low, adequate, and high	72.2
Iron (mg)	Low, adequate, and high	90.2
Vitamin A (µg)	Low, adequate, and high	76.8
Folate (µg)	Low, adequate, and high	72.1
Thiamin (mg)	Low and adequate	91.1
Riboflavin (mg)	Low and adequate	90.9
Vitamin B12 (µg)	Low and adequate	88.7
Vitamin C (mg)	Low, adequate, and high	72.7

Participants were placed in categories according to nutrient-based guidelines. The percentages in the table give the percentage of participants who were placed in the same category according to both the original and generic datasets.

4.8 Figures

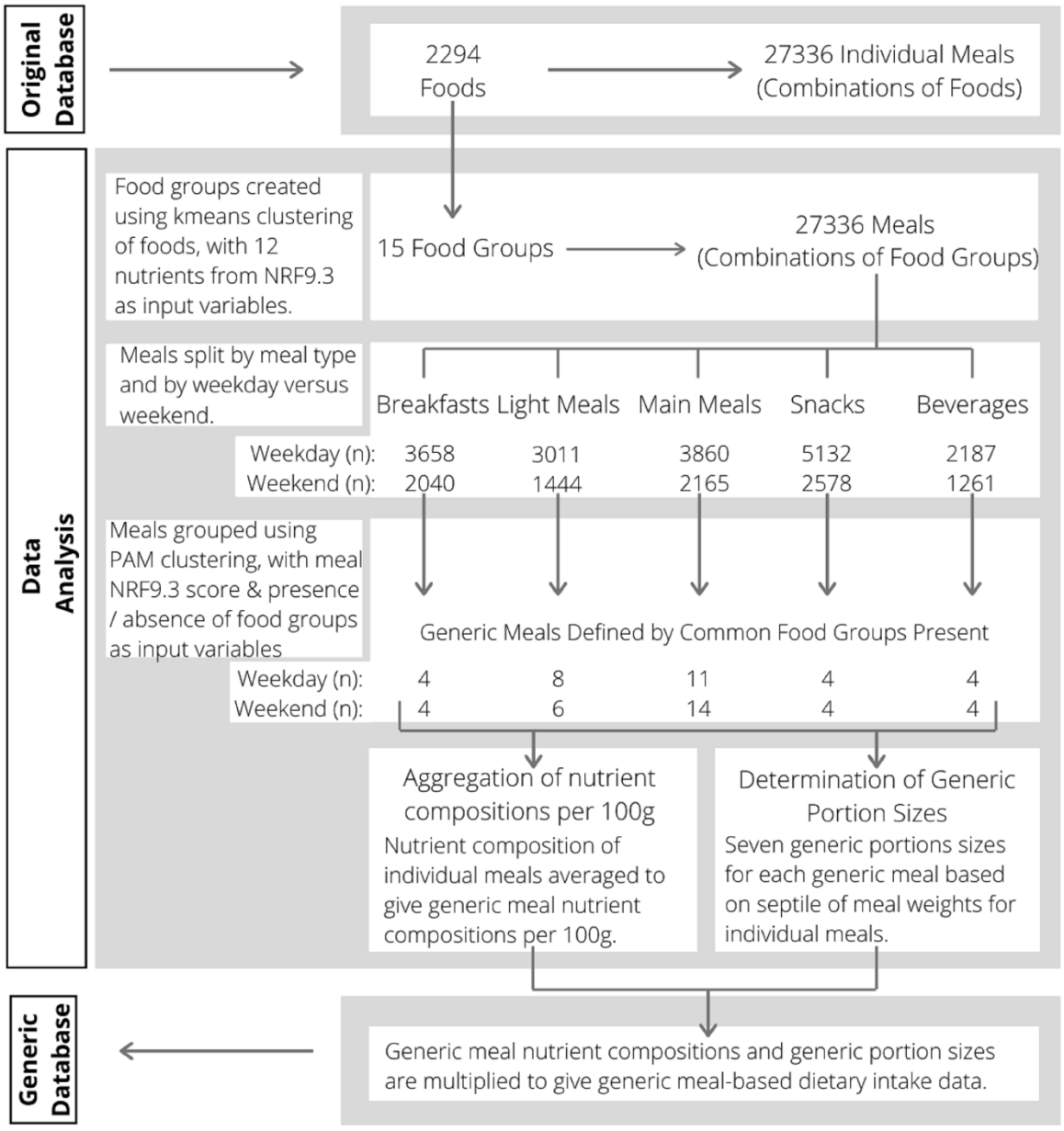


Figure 4.1: Flowchart of process for deriving generic meals.

The original database is from 1500 participants recording their dietary intakes over a period of 4 days during the National Adult Nutrition Survey in Ireland, 2008 to 2010. NRF9.3: Nutrient Rich Foods Index; PAM: partitioning around the medoids.

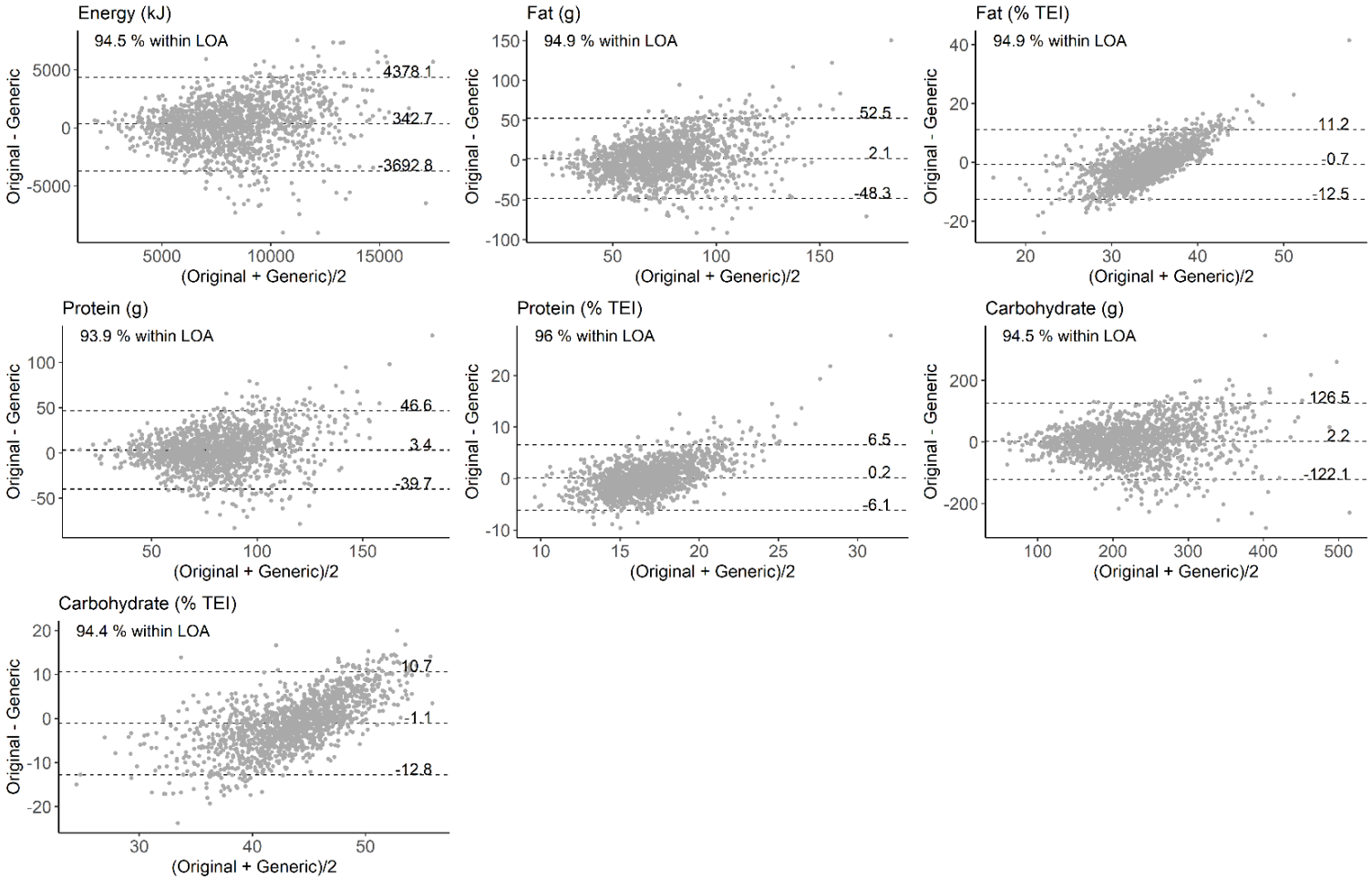


Figure 4.2: Bland-Altman plots for energy and macronutrients.

The middle dashed line, and associated number, represents the mean difference in mean daily intakes between the original and generic databases. The upper and lower dashed lines, and associated numbers, represent the upper and lower limits of agreement, respectively. Original and generic refer to the original and generic datasets. Each point represents an individual participant (n = 1500). TEI, Total Energy Intake.

4.9 References

1. Leech RM, Worsley A, Timperio A*, et al.* (2015) Understanding Meal Patterns: Definitions, Methodology and Impact on Nutrient Intake and Diet Quality. *Nutr Res Rev*, **28**(1), 1-21.

2. Dietary Guidelines Advisory Committee (2020) Scientific Report of the 2020 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture and the Secretary of Health and Human Services. Washington, DC: *U.S. Department of Agriculture, Agricultural Research Service*.

3. Dietary Guidelines Advisory Committee (2015) Scientific Report of the 2015 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Health and Human Services and the Secretary of Agriculture. Washington, DC: *U.S. Department of Agriculture, Agricultural Research Service*.

4. Camargo AM, Botelho AM, Dean M*, et al.* (2020) Meal Planning by High and Low Health Conscious Individuals During a Simulated Shop in the Supermarket: A Mixed Methods Study. *Appetite*, **144**, 104468.

5. Bisogni CA, Falk LW, Madore E*, et al.* (2007) Dimensions of Everyday Eating and Drinking Episodes. *Appetite*, **48**(2), 218-231.

6. Gibney MJ and Walsh MC (2013) The Future Direction of Personalised Nutrition: My Diet, My Phenotype, My Genes. *Proc Nutr Soc*, **72**(2), 219-225.

7. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

8. Uzhova I, Woolhead C, Timon CM*, et al.* (2018) Generic Meal Patterns Identified by Latent Class Analysis: Insights from NANS (National Adult Nutrition Survey). *Nutrients*, **10**(3), 310.

9. Murakami K, Livingstone MBE, Sasaki S*, et al.* (2018) Applying a Meal Coding System to 16-D Weighed Dietary Record Data in the Japanese Context: Towards the Development of Simple Meal-Based Dietary Assessment Tools. *J Nutr Sci*, **7**(e29), 1-10.

10. Leech RM, Boushey CJ, and McNaughton SA (2021) What Do Australian Adults Eat for Breakfast? A Latent Variable Mixture Modelling Approach for Understanding Combinations of Foods at Eating Occasions. *Int J Behav Nutr Phys Act*, **18**(1), 46.

11. Wei W, Jiang W, Huang J*, et al.* (2021) Association of Meal and Snack Patterns with Mortality of All-Cause, Cardiovascular Disease, and Cancer: The US National Health and Nutrition Examination Survey, 2003 to 2014. *J Am Heart Assoc*, **10**(13), e020254.

12. Forster H, Walsh MC, O'Donovan CB*, et al.* (2016) A Dietary Feedback System for the Delivery of Consistent Personalized Dietary Advice in the Web-Based Multicenter Food4Me Study. *J Med Internet Res*, **18**(6), e150.

13. Thompson FE, Subar AF, Loria CM*, et al.* (2010) Need for Technological Innovation in Dietary Assessment. *J Am Diet Assoc*, **110**(1), 48-51.

14. Michels KB and Schulze MB (2005) Can Dietary Patterns Help Us Detect Diet-Disease Associations? *Nutr Res Rev*, **18**(2), 241-248.

15. O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

16. Irish Universities Nutrition Alliance (2011) National Adult Nutrition Survey (2008–2010). Available at: https://www.iuna.net/surveyreports (accessed 14 December 2023).

17. James G, Witten D, Hastie T*, et al.* (2013) An Introduction to Statistical Learning with Applications in R. New York, USA: Springer.

18. Food Safety Authority of Ireland (2019) Healthy Eating, Food Safety and Food Legislation: A Guide Supporting the Healthy Ireland Food Pyramid. Dublin, Ireland: *Food Safety Authority of Ireland*.

19. Fulgoni VL, Keast DR, and Drewnowski A (2009) Development and Validation of the Nutrient-Rich Foods Index: A Tool to Measure Nutritional Quality of Foods. *J Nutr*, **139**(8), 1549-1554.

20. Charrad M, Ghazzali N, Boiteau V*, et al.* (2014) Nbclust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, **61**(6), 1-36.

21. Regulation (EU) No 1169/2011 (2011) Regulation of the European Parliament and of the Council of 25 October 2011 on the Provision of Food Information to Consumers. European Union.

22. EFSA (European Food Safety Authority) (2019) Dietary Reference Values for Nutrients. Summary Report. no. e15121.

23. Kaufman L and Rousseeuw PJ (1990) Partitioning around Medoids (Program PAM). In *Finding Groups in Data: An Introduction to Cluster Analysis*, pp. 68-125. Hoboken, NJ, USA: John Wiley & Sons Inc.

24. Maechler M, Rousseeuw PJ, Struyf A*, et al.* (2022) Cluster: Cluster Analysis Basics and Extensions. R Package Version 2.1.3. Available at: https://CRAN.R-project.org/package=cluster (accessed 30 May 2022).

25. R Core Team (2022) R: A Language and Environment for Statistical Computing, 4.2.2 ed. Vienna, Austria: R Foundation for Statistical Computing.

26. RStudio Team (2022) Rstudio: Integrated Development for R., 2022.07.2+576 ed. Boston, MA, USA: RStudio PBC.

27. Bland MJ and Altman DG (1999) Measuring Agreement in Method Comparison Studies. *Stat Methods Med Res*, **8**(2), 135-160.

28. Institute of Medicine (2005) Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids. Washington, DC, USA: The National Academies Press.

29. Institute of Medicine (2000) Dietary Reference Intakes: Applications in Dietary Assessment. Washington, DC, USA: The National Academies Press.

30. World Health Organization (2007) Protein and Amino Acid Requirements in Human Nutrition, *WHO Technical Report Series*. Geneva, Switzerland: WHO Press.

31. Black AE (2000) Critical Evaluation of Energy Intake Using the Goldberg Cut-Off for Energy Intake:Basal Metabolic Rate. A Practical Guide to Its Calculation, Use and Limitations. *Int J Obes*, **24**, 1119-1130.

32. Newby PK, Sc D, Tucker KL*, et al.* (2004) Empirically Derived Eating Patterns Using Factor or Cluster Analysis: A Review. *Nutr Rev*, **62**(5), 177-203.

33. Pennington JAT and Fisher RA (2009) Classification of Fruits and Vegetables. *J Food Compost Anal*, **22**, S23-S31.

34. Burden S, Probst YC, Steel DG*, et al.* (2009) Identification of Food Groups for Use in a Self-Administered, Computer-Assisted Diet History Interview for Use in Australia. *J Food Compost Anal*, **22**(2), 130-136.

35. Patras A, Brunton NP, Downey G*, et al.* (2011) Application of Principal Component and Hierarchical Cluster Analysis to Classify Fruits and Vegetables Commonly Consumed in Ireland Based on in Vitro Antioxidant Activity. *J Food Compost Anal*, **24**(2), 250-256.

36. Dy JG (2008) Unsupervised Feature Selection. In *Computational Methods of Feature Selection*, 1st ed. [H Liu and H Motoda, editors]. Boca Raton, FL, USA: Chapman & Hall/CRC.

37. Pfaffel O (2021) Feature Importance for Partitional Clustering. Available at: https://CRAN.R-project.org/package=FeatureImpCluster (accessed 21 June 2021).

38. An R (2016) Weekend-Weekday Differences in Diet among U.S. Adults, 2003–2012. *Ann Epidemiol*, **26**(1), 57-65.

39. Haines PS, Hama MY, Guilkey DK*, et al.* (2003) Weekend Eating in the United States Is Linked with Greater Energy, Fat, and Alcohol Intake. *Obes Res*, **11**(8), 945-949.

40. Racette SB, Weiss EP, Schechtman KB*, et al.* (2008) Influence of Weekend Lifestyle Patterns on Body Weight. *Obesity*, **16**(8), 1826-1830.

41. U.S. Department of Agriculture and U.S. Department of Health and Human Services (2020) Dietary Guidelines for Americans, 2020-2025, 9th ed.

42. Food Safety Authority of Ireland (2011) Scientific Recommendations for Healthy Eating Guidelines in Ireland. Dublin, Ireland: *Food Safety Authority of Ireland*.

43. National Health and Medical Research Council (2013) Australian Dietary Guidelines. Canberra, Australia: *Australian Department of Health and Ageing*.

44. Babic D, Sindik J, and Missoni S (2014) Development and Validation of a Self-Administered Food Frequency Questionnaire to Assess Habitual Dietary Intake and Quality of Diet in Healthy Adults in the Republic of Croatia. *Coll Antropol*, **38**(3), 1017-1026.

45. Park MK, Noh HY, Song NY*, et al.* (2012) Validity and Reliability of a Dish-Based, Semi-Quantitative Food Frequency Questionnaire for Korean Diet and Cancer Research. *Asian Pac J Cancer Prev*, **13**(2), 545-552.

46. Knudsen VK, Hatch EE, Cueto H*, et al.* (2016) Relative Validity of a Semi-Quantitative, Web-Based FFQ Used in the 'Snart Foraeldre' Cohort - a Danish Study of Diet and Fertility. *Public Health Nutr*, **19**(6), 1027-1034.

47. Yang YJ, Kim MK, Hwang SH*, et al.* (2010) Relative Validities of 3-Day Food Records and the Food Frequency Questionnaire. *Nutr Res Pract*, **4**(2), 142-148.

48. Shim JS, Oh K, and Kim HC (2014) Dietary Assessment Methods in Epidemiologic Studies. *Epidemiol Health*, **36**, e2014009.

49. Timon CM, Blain RJ, McNulty B*, et al.* (2017) The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. *J Med Internet Res*, **19**(5), e158.

50. Foster E, Matthews JN, Lloyd J*, et al.* (2008) Children's Estimates of Food Portion Size: The Development and Evaluation of Three Portion Size Assessment Tools for Use with Children. *Br J Nutr*, **99**(1), 175-184.

51. Bradley J, Simpson E, Poliakov I*, et al.* (2016) Comparison of Intake24 (an Online 24-H Dietary Recall Tool) with Interviewer-Led 24-H Recall in 11–24 Year-Old. *Nutrients*, **8**(6), 358.

52. Carter MC, Albar SA, Morris MA*, et al.* (2015) Development of a UK Online 24-H Dietary Assessment Tool: Myfood24. *Nutrients*, **7**(6), 4016-4032.

53. 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**(7), 1055-1064.

54. Molag ML, de Vries JH, Ocke MC*, et al.* (2007) Design Characteristics of Food Frequency Questionnaires in Relation to Their Validity. *Am J Epidemiol*, **166**(12), 1468-1478.

55. Cade J, Thompson R, Burley V*, et al.* (2002) Development, Validation and Utilisation of Food-Frequency Questionnaires – a Review. *Public Health Nutr*, **5**(4), 567-587.

56. Thompson FE and Subar AF (2017) Dietary Assessment Methodology. In *Nutrition in the Prevention and Treatment of Disease*, 4th ed., pp. 5-48 [A Coulston, C Boushey, M Ferruzzi and L Delahanty, editors]. Amsterdam, Netherlands: Academic Press.

57. Cummings SR, Block G, McHenry K*, et al.* (1987) Evaluation of Two Food Frequency Methods of Measuring Dietary Calcium Intake. *Am J Epidemiol*, **126**(5), 796-802.

58. Flegal KM and Larkin FA (1990) Partitioning Macronutrient Intake Estimates from a Food Frequency Questionnaire. *Am J Epidemiol*, **131**(6), 1046-1058.

59. Prentice RL, Mossavar-Rahmani Y, Huang Y*, et al.* (2011) Evaluation and Comparison of Food Records, Recalls, and Frequencies for Energy and Protein Assessment by Using Recovery Biomarkers. *Am J Epidemiol*, **174**(5), 591-603.

60. Araujo MC, Yokoo EM, and Pereira RA (2010) Validation and Calibration of a Semiquantitative Food Frequency Questionnaire Designed for Adolescents. *J Am Diet Assoc*, **110**(8), 1170-1177.

61. Fallaize R, Forster H, Macready AL*, et al.* (2014) Online Dietary Intake Estimation: Reproducibility and Validity of the Food4Me Food Frequency Questionnaire against a 4-Day Weighed Food Record. *J Med Internet Res*, **16**(8), e190.

62. Dehghan M, del Cerro S, Zhang X*, et al.* (2012) Validation of a Semi-Quantitative Food Frequency Questionnaire for Argentinean Adults. *PLoS One*, **7**(5), e37958.

63. Liu L, Wang PP, Roebothan B*, et al.* (2013) Assessing the Validity of a Self-Administered Food-Frequency Questionnaire (FFQ) in the Adult Population of Newfoundland and Labrador, Canada. *Nutr J*, **12**(49).

64. Herforth A, Arimond M, Álvarez-Sánchez C*, et al.* (2019) A Global Review of Food-Based Dietary Guidelines. *Adv Nutr*, **10**(4), 590-605.

65. O'Donovan CB, Walsh MC, Woolhead C*, et al.* (2017) Metabotyping for the Development of Tailored Dietary Advice Solutions in a European Population: The Food4Me Study. *Br J Nutr*, **118**(8), 561-569.

66. O'Donovan CB, Walsh MC, Nugent AP*, et al.* (2015) Use of Metabotyping for the Delivery of Personalised Nutrition. *Mol Nutr Food Res*, **59**(3), 377-385.

67. Celis-Morales C, Livingstone KM, Marsaux CF*, et al.* (2017) Effect of Personalized Nutrition on Health-Related Behaviour Change: Evidence from the Food4Me European Randomized Controlled Trial. *Int J Epidemiol*, **46**(2), 578-588.

68. Nielsen DE and El-Sohemy A (2014) Disclosure of Genetic Information and Change in Dietary Intake: A Randomized Controlled Trial. *PLoS One*, **9**(11), e112665.

69. Horne JR, Gilliland JA, O'Connor CP*, et al.* (2020) Change in Weight, Bmi, and Body Composition in a Population-Based Intervention Versus Genetic-Based Intervention: The Now Trial. *Obesity*, **28**(8), 1419-1427.

70. Wright JL, Sherriff JL, Dhaliwal SS*, et al.* (2011) Tailored, Iterative, Printed Dietary Feedback Is as Effective as Group Education in Improving Dietary Behaviours: Results from a Randomised Control Trial in Middle-Aged Adults with Cardiovascular Risk Factors. *Int J Behav Nutr Phys Act*, **8**, 43.

71. Kroeze W, Oenema A, Campbell M*, et al.* (2008) The Efficacy of Web-Based and Print-Delivered Computer-Tailored Interventions to Reduce Fat Intake: Results of a Randomized, Controlled Trial. *J Nutr Educ Behav*, **40**(4), 226-236.

72. Robroek SJ, Bredt FJ, and Burdorf A (2007) The (Cost-)Effectiveness of an Individually Tailored Long-Term Worksite Health Promotion Programme on Physical Activity and Nutrition: Design of a Pragmatic Cluster Randomised Controlled Trial. *BMC Public Health*, **7**, 259.

73. Xu Z, Geng J, Zhang S*, et al.* (2020) A Mobile-Based Intervention for Dietary Behavior and Physical Activity Change in Individuals at High Risk for Type 2 Diabetes Mellitus: Randomized Controlled Trial. *JMIR Mhealth Uhealth*, **8**(11), e19869.

74. Rollo ME, Haslam RL, and Collins CE (2020) Impact on Dietary Intake of Two Levels of Technology-Assisted Personalized Nutrition: A Randomized Trial. *Nutrients*, **12**(11), 3334.

75. Subar AF, Freedman LS, Tooze JA*, et al.* (2015) Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr*, **145**(12), 2639-2645.

76. Bingham SA, Gill C, Welch A*, et al.* (1994) Comparison of Dietary Assessment Methods in Nutritional Epidemiology: Weighed Records V. 24 H Recalls, Food-Frequency Questionnaires and Estimated-Diet Records. *Br J Nutr*, **72**(4), 619-643.

4.10 Supplementary Material

Supplementary Table 4.1: The values from the 24 indices used to determine the number of clusters when clustering foods in the fruit and vegetables level of the Irish food pyramid.

Index	Index Value for the Top 3 Cluster Numbers that were Deemed Optimal Most Frequently			Interpretation [a]
	2	5	10	
KL	83.1*	17.8	35.9	Max value
CH	148.4*	92.9	82.5	Max value
Hartigan	31.2	17.2	9.0*	Max difference between hierarchy levels of index
CCC	-3.8	2.1	21.4*	Max value
Scott	551.3	2159.6	3909.0*	Max difference between hierarchy levels of index
Marriot	4.02^{E+29}	6.07^{E+28}*	4.23^{E+27}	Max value of second differences between levels of the index
TrCovW	117175.8	69270.3	26588.9*	Max difference between hierarchy levels of index
TraceW	3845.3	2765.5*	1873.9	Max value of second differences between levels of the index
Friedman	7.8	18.6*	28.7	Max difference between hierarchy levels of index
Rubin	1.3	1.9*	2.8	Min value of second differences between levels
Cindex	0.14	0.11	0.13	Min value
DB	1.62	1.49	1.38*	Min value
Silhouette	0.5*	0.3	0.2	Max value
Duda	1.2*	1.1	1.9	Smallest number of clusters such that index > critical value
Beale	-1.5*	-0.6	-3.7	Number of clusters such that critical value ≥ alpha
Ratkowsky	0.31*	0.29	0.25	Max value
Ball	1922.7	553.1	187.4	Max difference between hierarchy levels of index
Ptbiserial	0.5*	0.5	0.5	Max value
McClain	0.2*	0.5	0.7	Min value
Dunn	0.0*	0.0	0.0	Max value
Hubert				Graphical method [b]
SDindex	2.27*	2.94	3.29	Min value
Dindex				Graphical method [b]
SDbw	1.5	1.7	1.7	Min value
Total	10	4	5	

The optimal number of clusters as chosen by each index is indicated by an asterisk. The number of clusters chosen for the subsequent clustering step was the cluster number that was deemed to be optimal most frequently by the various indices. The range of cluster numbers that was assessed was 2–10. [a] Information on the interpretation of the various indices was adapted from Charrad *et al.* [20]. [b] The two graphical methods used involved choosing the cluster number where a peak occurred on the second differences plot for those indices.

Supplementary Table 4.2: The values from the 24 indices used to determine the number of clusters when clustering meals consumed at breakfast on weekends.

Index	Index Value for the Top 3 Cluster Numbers that were Deemed Optimal Most Frequently			Interpretation [a]
	4	8	12	
KL	1.27	0.78	1.07	Max value
CH	0.21	2.97*	2.02	Max value
Hartigan	1.0	0.0*	0.2	Max difference between hierarchy levels of index
CCC	-243.8*	-472.6	-563.4	Max value
Scott	5900.9	7532.8	9252.3*	Max difference between hierarchy levels of index
Marriot	4.64^{E+37}	8.34^{E+37}	8.07^{E+37}*	Max value of second differences between levels of the index
TrCovW	3.60^{E+11}	3.53^{E+11}	3.53^{E+11}	Max difference between hierarchy levels of index
TraceW	2.40^{E+06}	2.38^{E+06}	2.38^{E+06}	Max value of second differences between levels of the index
Friedman	20.6	21.9	23.2*	Max difference between hierarchy levels of index
Rubin	2.38	2.40*	2.40	Min value of second differences between levels
Cindex	0.23	0.22	0.22	Min value
DB	2.87	3.48	5.20	Min value
Silhouette	0.09*	0.02	-0.03	Max value
Duda	1.0*	1.6	1.0	Smallest number of clusters such that index > critical value
Beale	0.0*	-2.6	0.0	Number of clusters such that critical value ≥ alpha
Ratkowsky	0.031	0.075	0.076*	Max value
Ball	6.01^{E+05}	2.97^{E+05}	1.98^{E+05}	Max difference between hierarchy levels of index
Ptbiserial	0.05	0.23	0.25	Max value
McClain	0.002*	0.015	0.018	Min value
Dunn	0.17*	0.10	0.10	Max value
Hubert				Graphical method [b]
SDindex	1.43*	1.80	2.34	Min value
Dindex		*		Graphical method [b]
SDbw	1.74	1.51	0.68*	Min value
Total	7	4	5	

The optimal number of clusters as chosen by each index is indicated by an asterisk. The number of clusters most frequently deemed to be optimal by the various indices was chosen for the subsequent clustering step. The range of cluster numbers that was assessed was 4–15. [a] Information on the interpretation of the various indices was adapted from Charrad *et al.* [20]. [b] The two graphical methods used involved choosing the cluster number where a peak occurred on the second differences plot for those indices.

Supplementary Table 4.3: Food-based descriptions of generic meals and their portion sizes.

Time of Week	Meal Type	Food-Based Description	Portion Sizes (g)						
			1	2	3	4	5	6	7
Weekend	Breakfast	Bread/toast with fat spread, tea with milk, eggs, sausages, rashers, and pudding.	275	372	440	515	600	747	1040
Weekend	Breakfast	Bread/toast, with fat spread and marmalade/jam. Tea with milk and sugar.	244	338	391	480	554	678	939
Weekend	Breakfast	Bread/porridge, tea with milk, and fruit juice/whole fruit	214	379	473	568	676	845	1135
Weekend	Breakfast	Cereal with tea and milk.	151	252	334	425	499	594	831
Weekday	Breakfast	Toast or porridge.	74	137	200	259	356	462	695
Weekday	Breakfast	Bread/toast with fat spread and jam/marmalade, tea with milk and sugar.	270	340	391	458	549	649	911
Weekday	Breakfast	Bread or porridge, tea with milk, and fruit juice/whole fruit	350	457	542	619	715	882	1191
Weekday	Breakfast	Cereal with milk and tea.	170	275	360	443	512	630	917
Weekend	Light Meal	Sandwich with bread, mayonnaise/dressing, meat with veg and tea with milk.	422	509	574	631	709	846	1190
Weekend	Light Meal	Sandwich with bread, fat spread, meat/eggs, tea and milk.	294	378	438	500	612	741	996
Weekend	Light Meal	Bread and soup and fruit.	122	235	300	384	456	648	932
Weekend	Light Meal	Sandwich with bread, cheese, ham, and tea.	157	296	392	441	534	715	951
Weekend	Light Meal	Sandwich with fat spread and meat, with cookie/chocolate, tea, and milk.	334	406	463	528	627	745	959
Weekend	Light Meal	Cookies/chocolate/cake with tea and milk.	236	318	406	486	581	713	1040
Weekday	Light Meal	Sandwich with bread, fat spread, eggs, sausages, and bacon, with tea and milk and sugar.	334	404	456	527	601	679	853
Weekday	Light Meal	Bread, fat spread, with banana, and tea with milk.	381	458	545	635	714	854	1133
Weekday	Light Meal	Bread with chicken and water.	221	350	438	544	640	757	949
Weekday	Light Meal	Ham sandwich with tea and milk.	194	365	435	480	592	709	935
Weekday	Light Meal	Bread with soup and/or whole fruit.	148	243	301	380	481	620	916

Time of Week	Meal Type	Food-Based Description	Portion Sizes (g)						
			1	2	3	4	5	6	7
Weekday	Light Meal	Sandwich with bread, mayonnaise/dressing, meat with veg and water.	242	417	508	588	734	853	1120
Weekday	Light Meal	Sandwich with bread, cheese, fat spread, ham, with tea and milk.	300	400	449	515	615	710	930
Weekday	Light Meal	Tea and milk with cookie/chocolate/cake.	237	321	408	469	577	676	947
Weekend	Main Meal	Potatoes with beef, veg, tea, milk, and sweet dessert.	675	771	935	1062	1154	1269	1604
Weekend	Main Meal	Bread/rice with chicken curry /stir fry.	358	518	668	795	933	1123	1620
Weekend	Main Meal	Chips/potatoes with meat and tea.	433	508	607	678	772	943	1300
Weekend	Main Meal	Potatoes, ham/beef, green veg, and water.	390	595	740	825	934	1109	1528
Weekend	Main Meal	Potatoes, beef, gravy, carrots, and green veg.	315	404	500	621	721	861	1132
Weekend	Main Meal	Potatoes, beef, veg, gravy, water.	481	637	756	909	1073	1241	1483
Weekend	Main Meal	Potatoes with chicken, veg, and water.	451	651	765	867	975	1176	1347
Weekend	Main Meal	Bread with fat spread, ham, egg, onions, coleslaw, tea and milk.	486	641	760	834	888	1007	1346
Weekend	Main Meal	Potatoes/chips, beef, veg, sugar-sweetened beverage.	382	498	601	716	811	936	1124
Weekend	Main Meal	Bread, rice/spaghetti, sauce, chicken, pulses, veg, and water.	328	578	713	817	960	1082	1613
Weekend	Main Meal	Potatoes, beef/bacon/ham, vegetables, cookie, tea, and milk.	588	691	808	894	1010	1160	1439
Weekend	Main Meal	Bread with meat, tea, milk, and cookies.	410	596	690	771	908	1108	1442
Weekend	Main Meal	Rice/bread/pizza with beef/poultry dish.	200	346	429	547	700	854	1298
Weekend	Main Meal	Potatoes with peas and meat.	241	348	454	571	700	825	1006
Weekday	Main Meal	Bread, beef dishes, tea, milk, and sugar.	411	533	630	740	895	1041	1358
Weekday	Main Meal	Potatoes, meat/fish, and veg	338	430	520	605	701	878	1165
Weekday	Main Meal	Potatoes, chicken, peas.	264	382	452	571	656	790	1113
Weekday	Main Meal	Rice/pasta/bread, meat products, beg, sauce, and water.	436	594	698	823	948	1083	1291
Weekday	Main Meal	Rice/pasta, chicken dishes, water.	368	542	654	764	884	1025	1322
Weekday	Main Meal	Rice/pasta/bread, chicken/fish, veg, tea, milk, sugar, chocolate.	502	682	763	835	909	1036	1314
Weekday	Main Meal	Potatoes, chicken/fish, veg, water, tea, milk.	560	663	772	897	989	1097	1434
Weekday	Main Meal	Rice/pasta/bread, beef/chicken.	210	347	437	525	657	777	1067

Time of Week	Meal Type	Food-Based Description	Portion Sizes (g)						
			1	2	3	4	5	6	7
Weekday	Main Meal	Potatoes, beef/ham, veg.	534	702	795	909	1011	1163	1430
Weekday	Main Meal	Chips/potatoes, fish/beef/bacon/chicken, water.	353	527	629	716	840	979	1310
Weekday	Main Meal	Potatoes, beef/fish/chicken.	375	565	657	782	901	1043	1374
Weekend	Snack	Bread/scone with fat spread and tea with milk.	169	296	348	398	455	546	767
Weekend	Snack	Cookie/chocolate with tea and milk.	202	252	283	312	345	403	632
Weekend	Snack	Cookie/chocolate with water.	26	66	200	267	354	522	809
Weekend	Snack	Piece of fruit.	49	113	160	250	359	480	790
Weekday	Snack	Cookies/chocolate.	22	40	64	129	250	363	592
Weekday	Snack	Bread with fat spread and tea and milk.	201	301	337	384	437	527	766
Weekday	Snack	Cookies/chocolate and tea and milk and sugar.	176	247	277	307	340	393	601
Weekday	Snack	Fruit and water.	97	140	200	275	372	519	761
Weekend	Beverage	Spirits or lager with cola or red bull.	200	250	330	468	600	915	1125
Weekend	Beverage	Lager/stout.	175	378	574	663	1144	1665	2272
Weekend	Beverage	Milk and tea/coffee.	220	245	260	275	309	332	542
Weekend	Beverage	Water.	160	250	284	366	568	628	1000
Weekday	Beverage	Coffee/tea with milk and sugar.	211	255	272	284	308	330	470
Weekday	Beverage	Water.	119	225	284	330	500	568	1000
Weekday	Beverage	Tea and milk.	210	232	245	262	287	315	462
Weekday	Beverage	Lager/stout.	145	375	568	638	990	1148	1722

CHAPTER 5

Comparison of Statistical Approaches in the Derivation of Meal Patterns

5.1 Abstract

Research on meal patterns investigating the content of different combinations of meals consumed over course of a day has increased in recent years. A variety of statistical techniques have been used to identify these patterns. Despite this, no study has assessed whether applying different statistical approaches to the same data gives rise to different outcomes. The objective was to identify meal patterns using different methods and compare the resulting meal patterns that were identified. This study is a secondary analysis of data from NHANES 2017–2018. A small number of generic meals were identified that were representative of the larger number of actual meals consumed with regard to their food group and nutrient content. Combinations of these generic meals consumed (meal patterns) were identified using three different statistical approaches: partitioning around the medoids clustering, principal component analysis (PCA), and latent class analysis (LCA). The number of meal patterns identified were 26 by clustering, 18 by PCA, and 17 by LCA. Meal patterns in which individuals skipped certain meal types were observed using clustering and LCA, but not PCA. There was only one meal pattern that was identical when comparing clustering patterns with LCA patterns, i.e., the generic meals consumed at all five meal types (breakfast, lunch, dinner, snacks, and beverages) were the same. No other identical patterns were identified. However, for all comparisons (clustering v. PCA, clustering v. LCA, and PCA v. LCA), there were two meal patterns in each in which identical generic meals were consumed in four of the five meal types. In conclusion, using different statistical approaches to meal pattern analysis gave rise to a differing number of meal patterns. A similar identification of meal skipping was observed between clustering and LCA, but not PCA, but considerable differences in the content of the meal patterns were observed across the three methods.

5.2 Introduction

Meal-based research provides insights into aspects of dietary intake that cannot be provided by research that focuses on nutrients or foods [1, 2]. For example, temporal meal patterns provide information on the timing and distribution of meals; content meal patterns provide information on what meals are combined to make meals and what meals are combined over time; and context meal patterns considers those aspects that are external to the meal itself such as whether someone eats alone or in company, the location of meals, and any other activities being carried out during meals [1].

A variety of studies have been conducted using various statistical techniques including clustering [3], LCA [4], and PCA [5, 6] to identify meal patterns. These methods identify meal patterns in different ways. Both clustering [7] and LCA [8] classify individuals into mutually exclusive groups in such a way that individuals within groups have similar meal patterns to each other. In clustering this similarity is based on the distance among observations in a geometric space [7]. In LCA, similarity is based on the probability of individuals consuming similar combinations of meals [8, 9]. PCA, does not inherently classify individuals into groups. Instead, it arranges meals into principal components (meal patterns) based on the degree to which the various meals are correlated with each other [7]. A principal component score can be calculated for each individual, however, representing how closely their meal intake pattern aligns with each of the principal component meal patterns; in practice these scores are used as a proxy for group membership in PCA research [10, 11]. Despite these varied techniques being applied to answer the same research question in different populations, it is not clear whether the results from these techniques are comparable given that they operate in different ways.

Clustering [12], PCA [13], and LCA [10], have also been used in dietary pattern research which investigates the combinations of foods or food groups that people consume over a day. In the area of dietary patterns, research has been conducted in which these statistical methods have been applied to the same datasets to compare and contrast the dietary patterns that they identify [14]. Broadly speaking similar patterns have been identified with the number of patterns typically being in the range of two to five. Patterns described as “Western” (high intakes of fried foods, low fibre foods, and sugar sweetened beverages) and “Prudent” (high intakes of fish, fruit, vegetables, and wholegrains) have been commonly identified [15-18].

A much smaller volume of research has been carried out in meal patterns compared to dietary patterns, as highlighted by a recent review that identified 10 studies using data-driven methods for meal pattern analysis [1]. No studies have been identified that have

attempted to identify meal patterns within a single dataset using multiple different data-driven techniques. This is in contrast with the area of dietary patterns research where a recent review identified 153 studies that used data-driven methods to identify those patterns, 19 of which used more than one method on the same dataset [14]. Despite the use of different statistical methods to derive meal patterns, it is currently unknown whether the methods have comparable outcomes when used on the same dataset. The aim of this study is to apply clustering, PCA, and LCA to the NHANES dataset to identify meal patterns and compare the outcomes from the different methods.

5.3 Methods

The current study is a secondary analysis of data from the 2017–2018 National Health and Nutrition Examination Survey (NHANES) from the USA [19]. In summary, it includes data from non-institutionalised civilian residents of the USA aged 18–80 years. Each participant completed two 24-hour recalls conducted by trained interviewers, with the first being in person and the second being over the telephone using a standardised process. The nutrient composition of the foods reported in the 24-hour recalls was derived from the food and nutrient database for dietary studies (FNDDS) 2017–2018 [20].

The meal pattern analysis within the current study was conducted in three broad steps. Firstly, foods were grouped into food groups based on how similar they were with regard to their nutrient composition. Secondly, individual meals were grouped into generic meals based on how similar they were with regard to their nutrient profile and food group content. Thirdly, participants are grouped into meal patterns based on having consumed similar combinations of generic meals. The first two steps have previously been described elsewhere [21]. The third step of identifying meal patterns was carried out using three different methods (clustering, PCA, and LCA) and comparisons were made among the outcomes for all three methods. All steps are described in the following sections.

5.3.1 Step 1. Food Groups

Participants' reported food intakes were grouped based on the five groups used in the USA dietary guidelines: fruit; vegetables; grains; protein foods; dairy. Foods that did not belong in any of those groups were assigned to one of the following five groups: foods high in fat, sugar, or salt; fats and oils; non-alcoholic beverages; alcoholic beverages; and other, with 10 food groups in total. Each of the 10 food groups were analysed separately using *k*-means clustering to create subgroups in such a way that foods within a subgroup were most similar to each other in relation to nutrients listed within the Nutrient Rich Foods (NRF9.3)

index: protein, fibre, vitamin A, vitamin C, vitamin E, calcium, iron, magnesium, potassium, saturated fatty acids, total sugar, and sodium [22]. Food nutritional composition was expressed in units per 100kcal and were z-standardised before clustering. The number of clusters was determined by applying 24 indices and choosing the number of clusters that was most frequently proposed among the indices [23]. The range of values assessed for the number of clusters in the data was between two and eight inclusive. The clustering of the non-alcoholic beverages group resulted in unbalanced clusters with very small numbers of foods in some clusters, so this group was split into energy containing and energy free non-alcoholic beverages instead of clustering. Two of the original food groups (alcoholic beverages and other) were not further split due to the small number of foods in each of these groups. Fats and oils were not further split, instead their original three subgroups from the FNDDS were kept: animal fats; margarines; and dressings and vegetable oils [20]. In total, 23 food groups were created. The original food groups and their subsequent subgroups after *k*-means clustering are summarised in **Table 5.1**.

5.3.2 Step 2. Generic Meals

The descriptions of participants' individual meals were changed so that the individual foods reported in the dietary recalls were replaced by the food group to which that food belonged. From this point the data were analysed separately for males and females. Meals were also split according to meal type. The original NHANES data used in this analysis included 18 different meal types. These meal types were condensed to five for the purpose of this analysis: breakfast (comprising breakfast, desayuno, and amuerzo), lunch (lunch, comida, and brunch), dinner (dinner, supper, and cena), snack (snack, merienda, entre comida, botana, bocadillo, tentempie, and extended consumption), and beverages (drink and bebida). These categorisations were based on previous methods from the literature [24-27].

Similar meals within each meal type were grouped using partitioning around the medoids (PAM) clustering. Similarity was based on 24 variables: 23 binary variables, each one representing one of the food groups and indicating whether that food group was present or not in the meal, and a single numeric value, namely the NRF9.3 index score [22] for each meal, which was used as an indication of the nutritional quality of the meal. The same approach described in the previous section was taken to determine the number of clusters present. The range of values assessed for the number of clusters in the data was between 4 and 15 inclusive.

Each group (cluster) of meals was used to define a generic meal. The nutrient content of a given generic meal was calculated as the mean nutrient content per 100g of the individual

meals that make up that generic meal, as described previously [21]. Seven portion sizes were determined for each generic meal by ordering the individual meals in that generic meal by weight (g) and dividing them into sevenths based on septiles. The median weights of each seventh were assigned as the generic portion sizes for that meal. The nutrient content of the generic meal per 100g consumed by participants was multiplied by the portion size consumed by participants to estimate energy intakes based on the generic meal data.

5.3.3 Step 3. Meal Patterns

5.3.3.1 Datasets

Due to the requirements of the different statistical methods used, the datasets used as input for the analysis varied slightly depending on the method being used. For clustering and PCA, the mean percentage of total energy intake (%TEI) that each participant consumed from each generic meal over the two days was calculated. Each observation represented the mean %TEI consumed from each generic meal, i.e., there was one observation for each participant. Participants were also assigned a value for five binary variables indicating whether or not they consumed each of the five meal types (yes or no for breakfast, lunch, dinner, snacks, beverages). While PAM clustering is designed to be able to handle binary data, numeric values were used in the data for PCA to indicate binary variables.

For LCA, categorical variables, rather than numeric, are required. Therefore, the dataset for input to LCA, consisted of five nominal categorical variables and five binary variables. The nominal variables represented each of the five meal types, with the values being the specific generic meal consumed by the individual for that meal type. Some participants consumed more than one of each meal type; when this occurred with more than 10% of participants then this was included in the LCA analysis as an additional meal type. The nine categorical variables for the male data were therefore breakfast, lunch, dinner, snack 1, snack 2, snack 3, beverage 1, beverage 2, and beverage 3. The eight variables for the female data were breakfast, lunch, dinner, snack 1, snack 2, snack 3, beverage 1, and beverage 2. The binary variables indicated whether or not each of the 5 meal types were consumed. Each observation represented values from a single day for an individual participant. Because it was not possible to calculate a mean for categorical variables, there were two observations for each participant for each meal type and binary variable, representing the two 24-hour recalls they completed.

Two generic meals describing two common beverages (the breakfast meal “coffee and sugar” and the snack meal “water”), were not used in the identification of meal patterns, as described below, because when included, these two meals dominated all meal patterns.

5.3.3.2 Step 3a. Clustering Analysis

PAM clustering was applied to the data described above (%TEI consumed from each generic meal and five binary variables indicating whether or not they consumed each of the five meal types), to group individuals with similar combinations of generic meal intakes, i.e., the identification of meal patterns. The same approach, as described in the previous sections on food groups and generic meals, was used to identify the number of clusters, i.e., 24 indices of model fit were applied and the number of clusters that was most frequently proposed among the indices was chosen [23]. The model fit indices were applied to consider the most appropriate number of clusters between 7 and 13 inclusive, based on previous meal pattern research [5, 6].

5.3.3.3 Step 3b. Principal Component Analysis

PCA was applied to the data described above (%TEI consumed from each generic meal and five binary variables indicating whether or not they consumed each of the five meal types). No rotation was used. Scree plots were used to determine the number of components to retain. To assign individuals to a specific principal component a process previously used in dietary patterns [10, 11] was adapted for use in meal pattern analysis as follows. A score was calculated for each participant for each principal component (principal component score); a higher score indicated that the meals consumed by that participant were similar to the meals that had a high loading on that principal component. The principal component score is the sum of the %TEI values of the generic meals multiplied by their respective loading value on that principal component. Within each principal component participants were ranked based on their score for that component. Each participant was then assigned to the component (meal pattern) for which they had the highest ranking.

5.3.3.4 Step 3b. Latent Class Analysis

The number of classes must be defined prior to conducting the LCA, and this was determined based on two model fit indices, namely the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The range of classes assessed was between 7 and 13. If there was disagreement between AIC and BIC on the optimal number of classes, then the principle of parsimony was applied, i.e., the simpler solution, or the solution with the smaller number of classes, was chosen. For each observation, LCA calculates the probability of it belonging to each latent class (meal pattern). Each observation was assigned to the meal pattern for which it had the highest probability. Each observation represents a single day's dietary intake; therefore, there were two observations

per participant. For the male participants 35.3% had the same meal pattern on both days, and 39.7% of female participants had the same meal pattern on both days.

5.3.4 Comparison of Meal Patterns Across Statistical Methods

A number of approaches were taken to compare the meal patterns that were identified by the statistical methods described above (**Figure 5.1**). Firstly, the number of meal patterns identified by each of the methods was compared. The identification of meal skipping was also compared. For a given meal pattern, a meal type (breakfast, lunch, dinner, snacks, or beverages) was deemed to have been skipped if >98% of the participants belonging to that meal pattern did not consume that meal type.

Finally, the content of the meal patterns derived from each method was compared. To do this each meal pattern was described according to the top two most commonly consumed generic meals at each meal type, and every meal pattern arising from each method was compared with every meal pattern arising from the other two methods. In other words, every meal pattern arising from clustering was compared with every meal pattern arising from PCA (117 comparisons for patterns consumed by males and 117 for females); similarly, clustering was compared with LCA (117 and 104 comparisons for males and females, respectively), and PCA was compared with LCA (81 and 72 comparisons for males and females, respectively). For each comparison, the five different meal types were examined, i.e., breakfast compared with breakfast, lunch compared with lunch etc. For each of these five meal types it was assessed whether the same or different generic meals were consumed for that meal type. The number of matching meal types were counted and could range from zero if none of the meal types matched to five if the same generic meals were consumed at all five meal types for the two meal patterns being compared.

5.4 Results

The results refer to analysis of data from a total of 4332 participants aged 18 years of age or older with dietary recall data for two days. They had a median (interquartile range) age of 52 years (34–64 years) and were comprised of 2068 males and 2264 females (**Table 5.1**). The most commonly consumed meal patterns consumed by males arising from the different methods are given in **Table 5.2** and those consumed by females are given in **Table 5.3**. A full list of all meal patterns is given in Supplementary Tables 5.1–5.6.

5.4.1 Number of Meal Patterns

The clustering method identified 13 meal patterns among the male participants and 13 among the females or a total of 26, the PCA method identified 9 patterns among the males and 9 among the females or a total of 18, and the LCA method identified 9 patterns among the males and 8 among the females or a total of 17.

5.4.2 Meal Skipping

Eighteen of the 26 clustering patterns featured one or more of the five meal types being skipped, i.e., not consumed by >98% of participants in that meal pattern. Twelve of the 17 LCA patterns featured one or more of the five meal types being skipped. No patterns were identified by PCA in which >98% of participants had skipped a meal. The largest proportion of individuals skipping a meal in any one PCA meal pattern was 62.6%. Patterns were identified in which one or more of breakfast, lunch, snacks, or beverages were skipped, but no method identified a pattern in which the dinner meal was skipped by >98% of participants (**Table 5.4**).

5.4.3 Comparison of Meal Types

In the male participant data, on comparing meal patterns arising from clustering (n = 13) with meal patterns arising from PCA (n = 9), there were a total of 117 possible pairs for comparison. There were 10 pairs of meal patterns that had three meal types in common and no pairs had four or five meal types in common (**Figure 5.2**). Breakfasts accounted for 26 matches, lunches accounted for 5, dinners accounted for 1, snacks accounted for 49, and beverages accounted for 72. In the female participant data, on comparing meal patterns arising from clustering (n = 13) with PCA (n = 9) there were a total of 117 possible pairs for comparison. There were 18 pairs of meal patterns that had three meal types in common and two meal patterns that had four meal types in common (**Figure 5.3**). Breakfasts accounted for 42 matches, lunches accounted for 8, dinners accounted for 3, snacks accounted for 58, and beverages accounted for 90.

On comparing clustering meal patterns (n = 13) with LCA patterns (n = 9) in the male data, there were a total of 117 possible pairs for comparison. There were eight pairs of meal patterns that had three meal types in common, and one pair that had four meal types in common (**Figure 5.2**). Breakfasts accounted for 36 matches, lunches accounted for 3, dinners accounted for 5, snacks accounted for 38, and beverages accounted for 48. On comparing clustering patterns (n = 13) with LCA patterns (n = 8) in the female participant data, there were a total of 104 possible pairs for comparison (**Figure 5.3**). There were 17

pairs of meal patterns that had three meal types in common, one pair that had four meal types in common, and one pair that had all five meal types in common. Breakfasts accounted for 45 matches, lunches accounted for 21, dinners accounted for 7, snacks accounted for 38, and beverages accounted for 59.

On comparing PCA meal patterns (n = 9) with LCA patterns (n = 9) in the male participant data, there were a total of 81 possible pairs for comparison. There were eight pairs of meal patterns that had three meal types in common, and two pairs that had four meal types in common. There were no meal patterns that had all five meal types in common (**Figure 5.2**). Breakfasts accounted for 28 matches, lunches accounted for 4, dinners accounted for 0, snacks accounted for 48, and beverages accounted for 45. On comparing PCA (n = 9) and LCA (n = 8) meal patterns in the female participant data, there were a total of 72 possible pairs for comparison. There were 14 pairs of meal patterns that had three meal types in common, but no pair that had either four or five meal types in common (**Figure 5.3**). Breakfasts accounted for 42 matches, lunches accounted for 1, dinners accounted for 5, snacks accounted for 30, and beverages accounted for 45.

5.4.4 Description of the Foods in Meal Patterns Consumed by Males

This section describes the foods that were consumed as part of the most similar meal patterns, i.e., meal patterns that have four or more meal types in common between methods. Clustering pattern 2 (CL2) (Supplementary Table 5.1) and latent class pattern 7 (LC7) (Supplementary Table 5.5) had breakfast, lunch, dinner, and snack meals, but not beverages in common. The common meals characterising these meal patterns were toast- or egg-based breakfast, skipped lunch, pizza with or without pizza for dinner, and fruit or chocolate for snack. There were 116 participants consuming meal pattern CL2 and 342 consuming LC7. Of these, 93 were consuming both CL2 and LC7.

PCA pattern 7 (PC 7) (Supplementary Table 5.3) had breakfast, lunch, snack, and beverage meals, but not dinner, in common with both LC1 and LC9 (Supplementary Table 5.5). The common meals characterising these meal patterns were toast- or egg-based breakfast; rice-, pizza-, or burger-based lunch; fruit or chocolate snack; and water, soft drink, tea or coffee beverage. There were 162 participants consuming meal pattern PC7, 469 consuming pattern LC1, and 484 consuming pattern LC9. Of these, 29 were consuming both PC7 and LC1, 34 were consuming both PC7 and LC9, and 61 consuming both LC1 and LC9.

5.4.5 Description of the Foods in Meal Patterns Consumed by Females

This section describes the foods that were consumed in the most similar meal patterns, i.e., those that have four or more meal types in common between methods. Clustering pattern 2 (CL2) (Supplementary Table 5.2) and PCA pattern 3 (PC3) (Supplementary Table 5.4) had breakfast, lunch, snack, and beverage meal types, but not dinner, in common. The common meals characterising these patterns were egg- or toast-based breakfast, egg sandwich or fruit-based lunch, fruit- or biscuit-based snack, and water or soft drink or coffee or tea as beverage. There were 159 participants consuming meal pattern CL2 and 256 consuming pattern PC3. Of these, 73 were consuming both CL2 and PC3.

CL12 (Supplementary Table 5.2) and LC4 (Supplementary Table 5.6) had breakfast, lunch, snack, and beverage meal types, but not dinner, in common. The common meals characterising these meal patterns were toast- or egg-based breakfast; egg sandwich, rice-based, or pizza lunch; skipped snacks, and water, soft drink, tea or coffee as beverage. There were 106 participants consuming pattern CL12 and 480 consuming pattern LC4. Of these, 88 were consuming both CL12 and LC4.

CL6 (Supplementary Table 5.2) and LC5 (Supplementary Table 5.6) had all five meal types in common. These meal patterns were characterised by toast- or egg-based breakfast; skipped lunch; rice/pasta-based dish or pizza for dinner, with or without a soft drink; fruit or biscuit for snack, and water or soft drink, tea or coffee as beverage. There were 123 participants consuming pattern CL6 and 404 consuming pattern LC5. Of these, there were 116 consuming both CL6 and LC5.

5.5 Discussion

The current study provides a comparison of meal patterns derived from the same dataset using different statistical methods. The number of meal patterns varied among methods, as did the combinations of meals that were consumed in the different meal patterns. Clustering and LCA both identified meal patterns in which meal skipping was present, while PCA did not. A limited number of patterns were identified that were similar between methods with regard to the meals consumed. While previous studies have used these methods on various different datasets [4-6], this is the first study to use multiple methods on the same dataset for comparison.

The number of patterns identified in other studies of content meal patterns has varied from 12 meal patterns using PCA [5], 7 meal patterns using LCA [4], to 11 meal patterns using PCA [6]. This range is comparable to the current study with 9 to 13 meal patterns identified

in males and 8 to 13 meal patterns in females depending on the method. This variability may arise from subjective decisions required of the researchers regarding the different statistical approaches that are available to identify the number of meal patterns (clusters, components, or classes) in a dataset, with different researchers choosing different methods depending on the specifics of their data (e.g., categorical or numeric) and research aims (e.g., requiring mutually exclusive groups or not) [4-6]. In the current study, an objective statistical approach was taken where 24 different model fit indices were used to identify the number of clusters, two were used to identify the number of latent classes, and the scree plot was used to identify the number of principal components. Other methods are available, however, to determine the number of patterns. While the scree plot has been used in other meal pattern studies using PCA [4, 6] other methods include the use of eigenvalues or the interpretability of the patterns [14]. Similarly with clustering and LCA, different combinations of model fit statistics are available to researchers potentially giving rise to differing numbers of patterns, and other methods such as interpretability, quantifying the variance explained, or basing decisions on the sample sizes of the clusters or latent classes [14]. The methods chosen in the current thesis were those that have previously been used in meal pattern research, so that the comparisons would reflect current practices in this research area [4-6, 21]. It is clear from the results that the differing methods can give rise to differing numbers of meal patterns. While a clear description of the method used, as provided in this thesis, is required for reproducibility, further work is required to compare different methods of determining the number of groups present that have not yet been applied to meal patterns.

The current study identified meal patterns that were characterised by skipping of certain meal types. This has also been observed in other meal pattern studies. In a study of 1500 adults in Ireland using LCA, meal patterns were identified in which the lunch meal type was skipped [4]. In the current study, meal patterns characterised by the skipping of one or more of breakfasts, lunches, snacks, and beverages, but not dinners were identified. Two studies using PCA to identify content meal patterns did not refer to meal skipping in their results, so it is not clear whether this was identified or not [5, 6]. In the current study, when using PCA, no meal pattern was identified where meal skipping was a feature. This is in contrast to previous research that has reported meal skipping as a common feature of dietary intakes, estimating that 34% of American adults did not consume breakfast, 19% did not consume lunch, and 7% did not consume dinner [28]. This discrepancy may have arisen due to the ability of PAM clustering and LCA to use categorical input variables, while PCA can only use numeric variables. In the current study, binary input variables were used to indicate the consumption or non-consumption of each meal type. Numbers were used to represent

the binary variables in PCA, but this may not have had the same impact on the outcome as the categorical values used in clustering and LCA.

While the current study is the first to directly compare different methods for analysis of content meal pattern analysis, context can be provided by similar research carried out in dietary patterns. In that field of research, dietary patterns arising from different statistical methods have generally been considered similar, but with some differences. For example, in a study of 1379 adults on the island of Ireland, Hearty and Gibney [15] identified four dietary patterns using PCA and six using clustering. The most similar pattern between the two methods was the one that the authors titled “healthy”. The PCA healthy pattern was characterised by high intakes of nine different food groups and the clustering healthy pattern was characterised by high intakes of eight different food groups; they had six food groups in common [15]. In the current study meal patterns are characterised based on the most commonly consumed generic meals at five different meal types. Depending on the method there were between 17 and 26 meal patterns identified when considering both males and females, yet there were only a small number of patterns that were similar among methods. Across both males and females, this gave rise to between 153 and 234 comparisons of pairs of meal patterns arising from different methods. The most similar pair of meal patterns were identified in the female data using clustering and LCA and they were identical, i.e., they had all five meal types in common. There were six pairs of meal patterns that had four meal types in common: two were identified by clustering and PCA, two were identified by clustering and LCA, and two were identified by PCA and LCA.

A recent review identified only 10 studies that have used data-driven methods to identify meal patterns [1]. When considering that the term meal pattern refers to three different concepts (temporal, content, and context patterns), the number of studies focusing on any one concept is smaller still. While this is still an emerging area of research with regard to the volume of research that has been conducted, it is important that studies consider the variety of methods that are available for this type of analysis and justify their choice of method based on those considerations [14, 18]. Given that PCA identifies correlated meals, it may be most useful in the context of providing descriptive information, e.g., the meal patterns that exist in the population, but as identified in the current study is limited in identifying meal skipping. While PCA groups variables, clustering and LCA group observations, i.e., participants with similar meal intakes. These methods may be more appropriate when distinct groups are required and can be useful in the exploratory stages of research to explore differences among groups [29]. There are also numerous other methods that have not been applied to this area of research but could also be considered

as alternative approaches and include reduced rank regression, treelet transform analysis, and Gaussian graphical models, and random forest with classification tree analysis (18).

Some strengths and limitations are important to note. In the current study, some subjective decisions were made by the researchers during the process; these reflect the decisions that are required in this type of analysis and therefore provide a true reflection of how these methods are applied [14]. Furthermore, all measures of dietary intake assessment are prone to measurement error [30]. However, the use of 24-hour recall data from the NHANES dataset provided information on a large number of food items, allowing the researchers to create food groups in a data-driven manner that was specific to their use in meal pattern analysis [31].

5.6 Conclusion

The use of different statistical techniques to identify meal patterns gives rise to varying numbers of meal patterns. Some similarities exist in the content of the meal patterns identified, but these are limited. Clustering and LCA, but not PCA, identified meal patterns that were characterised by the skipping of one or more meal types. Researchers should consider the features of the various methods, the data type being used, and the research question being asked when choosing a method for meal pattern analysis. Further work is required to examine the impact of the subjective decisions required from researchers in these data-driven methods and to explore other methods that could be used in place of or in combination with methods that have been used to date.

5.7 Tables

Table 5.1: Participant demographics and anthropometry.

	Males	**Females**	**Total**
	n = 2068 (47.7%)	n = 2264 (52.3%)	n = 4332
Age (y)	52.5 (34–65)	54 (34–64)	52 (34–64)
Weight (kg)	85.1 (73.5–100.4)	74.2 (62.4–89.3)	79.7 (67.4–95.7)
Height (cm)	173.7 (168.5–178.9)	160.0 (155.3–164.6)	166.2 (159.1–173.8)
BMI (kg/m^2)	28.4 (24.9–32.9)	29.0 (24.6–34.8)	28.7 (24.8–33.7)

Values are given and median with interquartile range in parentheses.

Table 5.2: The most commonly consumed meal patterns among male participants based on the three different methods of identifying the patterns.

Meal Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
			Clustering		
CL8	Toast (53%) OR Scrambled egg or omelette. (48%)	Soft drink (cola) with chicken burger and chips (35%) OR Sandwich with luncheon ham and mayonnaise (20%)	Meat pizza (35%) OR Rice/pasta dish with water. (16%)	Banana/apple/grapes (60%) OR Chocolate (cookie) (39%)	Water (83%) OR Alcohol (beer) (30%)
			Principal Component Analysis		
PC1	Cereal and milk (23%) OR Toast (20%)	Soft drink (cola) and chicken burger. (37%) OR Soft drink with pizza (30%)	Soft drink (cola), fried chicken sandwich and chips. (75%) OR Soft drink (cola) or iced tea with meat pizza. (14%)	Chocolate (cookie) (36%) OR Banana/apple/grapes (34%)	Water (41%) OR Soft drink or coffee or tea (27%)
			Latent Class Analysis		
LC3	Toast (30%) OR Scrambled egg or omelette. (29%)	Chicken burger (13%) OR Rice-based dish or pizza with water (12%)	Meat pizza (11%) OR Fried chicken. (10%)	Banana/apple/grapes (33%) OR Chocolate (cookie) (19%)	

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

Table 5.3: The most commonly consumed meal patterns among female participants based on the three different methods of identifying the patterns.

Meal Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
			Clustering		
CL9	Scrambled egg or omelette (52%) OR Toast and water (52%)	Rice-based dish or pizza with bread (58%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (32%)	Water with rice/pasta-based dish (35%) OR Soft drink (cola) or iced tea with rice-based dish or pizza. (19%)	Chocolate cookie (39%) OR Ice cream / cheese (32%)	Water (84%) OR Soft drink or coffee or tea (38%)
			Principal Components Analysis		
PC4	Toast and water (42%) OR Cereal and milk (25%)	Fried egg sandwich and water (74%) OR Bread, cheese, luncheon ham, tomatoes, lettuce, mayonnaise, and water. (17%)	Water with rice/pasta-based dish (42%) OR Fried chicken and chips. (26%)	Chocolate cookie (34%) OR Banana/apple/grapes (26%)	Water (64%) OR Soft drink or coffee or tea (18%)
			Latent Class Analysis		
LC2	Toast and water (30%) OR Scrambled egg or omelette (25%)	Rice-based dish or pizza with bread (22%) OR Fried egg sandwich and water (16%)	Water with rice/pasta-based dish (11%) OR Cheese (11%)	Chocolate cookie (24%) OR Banana/apple/grapes (23%)	Water (76%) OR Soft drink or coffee or tea (25%)

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

Table 5.4: The number of meal patterns identified by the different analysis methods and the number of those patterns that in which one or more of the five meal types are not consumed.

	Number of Meal Patterns	**Number of patterns with skipped meal types**					
		Breakfast	Lunch	Dinner	Snacks	Beverages	Total
Clustering							
Male	13	5	1	0	2	4	9
Female	13	4	2	0	1	3	9
Total	26	9	3	0	3	7	18
PCA							
Male	9	0	0	0	0	0	0
Female	9	0	0	0	0	0	0
Total	18	0	0	0	0	0	0
LCA							
Male	9	2	1	0	2	2	6
Female	8	1	1	0	2	3	6
Total	17	3	2	0	4	5	12

A meal type was defined as skipped if >98% of people in that meal pattern did not consume that meal type. LCA, latent class analysis; PCA, principal component analysis.

5.8 Figures

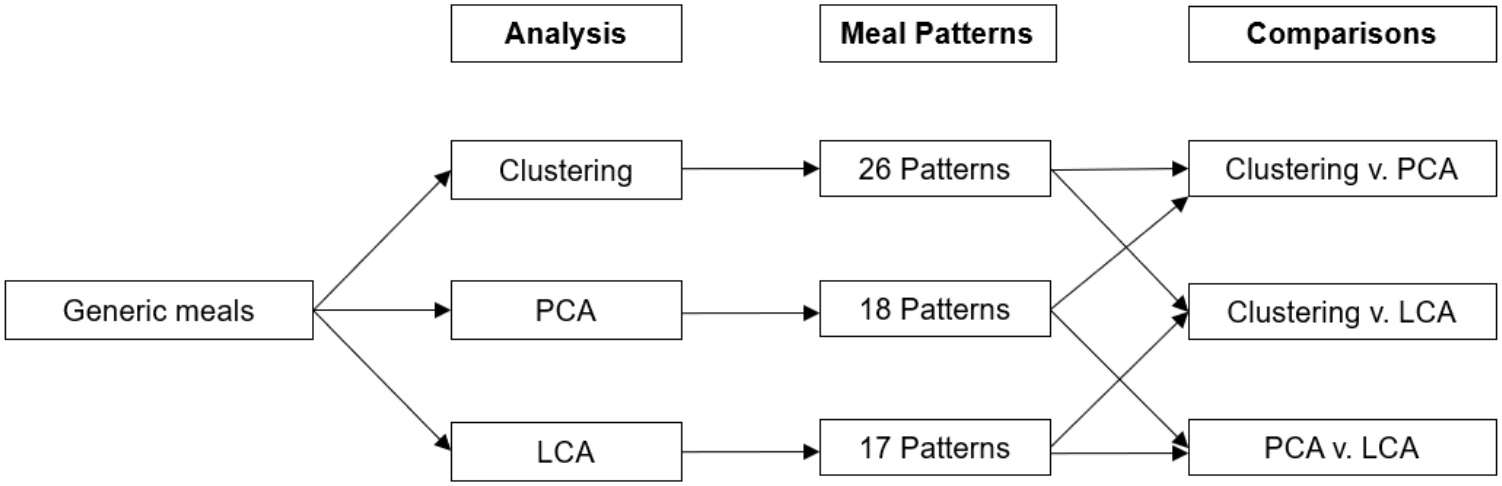


Figure 5.1: Overview of the analysis and comparisons for the different meal pattern analysis techniques.

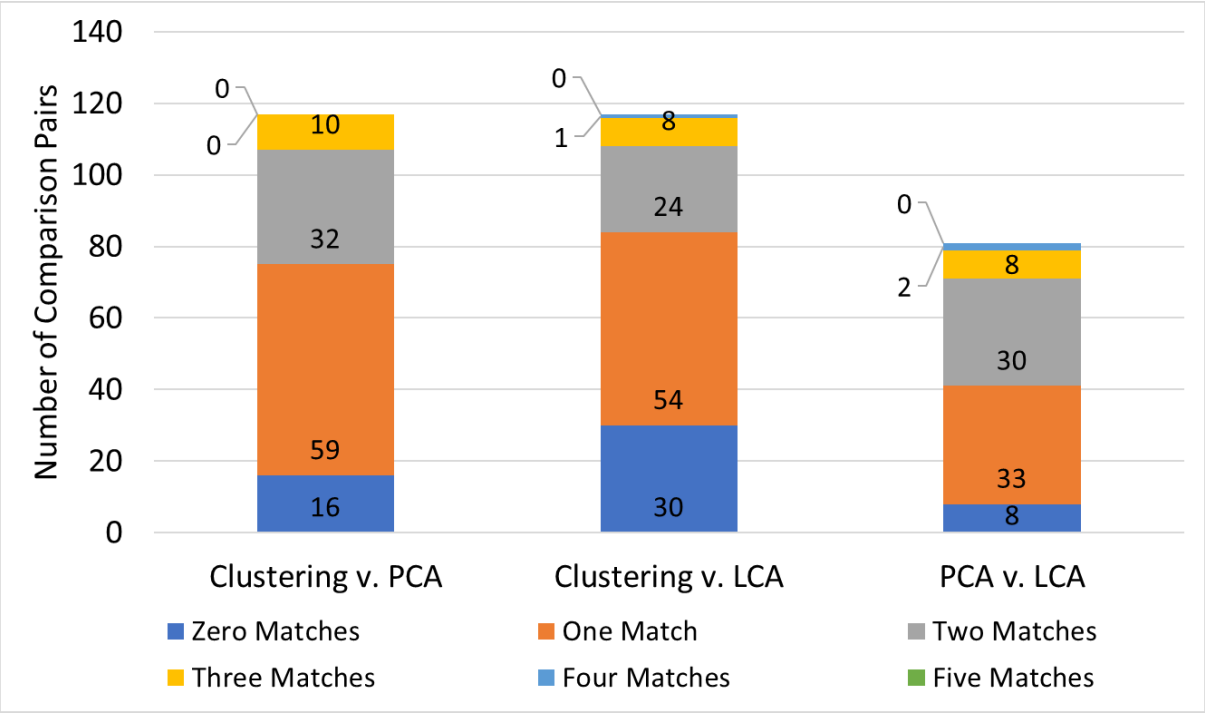


Figure 5.2: The number of matching meal types on comparing the different methods for meal pattern analysis with one another in the data from male participants.

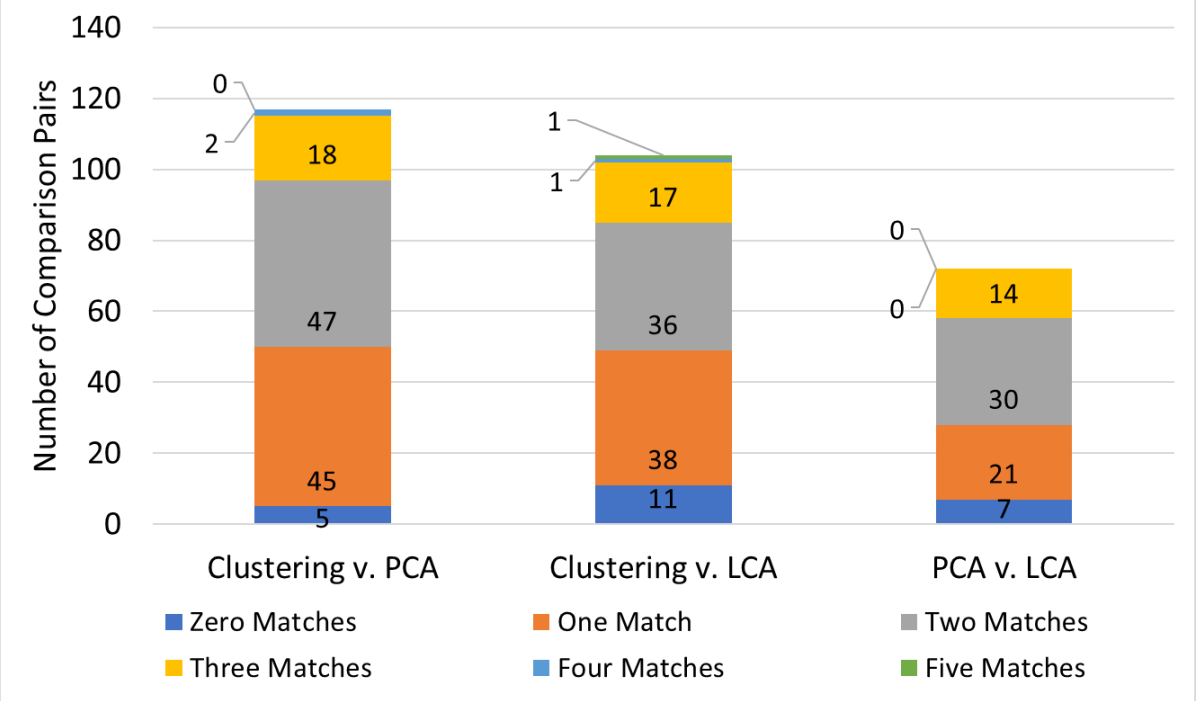


Figure 5.3: The number of matching meal types on comparing the different methods for meal pattern analysis with one another in the data from female participants.

5.9 References

1. O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

2. Leech RM, Worsley A, Timperio A*, et al.* (2015) Understanding Meal Patterns: Definitions, Methodology and Impact on Nutrient Intake and Diet Quality. *Nutr Res Rev*, **28**(1), 1-21.

3. Khanna N, Eicher-Miller HA, Boushey CJ*, et al.* (2011) Temporal Dietary Patterns Using Kernel K-Means Clustering. *IEEE International Symposium on Multimedia*, **2011**, Dana Point, CA, USA, 375-380.

4. Uzhova I, Woolhead C, Timon CM*, et al.* (2018) Generic Meal Patterns Identified by Latent Class Analysis: Insights from NANS (National Adult Nutrition Survey). *Nutrients*, **10**(3), 310.

5. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

6. Murakami K, Livingstone MBE, Sasaki S*, et al.* (2018) Applying a Meal Coding System to 16-D Weighed Dietary Record Data in the Japanese Context: Towards the Development of Simple Meal-Based Dietary Assessment Tools. *J Nutr Sci*, **7**(e29), 1-10.

7. James G, Witten D, Hastie T*, et al.* (2013) An Introduction to Statistical Learning with Applications in R. New York, USA: Springer.

8. Collins LM and Lanza ST (2010) Latent Class and Latent Transition Analysis with Applications in the Social, Behavioral, and Health Sciences, *Wiley Series in Probability and Statistics.* New Jersey, USA: John Wiley & Sons.

9. Lanza ST and Rhoades BL (2013) Latent Class Analysis: An Alternative Perspective on Subgroup Analysis in Prevention and Treatment. *Prev Sci*, **14**(2), 157-168.

10. Sotres-Alvarez D, Herring AH, and Siega-Riz AM (2010) Latent Class Analysis Is Useful to Classify Pregnant Women into Dietary Patterns. *The Journal of Nutrition*, **140**(12), 2253-2259.

11. Knudsen VK, Orozova-Bekkevold IM, Mikkelsen TB*, et al.* (2008) Major Dietary Patterns in Pregnancy and Fetal Growth. *Eur J Clin Nutr*, **62**(4), 463-470.

12. Bull CJ and Northstone K (2016) Childhood Dietary Patterns and Cardiovascular Risk Factors in Adolescence: Results from the Avon Longitudinal Study of Parents and Children (ALSPAC) Cohort. *Public Health Nutr*, **19**(18), 3369-3377.

13. Arabshahi S, Ibiebele TI, Hughes MCB*, et al.* (2017) Dietary Patterns and Weight Change: 15-Year Longitudinal Study in Australian Adults. *Eur J Nutr*, **56**(4), 1455-1465.

14. Wingrove K, Lawrence MA, and McNaughton SA (2022) A Systematic Review of the Methods Used to Assess and Report Dietary Patterns. *Front Nutr*, **9**, 892351.

15. Hearty AP and Gibney MJ (2009) Comparison of Cluster and Principal Component Analysis Techniques to Derive Dietary Patterns in Irish Adults. *Br J Nutr*, **101**(4), 598-608.

16. Meyer J, Doring A, Herder C*, et al.* (2011) Dietary Patterns, Subclinical Inflammation, Incident Coronary Heart Disease and Mortality in Middle-Aged Men from the Monica/Kora Augsburg Cohort Study. *Eur J Clin Nutr*, **65**(7), 800-807.

17. Stricker MD, Onland-Moret NC, Boer JM*, et al.* (2013) Dietary Patterns Derived from Principal Component- and K-Means Cluster Analysis: Long-Term Association with Coronary Heart Disease and Stroke. *Nutr Metab Cardiovasc Dis*, **23**(3), 250-256.

18. Schulz CA, Oluwagbemigun K, and Nothlings U (2021) Advances in Dietary Pattern Analysis in Nutritional Epidemiology. *Eur J Nutr*, **60**(8), 4115-4130.

19. Centers for Disease Control and Prevention (CDC) (2018) National Health and Nutrition Examination Survey (NHANES) Data 2017-2018. Available at: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 (accessed 14 December 2023).

20. U.S. Department of Agriculture (2020) USDA Food and Nutrient Database for Dietary Studies 2017-2018. Available at: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds-download-databases/ (accessed 14 December 2023).

21. O'Hara C, O'Sullivan A, and Gibney ER (2022) A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data. *J Nutr*, **152**(10), 2297-2308.

22. Fulgoni VL, Keast DR, and Drewnowski A (2009) Development and Validation of the Nutrient-Rich Foods Index: A Tool to Measure Nutritional Quality of Foods. *J Nutr*, **139**(8), 1549-1554.

23. Charrad M, Ghazzali N, Boiteau V*, et al.* (2014) Nbclust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, **61**(6), 1-36.

24. Kant AK and Graubard BI (2015) 40-Year Trends in Meal and Snack Eating Behaviors of American Adults. *J Acad Nutr Diet*, **115**(1), 50-63.

25. Mathias KC, Almoosawi S, and Karagounis LG (2017) Protein and Energy Intakes Are Skewed toward the Evening among Children and Adolescents in the United States: NHANES 2013–2014. *J Nutr*, **147**(6), 1160-1166.

26. Barrera CM, Moore LV, Perrine CG*, et al.* (2019) Number of Eating Occasions and Source of Foods and Drinks among Young Children in the United States: NHANES, 2009-2014. *Nutrients*, **11**(4), 897.

27. Sullivan VK, Na M, Proctor DN*, et al.* (2021) Consumption of Dried Fruits Is Associated with Greater Intakes of Underconsumed Nutrients, Higher Total Energy Intakes, and Better Diet Quality in US Adults: A Cross-Sectional Analysis of the National Health and Nutrition Examination Survey, 2007-2016. *J Acad Nutr Diet*, **121**(7), 1258-1272.

28. Zeballos E and Todd JE (2020) The Effects of Skipping a Meal on Daily Energy Intake and Diet Quality. *Public Health Nutr*, **23**(18), 3346-3355.

29. Gleason PM, Boushey CJ, Harris JE*, et al.* (2015) Publishing Nutrition Research: A Review of Multivariate Techniques—Part 3: Data Reduction Methods. *J Acad Nutr Diet*, **115**(7), 1072-1082.

30. Subar AF, Freedman LS, Tooze JA*, et al.* (2015) Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr*, **145**(12), 2639-2645.

31. Newby PK, Sc D, Tucker KL*, et al.* (2004) Empirically Derived Eating Patterns Using Factor or Cluster Analysis: A Review. *Nutr Rev*, **62**(5), 177-203.

5.10 Supplementary Material

Supplementary Table 5.1: Meal patterns identified from male participants using partitioning around the medoids clustering.

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
1 (n = 181)	Scrambled egg or omelette. (67%) OR Cereal and milk (30%)	Rice-based dish or pizza with water (50%) OR Soft drink with pizza (20%)	Soft drink (cola), fried chicken sandwich and chips. (39%) OR Rice/pasta dish with water. (18%)	Chocolate (cookie) (38%) OR Banana/apple/grapes (35%)	Water (86%) OR Soft drink or coffee or tea (36%)
2 (n = 116)	Toast (61%) OR Scrambled egg or omelette. (47%)		Soft drink (cola) or iced tea with meat pizza. (22%) OR Meat pizza (21%)	Banana/apple/grapes (51%) OR Chocolate (cookie) (40%)	Water (66%) OR Soft drink or coffee or tea (32%)
3 (n = 80)		Soft drink with pizza (2%) OR Cheese and chicken sandwich with water (1%)	Soft drink (cola), fried chicken sandwich and chips. (26%) OR Meat pizza (22%)	Chocolate (cookie) (45%) OR Soft drink (cola) (38%)	Water (44%) OR Soft drink or coffee or tea (34%)
4 (n = 211)		Chicken burger (38%) OR Cheese and chicken sandwich with water (22%)	Fried chicken and water. (18%) OR	Banana/apple/grapes (45%) OR Chocolate (cookie) (41%)	Water (69%) OR Soft drink or coffee or tea (36%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
			Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise. (17%)		
5 (n = 176)		Soft drink with pizza (30%) OR Chicken burger (24%)	Fried chicken. (26%) OR Soft drink (cola), fried chicken sandwich and chips. (16%)	Chocolate (cookie) (40%) OR Soft drink (cola) (34%)	
6 (n = 241)	Toast (54%) OR Cereal and milk (36%)	Chicken burger (39%) OR Soft drink (cola) and chicken burger. (20%)	Fried chicken. (39%) OR Soft drink (cola) or iced tea with meat pizza. (26%)	Chocolate (cookie) (42%) OR Ice cream or cheese (35%)	Water (64%) OR Soft drink or coffee or tea (42%)
7 (n = 73)	Toast (47%) OR Scrambled egg or omelette. (41%)	Rice-based dish or pizza with water (34%) OR Soft drink (cola) with chicken burger and chips (19%)	Meat pizza (29%) OR Potatoes and chicken (26%)		
8 (n = 309)	Toast (53%) OR Scrambled egg or omelette. (48%)	Soft drink (cola) with chicken burger and chips (35%) OR	Meat pizza (35%) OR Rice/pasta dish with water. (16%)	Banana/apple/grapes (60%) OR Chocolate (cookie) (39%)	Water (83%) OR Alcohol (beer) (30%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
		Sandwich with luncheon ham and mayonnaise (20%)			
9 (n = 152)	Scrambled egg or omelette. (71%) OR Cereal and milk (26%)	Chicken burger (35%) OR Soft drink (cola) and chicken burger. (19%)	Fried chicken. (28%) OR Fried chicken and water. (18%)	Banana/apple/grapes (50%) OR Chocolate (cookie) (37%)	
10 (n = 140)	Scrambled egg or omelette. (48%) OR Toast (46%)	Soft drink with pizza (21%) OR Sandwich with luncheon ham and mayonnaise (21%)	Meat pizza (19%) OR Chicken and cheese sandwich. (16%)		Water (79%) OR Soft drink or coffee or tea (26%)
11 (n = 164)	Toast (66%) OR Cereal and milk (30%)	Rice-based dish or pizza with water (28%) OR Soft drink with pizza (20%)	Meat pizza (26%) OR Chicken and cheese sandwich. (19%)	Banana/apple/grapes (41%) OR Chocolate (cookie) (34%)	
12 (n = 80)		Soft drink with pizza (19%) OR Soft drink (cola) and chicken burger. (19%)	Fried chicken and water. (30%) OR Soft drink (cola), fried chicken sandwich and chips. (21%)	Ice cream or cheese (2%) OR Coffee with sugar and cookie (1%)	Water (55%) OR Soft drink or coffee or tea (30%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
13 (n = 145)		Soft drink (cola) and chicken burger. (49%) OR Soft drink with pizza (34%)	Soft drink (cola), fried chicken sandwich and chips. (46%) OR Soft drink (cola) or iced tea with meat pizza. (21%)	Chocolate (cookie) (44%) OR Soft drink (cola) (32%)	Water (59%) OR Soft drink or coffee or tea (34%)

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

Supplementary Table 5.2: Meal patterns identified from female participants using partitioning around the medoids clustering.

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
CL1 (n = 176)		Rice-based dish or pizza with bread (49%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes). (22%)	Soft drink (cola) or iced tea with rice-based dish or pizza. (30%) OR Soft drink (cola) or iced tea with chicken burger, lettuce, tomatoes, chips, and mayonnaise. (23%)	Chocolate cookie (34%) OR Crackers/popcorn (25%)	Water (83%) OR Soft drink or coffee or tea (31%)
CL2 (n = 159)	Scrambled egg or omelette (72%) OR Toast and water (29%)	Fried egg sandwich and water (45%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes). (20%)	Fried chicken (67%) OR Soft drink (cola) or iced tea with chicken burger, lettuce, tomatoes, chips, and mayonnaise. (26%)	Chocolate cookie (38%) OR Banana/apple/grapes (28%)	Water (79%) OR Soft drink or coffee or tea (28%)
CL3 (n = 210)	Toast and water (65%) OR Scrambled egg or omelette (45%)	Potatoes and chicken. (41%) OR Rice-based dish or pizza with bread (25%)	Rice/pasta-based dish or bread or pizza (49%) OR Potatoes and chicken (35%)	Chocolate cookie (50%) OR Banana/apple/grapes (37%)	Water (84%) OR Soft drink or coffee or tea (28%)
CL4 (n = 147)		Fried egg sandwich and water (52%) OR Soft drink (cola) and chicken	Fried chicken (33%) OR Water with rice/pasta-based dish (33%)	Chocolate cookie (37%) OR Banana/apple/grapes (31%)	Water (84%) OR Soft drink or coffee or tea (23%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
		sandwich/burger or fried egg sandwich. (48%)			
CL5 (n = 237)	Scrambled egg or omelette (59%) OR Cereal and milk (32%)	Rice-based dish or pizza with bread (46%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (19%)	Soft drink (cola) or iced tea with rice-based dish or pizza. (30%) OR Fried chicken (24%)	Chocolate cookie (44%) OR Banana/apple/grapes (27%)	
CL6 (n = 123)	Toast and water (61%) OR Scrambled egg or omelette (59%)		Rice/pasta-based dish or bread or pizza (26%) OR Soft drink (cola) or iced tea with rice-based dish or pizza. (25%)	Banana/apple/grapes (42%) OR Chocolate cookie (34%)	Water (70%) OR Soft drink or coffee or tea (24%)
CL7 (n= 214)	Cereal and milk (55%) OR Toast and water (51%)	Fried egg sandwich and water (62%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes). (21%)	Water, fried chicken, chips (43%) OR Fried chicken and chips. (30%)	Chocolate cookie (38%) OR Crackers/popcorn (35%)	Water (82%) OR Soft drink or coffee or tea (28%)
CL8 (n = 186)		Rice-based dish or pizza with bread (38%) OR	Fried chicken (23%) OR	Chocolate cookie (41%) OR Banana/apple/grapes (27%)	

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
		Fried egg sandwich and water (31%)	Soft drink (cola) or iced tea with rice-based dish or pizza. (22%)		
CL9 (n = 320)	Scrambled egg or omelette (52%) OR Toast and water (52%)	Rice-based dish or pizza with bread (58%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (32%)	Water with rice/pasta-based dish (35%) OR Soft drink (cola) or iced tea with rice-based dish or pizza. (19%)	Chocolate cookie (39%) OR Ice cream / cheese (32%)	Water (84%) OR Soft drink or coffee or tea (38%)
CL10 (n = 118)		Rice-based dish or pizza with bread (42%) OR Fried egg sandwich and water (22%)	Fried chicken and chips. (62%) OR Rice/pasta-based dish or bread or pizza (47%)	Chocolate cookie (47%) OR Banana/apple/grapes (29%)	Water (70%) OR Soft drink or coffee or tea (33%)
CL11 (n = 194)	Toast and water (70%) OR Cereal and milk (32%)	Fried egg sandwich and water (38%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (32%)	Potatoes and chicken (26%) OR Water with rice/pasta-based dish (22%)	Banana/apple/grapes (32%) OR Chocolate cookie (32%)	
CL12 (n = 106)	Toast and water (54%) OR	Fried egg sandwich and water (42%) OR	Water with rice/pasta-based dish (40%) OR		Water (72%) OR

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
	Scrambled egg or omelette (38%)	Rice-based dish or pizza with bread (35%)	Water and soup (28%)		Soft drink or coffee or tea (25%)
CL13 (n = 74)	Toast and water (3%) OR Scrambled egg or omelette (1%)		Fried chicken (31%) OR Rice/pasta-based dish or bread or pizza (24%)	Ice cream / cheese (31%) OR Chocolate cookie (30%)	Water (53%) OR Soft drink or coffee or tea (20%)

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

Supplementary Table 5.3: Meal patterns identified from male participants using principal component analysis.

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
PC1 (n = 316)	Cereal and milk (23%) OR Toast (20%)	Soft drink (cola) and chicken burger. (37%) OR Soft drink with pizza (30%)	Soft drink (cola), fried chicken sandwich and chips. (75%) OR Soft drink (cola) or iced tea with meat pizza. (14%)	Chocolate (cookie) (36%) OR Banana/apple/grapes (34%)	Water (41%) OR Soft drink or coffee or tea (27%)
PC2 (n = 190)	Cereal and milk (17%) OR Scrambled egg or omelette. (15%)	Chicken burger (28%) OR Fried chicken and sauce/soup (21%)	Fried chicken. (82%) OR Chicken, bread or rice, and sauce. (16%)	Ice cream or cheese (35%) OR Chocolate (cookie) (35%)	Water (37%) OR Soft drink or coffee or tea (29%)
PC3 (n = 285)	Scrambled egg or omelette. (43%) OR Cereal and milk (22%)	Soft drink (cola) and chicken burger. (32%) OR Soft drink (cola), chicken/cheese burger, lettuce, tomatoes, chips, and mayonnaise. (22%)	Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise. (45%) OR Potatoes and chicken (23%)	Chocolate (cookie) (34%) OR Banana/apple/grapes (33%)	Water (51%) OR Soft drink or coffee or tea (25%)
PC4 (n = 253)	Toast (93%) OR Scrambled egg or omelette. (27%)	Chicken burger (25%) OR	Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise. (27%) OR	Banana/apple/grapes (39%) OR	Water (53%) OR Soft drink or coffee or tea (23%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
		Soft drink (cola) with chicken burger and chips (21%)	Potatoes and chicken (21%)	Chocolate (cookie) (28%)	
PC5 (n = 280)	Scrambled egg or omelette. (48%) OR Cereal and milk (18%)	Chicken burger (62%) OR Cheese and chicken sandwich with water (19%)	Fried chicken and water. (36%) OR Rice/pasta dish with water. (26%)	Chocolate (cookie) (40%) OR Banana/apple/grapes (36%)	Water (58%) OR Soft drink or coffee or tea (19%)
PC6 (n = 214)	Toast (29%) OR Scrambled egg or omelette. (26%)	Sandwich with luncheon ham and mayonnaise (24%) OR Rice-based dish or pizza with water (22%)	Meat pizza (95%) OR Bean-based salad and fruit/fruit juice. (10%)	Banana/apple/grapes (40%) OR Chocolate (cookie) (34%)	Water (57%) OR Soft drink or coffee or tea (22%)
PC7 (n = 162)	Toast (38%) OR Scrambled egg or omelette. (26%)	Chicken burger (30%) OR Rice-based dish or pizza with water (18%)	Chicken, bread or rice, and sauce. (93%) OR Soft drink (cola), fried chicken sandwich and chips. (15%)	Chocolate (cookie) (37%) OR Banana/apple/grapes (36%)	Water (59%) OR Soft drink or coffee or tea (23%)
PC8 (n = 204)	Scrambled egg or omelette. (57%) OR Toast (33%)	Rice-based dish or pizza with water (75%) OR	Fried chicken. (33%) OR Rice/pasta dish with water. (22%)	Banana/apple/grapes (41%) OR Chocolate (cookie) (29%)	Water (59%) OR Soft drink or coffee or tea (23%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
		Sandwich with luncheon ham and mayonnaise (21%)			
PC9 (n = 164)	Toast (30%) OR Cereal and milk (16%)	Rice-based dish or pizza with water (33%) OR Soft drink with pizza (24%)	Soft drink (cola) or iced tea with meat pizza. (88%) OR Soft drink with cookie or cake (15%)	Chocolate (cookie) (32%) OR Banana/apple/grapes (30%)	Water (50%) OR Soft drink or coffee or tea (28%)

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

Supplementary Table 5.4: Meal patterns identified from female participants using principal component analysis.

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
PC1 (n = 305)	Toast and water (37%) OR Scrambled egg or omelette (29%)	Rice-based dish or pizza with bread (100%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes). (15%)	Rice/pasta-based dish or bread or pizza (29%) OR Water with rice/pasta-based dish (23%)	Chocolate cookie (39%) OR Banana/apple/grapes (32%)	Water (58%) OR Soft drink or coffee or tea (23%)
PC2 (n = 175)	Scrambled egg or omelette (31%) OR Toast and water (14%)	Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (49%) OR Rice-based dish or pizza with bread (33%)	Fried chicken (78%) OR Fried chicken and chips. (37%)	Chocolate cookie (49%) OR Banana/apple/grapes (26%)	Water (44%) OR Soft drink or coffee or tea (24%)
PC3 (n = 256)	Scrambled egg or omelette (89%) OR Toast and water (40%)	Fried egg sandwich and water (35%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes). (25%)	Fried chicken (43%) OR Water with rice/pasta-based dish (23%)	Banana/apple/grapes (32%) OR Chocolate cookie (32%)	Water (61%) OR Soft drink or coffee or tea (20%)
PC4 (n = 326)	Toast and water (42%) OR Cereal and milk (25%)	Fried egg sandwich and water (74%) OR Bread, cheese, luncheon ham,	Water with rice/pasta-based dish (42%) OR Fried chicken and chips. (26%)	Chocolate cookie (34%) OR Banana/apple/grapes (26%)	Water (64%) OR Soft drink or coffee or tea (18%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
		tomatoes, lettuce, mayonnaise, and water. (17%)			
PC5 (n = 232)	Scrambled egg or omelette (39%) OR Toast and water (25%)	Chicken burger with lettuce, tomatoes, chips, mayonnaise, and soft drink (cola) or iced tea. (34%) OR Rice-based dish or pizza with bread (25%)	Soft drink (cola) or iced tea with chicken burger, lettuce, tomatoes, chips, and mayonnaise. (86%) OR Soft drink (cola) or iced tea with rice-based dish or pizza. (17%)	Chocolate cookie (38%) OR Ice cream / cheese (24%)	Water (54%) OR Soft drink or coffee or tea (23%)
PC6 (n = 230)	Toast and water (48%) OR Cereal and milk (17%)	Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (81%) OR Fried egg sandwich and water (25%)	Rice/pasta-based dish or bread or pizza (43%) OR Water with rice/pasta-based dish (21%)	Chocolate cookie (38%) OR Soft drink (cola) (25%)	Water (60%) OR Soft drink or coffee or tea (26%)
PC7 (n = 201)	Toast and water (50%) OR Scrambled egg or omelette (42%)	Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (29%) OR Bread, cheese, luncheon ham,	Soft drink (cola) or iced tea with rice-based dish or pizza. (76%) OR Fried chicken and chips. (19%)	Banana/apple/grapes (30%) OR Chocolate cookie (30%)	Water (58%) OR Soft drink or coffee or tea (18%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
		tomatoes, lettuce, mayonnaise, and water. (26%)			
PC8 (n = 226)	Toast and water (21%) OR Cereal and milk (19%)	Fruit juice (apple) or raw fruit (apple, banana, grapes). (34%) OR Chicken burger with lettuce, tomatoes, chips, mayonnaise, and soft drink (cola) or iced tea. (23%)	Rice/pasta-based dish or bread or pizza (55%) OR Soft drink (cola) or iced tea with rice-based dish or pizza. (40%)	Chocolate cookie (41%) OR Banana/apple/grapes (30%)	Water (50%) OR Soft drink or coffee or tea (20%)
PC9 (n = 313)	Toast and water (33%) OR Scrambled egg or omelette (29%)	Potatoes and chicken. (29%) OR Rice-based dish or pizza with bread (25%)	Potatoes and chicken (70%) OR Water, fried chicken, chips (18%)	Chocolate cookie (38%) OR Banana/apple/grapes (33%)	Water (58%) OR Soft drink or coffee or tea (22%)

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

Supplementary Table 5.5: Meal patterns identified from male participants using latent class analysis.

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
LC1 (n = 469)	Toast (20%) OR Scrambled egg or omelette. (18%)	Rice-based dish or pizza with water (14%) OR Chicken burger (10%)	Meat pizza (11%) OR Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise. (10%)	Banana/apple/grapes (31%) OR Chocolate (cookie) (28%)	Water (84%) OR Soft drink or coffee or tea (41%)
LC2 (n = 244)		Soft drink with pizza (12%) OR Fried chicken and sauce/soup (9%)	Soft drink (cola), fried chicken sandwich and chips. (15%) OR Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise. (13%)	Ice cream or cheese (9%) OR Soft drink (cola) (6%)	Water (54%) OR Soft drink or coffee or tea (39%)
LC3 (n = 666)	Toast (30%) OR Scrambled egg or omelette. (29%)	Chicken burger (13%) OR Rice-based dish or pizza with water (12%)	Meat pizza (11%) OR Fried chicken. (10%)	Banana/apple/grapes (33%) OR Chocolate (cookie) (19%)	
LC4 (n = 567)		Chicken burger (15%) OR Soft drink with pizza (13%)	Fried chicken. (12%) OR Soft drink (cola), fried chicken sandwich and chips. (11%)	Chocolate (cookie) (27%) OR Banana/apple/grapes (24%)	Water (16%) OR Soft drink or coffee or tea (6%)

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
LC5 (n = 437)	Toast (31%) OR Scrambled egg or omelette. (30%)	Rice-based dish or pizza with water (14%) OR Chicken burger (11%)	Rice/pasta dish with water. (10%) OR Meat pizza (9%)	Ice cream or cheese (1%)	Water (74%) OR Soft drink or coffee or tea (30%)
LC6 (n = 577)	Scrambled egg or omelette. (19%) OR Toast (18%)	Chicken burger (13%) OR Soft drink with pizza (13%)	Meat pizza (11%) OR Soft drink (cola), fried chicken sandwich and chips. (11%)		
LC7 (n = 342)	Toast (23%) OR Scrambled egg or omelette. (18%)		Meat pizza (11%) OR Soft drink (cola) or iced tea with meat pizza. (11%)	Banana/apple/grapes (23%) OR Chocolate (cookie) (15%)	Water (15%) OR Alcohol (beer) (6%)
LC8 (n = 350)	Scrambled egg or omelette. (19%) OR Toast (18%)	Chicken burger (12%) OR Sandwich with luncheon ham and mayonnaise (11%)	Meat pizza (10%) OR Fried chicken. (9%)	Chocolate (cookie) (66%) OR Banana/apple/grapes (51%)	Water (16%) OR Alcohol (beer) (11%)
LC9 (n = 484)	Toast (28%) OR Scrambled egg or omelette. (27%)	Chicken burger (13%) OR Rice-based dish or pizza with water (12%)	Rice/pasta dish with water. (11%) OR Meat pizza (10%)	Banana/apple/grapes (31%) OR Chocolate (cookie) (26%)	Water (71%) OR Soft drink or coffee or tea (17%)

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

Supplementary Table 5.6: Meal patterns identified from female participants using latent class analysis.

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
LC1 (n = 429)	Toast and water (22%) OR Scrambled egg or omelette (22%)	Rice-based dish or pizza with bread (19%) OR Fried egg sandwich and water (18%)	Fried chicken (13%) OR Water with rice/pasta-based dish (12%)		
LC2 (n = 819)	Toast and water (30%) OR Scrambled egg or omelette (25%)	Rice-based dish or pizza with bread (22%) OR Fried egg sandwich and water (16%)	Water with rice/pasta-based dish (11%) OR Cheese (11%)	Chocolate cookie (24%) OR Banana/apple/grapes (23%)	Water (76%) OR Soft drink or coffee or tea (25%)
LC3 (n = 439)	Toast and water (24%) OR Scrambled egg or omelette (22%)	Rice-based dish or pizza with bread (16%) OR Fried egg sandwich and water (15%)	Water with rice/pasta-based dish (11%) OR Fried chicken and chips. (10%)	Chocolate cookie (52%) OR Ice cream / cheese (34%)	Water (73%) OR Soft drink or coffee or tea (25%)
LC4 (n = 480)	Toast and water (21%) OR Scrambled egg or omelette (20%)	Fried egg sandwich and water (23%) OR Rice-based dish or pizza with bread (20%)	Water with rice/pasta-based dish (13%) OR Fried chicken (12%)		Water (75%) OR Soft drink or coffee or tea (26%)
LC5 (n = 404)	Toast and water (26%) OR		Soft drink (cola) or iced tea with rice-based dish or pizza. (14%)	Banana/apple/grapes (18%) OR	Water (35%) OR

Pattern	Breakfast	Lunch	Dinner	Snacks	Beverages
	Scrambled egg or omelette (22%)		OR Rice/pasta-based dish or bread or pizza (12%)	Chocolate cookie (12%)	Soft drink or coffee or tea (11%)
LC6 (n = 748)		Rice-based dish or pizza with bread (22%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich. (17%)	Soft drink (cola) or iced tea with rice-based dish or pizza. (12%) OR Fried chicken (11%)	Chocolate cookie (22%) OR Banana/apple/grapes (20%)	Water (37%) OR Soft drink or coffee or tea (11%)
LC7 (n = 435)	Toast and water (17%) OR Scrambled egg or omelette (15%)	Rice-based dish or pizza with bread (20%) OR Fried egg sandwich and water (13%)	Rice/pasta-based dish or bread or pizza (13%) OR Fried chicken (13%)	Chocolate cookie (53%) OR Ice cream / cheese (38%)	
LC8 (n = 774)	Toast and water (31%) OR Scrambled egg or omelette (26%)	Fried egg sandwich and water (18%) OR Rice-based dish or pizza with bread (18%)	Soft drink (cola) or iced tea with rice-based dish or pizza. (12%) OR Rice/pasta-based dish or bread or pizza (11%)	Banana/apple/grapes (24%) OR Chocolate cookie (18%)	

Values in parentheses are the percentage of participants in that meal pattern consuming that generic meal.

CHAPTER 6

An Exploratory Analysis of Meal Patterns and their Relationships with Health: Findings from the NHANES Cohort

6.1 Abstract

Meal pattern research can identify aspects of dietary intakes not evident from food- or nutrient-based research. Previous research indicates that meal patterns impact on health; however, such research often focuses on single constructs (temporal, context, or content) of meal patterns, missing the complexity of examining temporal and content aspects together. The objective of this study was to identify meal patterns consumed by American adults, combining both temporal and content aspects, and explore associations between meal patterns and health. This study analysed data from NHANES 2017–2018. Meal patterns were identified from 24-hour recall data using a three-step approach. Firstly, *k*-means clustering was used to identify food groups; secondly, partitioning around the medoids (PAM) clustering was used to identify common combinations of food groups consumed within meals (generic meals); and thirdly, PAM clustering was used to identify the combinations of generic meals consumed (meal patterns). Summary statistics and parallel plots were generated to compare dietary intake and health variables among meal patterns. A total of 26 meal patterns were identified. Patterns with a skipped breakfast were more common than skipping of other meal types. In the male data, those consuming the meal pattern characterised by breakfast skipping, soft drink and burger or fried meat at lunch and dinner, chocolate or soft drink as snack had the highest daily energy intake and the second lowest dietary quality. In the female data, those consuming the meal pattern characterised by an egg-based breakfast, sandwich or fruit at lunch, fried chicken at dinner, chocolate or fruit as snacks, had the highest median energy intakes. No differences were observed among the meal patterns for the health variables. In conclusion, clustering is a feasible approach to identify combined aspects of temporal and content meal patterns and identify differences in nutrient quality, but not health outcomes among meal patterns.

6.2 Introduction

An increasing number of studies have emerged in recent years investigating meal patterns in nutrition [1]. This area of research provides insights into dietary intakes not evident from more traditional approaches that focus on foods and nutrients [1, 2]. Temporal meal patterns provide information regarding the timing and distribution of meals, context meal patterns give information about the location, the presence or absence of others, or activities during meals, while content meal patterns provide information about the combination of foods in meals and the combinations of meals over a day [1].

Some countries now consider the specific role of meals in the development of national dietary guidelines to complement what is already known about foods and nutrients, such as those in the United States of America (USA) [3] and Brazil [4]. While acknowledging the limited literature in the area and the need for further research, these guidelines refer to the timing and distribution of meals, the contexts in which meals are consumed, and the different combinations of foods in meals and meals over a day. Furthermore, in the area of technology and personalised nutrition, recommender systems must ultimately be able to provide recommendations for meal intakes as opposed to food intakes, to support optimal behavioural change [5]. However, before these can be fully realised, meal pattern research is required to provide the evidence upon which meal-based nutrition recommendations can be made [3, 4].

Advanced statistical approaches are required to identify content meal patterns due to the large numbers of possible combinations of foods and meals that could make up an individual's or population's dietary intakes. While research remains limited in the area, a number of statistical techniques have been used to date, including principal components analysis (PCA) in Irish and Japanese cohorts [6, 7] and latent class analysis (LCA) [8] and clustering [9] in Irish cohorts. In the USA, temporal meal patterns, but not content meal patterns, have been identified using clustering [10, 11]. In that research, the pattern associated with the highest Healthy Eating Index (HEI) score was characterised by three relatively evenly spread meals throughout the day. The pattern associated with the lowest HEI score was one characterised by those likely to have five meals in the day, often consumed late in the day or during the night. The other two patterns identified are characterised by those likely to have their main intake later in the day with their second largest intake earlier in the day, or vice versa [11]. No information is available, however, detailing what content meal patterns are consumed in the USA, how these may relate to temporal patterns, or whether there are associations between those patterns and health.

Linking specific meal patterns to health outcomes is important in the development of meal-based dietary guidelines or personalised recommender systems. For example, this could allow recommender systems to suggest specific meals or combinations of meals (meal patterns) for an individual to consume based on their potential health impact instead of suggesting individual foods or recipes [5]. Some limited associations between health outcomes and content meal patterns have also been identified [8, 12]. For example, research in an Irish cohort identified that those consuming a meal pattern characterised by "a cooked breakfast, skipped light meal, and protein-carbohydrate main meal" were likely to have higher diastolic blood pressure compared with those consuming the meal pattern characterised by "cereal and/or toast for breakfast, sandwich for light meal and protein-carbohydrate or just protein main meal" [8]. Further work is required to explore the potential relationships between content meal patterns and health outcomes in various populations to inform meal-based dietary recommendations.

Therefore, the aim of the current study was to identify meal patterns that exist in the USA, combining both temporal and content aspects, and explore whether associations exist between those meal patterns and health.

6.3 Methods

6.3.1 Dietary Intake Data

The current study analysed previously collected data from non-institutionalised civilian residents of the USA from the 2017-2018 NHANES. NHANES was approved by the National Center for Health Statistics Ethics Review Board [13]. The analysis reported in the current study includes adults (18–80 years) who had completed two 24-hour recalls; these were conducted by trained interviewers in person for the first recalls and over telephone for the second recalls using a standardised process. The nutrient compositions of the foods reported in the 24-hour recalls are from the food and nutrient database for dietary studies (FNDDS) 2017-2018 [14]. Nutrient intakes are reported with the exclusion of intakes from dietary supplements.

6.3.2 Meal Pattern Analysis

There were three broad steps involved in identifying meal patterns, that is, the clustering of foods into food groups, the clustering of meals into generic meals, and finally the clustering of generic meals into meal patterns (**Figure 6.1**). The first two steps were based on a previously developed method, as described in detail by O'Hara *et al.* [9], adapted for use

with the NHANES data in the current study. All steps are described in the following subsections.

6.3.2.1 Step 1. Food Groups

Foods listed within the reported dietary recalls were grouped based on the five groups used in the USA dietary guidelines: fruit; vegetables; grains; protein foods; dairy. Foods that did not belong in any of those groups were assigned to one of the following five groups: foods high in fat, sugar, or salt; fats and oils; non-alcoholic beverages; alcoholic beverages; and other, with 10 food groups in total. *K*-means clustering was applied to each of the 10 food groups separately to create subgroups containing foods that were most similar to each other with regard to nutrients listed within the Nutrient Rich Foods (NRF9.3) index, namely protein, fibre, vitamin A, vitamin C, vitamin E, calcium, iron, magnesium, potassium, saturated fatty acids, total sugar, and sodium [15]. The nutrient contents of the foods were expressed in units per 100kcal and were z-standardised before clustering. The number of clusters was determined by applying 24 indices and choosing the number of clusters that was most frequently proposed among the indices [16] (Supplementary table 6.1). The range of values assessed for the number of clusters in the data was between two and eight inclusive. The clustering of the non-alcoholic beverages group resulted in unbalanced clusters with very small numbers of foods in some clusters, so this group was split into energy containing and energy free non-alcoholic beverages instead of clustering. Two of the original food groups (alcoholic beverages and other) were not further split due to the small number of foods in each of these groups. Fats and oils were not further split, instead their original three subgroups from the FNDDS were kept: animal fats; margarines; and dressings and vegetable oils [14]. This gave rise to a total of 23 food groups. The food groups and their subsequent subgroups after *k*-means clustering are summarised in **Table 6.1**.

6.3.2.2 Step 2. Generic Meals

After identifying the food groups described above, meals were then defined based on the food groups they contained, i.e., the individual foods reported by the participants were replaced by the food group to which that food belonged. The meal intake data were split into two groups based on sex and from this point analysed separately for males and females. Meals were also split according to meal type. The original NHANES data used in this analysis included 18 different meal types. These meal types were condensed to five for the purpose of this analysis: breakfast (comprising breakfast, desayuno, and amuerzo), lunch (lunch, comida, and brunch), dinner (dinner, supper, and cena), snack (snack, merienda, entre comida, botana, bocadillo, tentempie, and extended consumption), and

beverages (drink and bebida). These categorisations were based on previous methods from the literature [17-20].

Partitioning around the medoids (PAM) clustering was applied separately to each meal type, allowing similar meals in each meal type to be grouped. Similarity was based on 24 variables: 23 binary variables, each one representing one of the food groups and indicating whether that food group was present or not in the meal, and a single numeric value, namely the NRF9.3 index score [15] for each meal, which was used as an indication of the nutritional quality of the meal. As with *k*-means clustering, PAM clustering also requires that the number of clusters is specified before clustering. The same approach described in the previous section was taken. The range of values assessed for the number of clusters in the data was between 4 and 15 inclusive.

The nutrient content of a given generic meal was then calculated as the mean nutrient content per 100g of the individual meals that make up that generic meal, as described previously [9]. To incorporate a range of portion sizes, seven portion sizes were determined for each generic meal by ordering the individual meals in that generic meal by weight (g) and dividing them into sevenths based on septiles. The median weights of each seventh were assigned as the generic portion sizes for that meal. The nutrient content of the generic meal per 100g consumed by participants was multiplied by the portion size to estimate energy intakes based on the generic meals.

6.3.2.3 Step 3. Meal Patterns

Meal patterns refer to the combinations of generic meals that are consumed over the course of the day. For this final step, the mean percentage of total energy intake (%TEI) that each participant consumed from each generic meal over the two days was calculated. Participants were also assigned a value for five binary variables indicating whether or not they consumed each of the five meal types (breakfast, lunch, dinner, snacks, beverages). PAM clustering was applied to the data using the %TEI values for generic meals and the binary variables for each meal type as input variables. This allowed the grouping of individuals with similar combinations of generic meal intakes, i.e., the identification of meal patterns.

The same approach, as described in the previous clustering sections, was used to identify the number of clusters. The range of values assessed for the number of clusters in the data was between 7 and 13 inclusive. Two generic meals from the male data and two meals from the female data were not used in the clustering, as when included these 2 meals

dominated all meal patterns. These 2 meals were the breakfast meal “coffee and sugar” and the snack meal “water”.

6.3.3 Anthropometry and Clinical Measurements

Height was measured to the nearest 0.1cm using a stadiometer with a fixed vertical backboard and an adjustable head piece. Participants’ weight was measured to the nearest 0.1kg in kilograms using a calibrated digital scale. Waist (abdominal) circumference measurements were taken on participants’ bare abdomen. The measurements were taken at the point of the uppermost lateral border of the right ilium of the pelvis, and the measuring tape extended horizontally around the waist from that point. Hip (buttocks) circumference measurements were taken around the maximum protuberance of the buttocks when viewed from participants’ right side. Both measurements were taken to the nearest 0.1cm at the end of a normal exhalation from the participant. Percentage body fat was estimated using dual energy x-ray absorptiometry (DXA) using the Hologic Discovery A fan beam x-ray bone densitometer [13].

Prior to blood pressure measurement, participants rested quietly in a seated position for five minutes. Three blood pressure measurements were then taken with one minute gap between each one. The mean blood pressure value of the three measurements was used in the current study. Measurements were taken on the bare upper right arm unless there were any specific conditions that did not allow for this, in which case the left arm was used. The Omron IntelliSense Blood Pressure Monitor (Model: HEM-907XL) was used to take the blood pressure measurements.

For biochemical measurements, a fasting questionnaire was administered prior to venipuncture to confirm fasting status. For fasting samples, participants are asked to fast for nine hours. The vacutainer blood collection device and blood tubes were used to draw the blood. The samples used for plasma glucose were collected in sodium fluoride tubes and measured with a UV in vitro test using the Roche/Hitachi Cobas C System (c311). The samples used for serum total cholesterol and HDL were collected in EDTA or heparin tubes and measured photometrically using the Roche/Hitachi Cobas 6000 Chemistry Analyser. LDL Cholesterol was not measured directly but calculated using the Friedewald equation [13].

6.3.4 Statistical Analysis

All analysis was carried out using R version 4.2.2 [21] in the RStudio integrated development environment (version 2022.07.2+576) [22]. Medians and 95% confidence intervals (CI) of

nutrient intakes and health parameters were calculated separately for those belonging to each of the different meal patterns. Z-scores of these values were calculated and plotted on parallel plots to compare multiple variables of differing units among the different meal patterns. Given the exploratory nature of the current study, confirmatory statistics and *p* values were not deemed appropriate in the analysis. Confidence intervals were calculated to quantify the uncertainty around the medians. Bonferroni correction for multiple comparisons was used when calculating confidence intervals. Given the right-skewed distribution of the data, the bias corrected and accelerated bootstrap method for calculating confidence intervals was used [23].

Basal metabolic rate was estimated using the Oxford equations [24]. The Goldberg equation was used to calculate cut-offs to identify those with implausible reports of energy intake. Those with reported energy intakes resulting in an energy intake to basal metabolic rate ratio of less than 0.96 or greater than or equal to 2.49 were deemed to be implausible [25]. The inclusion of all participants regardless of the plausibility of their reported energy intakes did not impact on the identification of generic meals, therefore all participants were included in that analysis. Those with implausible reports of energy intakes, however, were not evenly distributed among the various meal patterns, with the proportion of implausible reporters of energy intake ranging from 18.2 to 49.3% in the male data and 18.8 to 58.1% in the female data depending on the meal pattern. Therefore, for the purposes of comparing energy and nutrient intakes and health parameters among meal patterns, those with implausible reports of energy intakes were excluded.

While data on nutrient intake and age was available for all participants, only a portion of participants with plausible reports of energy intake had data for BMI (n = 1415 for males and n = 1508 for females), WHR (n = 1389 for males and n = 1462 for females), percentage body fat by DXA (n = 661 for males and n = 732 for females), systolic and diastolic blood pressure (n = 1336 for males and n = 1385 for females), fasting glucose (n = 678 for males and n = 721 for females), HDL and total cholesterol (n =1360 for males and 1449 for females), and LDL cholesterol (n = 656 for males and n = 705 for females). The summary data reported excludes those with missing data for the variable in question.

6.4 Results

6.4.1 Generic Meals and Meal Patterns

A total of 4332 participants aged 18 years of age or older with recall data for two days were included in the identification of generic meals and meal patterns. They had a median

(interquartile range) age of 52 years (34–64 years) and were comprised of 2068 males and 2264 females (Supplementary Table 6.2).

In total, 42 generic meals were identified in the male data which included 4 breakfast meals, 10 lunch meals, 15 dinner meals, 9 snack meals, and 4 beverage meals. In the female data 41 generic meals were identified which included 4 breakfast meals, 7 lunch meals, 12 dinner meals, 14 snack meals, and 4 beverage meals (Supplementary Table 6.3).

In the male data 13 meal patterns were identified; three included all five of the meal types (breakfast, lunch, dinner, snacks, and beverages), two excluded only breakfast, one excluded only lunch, no meal patterns excluded dinner, one excluded only snacks, two excluded only beverages, one excluded breakfast and lunch, one excluded breakfast and snacks, one excluded breakfast and beverages, and one excluded snacks and beverages. Food-based descriptions for the three most frequently consumed meal patterns are given in **Table 6.2**, with the remainder given in Supplementary Table 6.4. The most frequently consumed meal pattern (meal pattern 8) is characterised by toast or egg at breakfast, chicken burger with soft drink at lunch, meat pizza at dinner, fruit as snacks, and water as a beverage.

In the female data 13 meal patterns were also identified; four included all five of the meal types assessed, three excluded only breakfast, one excluded only lunch, no meal patterns excluded dinner, one meal pattern excluded only snacks, two excluded only beverages, one excluded both breakfast and lunch, and one excluded both breakfast and beverages. Food-based descriptions for the three most frequently consumed meal patterns are given in **Table 6.3**, with the remainder given in Supplementary Table 6.5. The most frequently consumed meal pattern (meal pattern 9) by females was likely to include egg, toast, and water at breakfast, a rice-based dish or pizza at lunch, a rice- or pasta-based dish with water at dinner, cookies, ice-cream or cheese as snacks, and water as a beverage.

6.4.2 Nutrient Quality and Health

The values given in this section of the results, including subsections, refer only to those with plausible reports of energy intake. This includes 2964 participants with a median (interquartile range) age of 52 years (34–64 years), comprising of 1432 males and 1532 females (**Table 6.4**). The values given below are given as median and 95% CI. For age and the health variables, no differences among the meal patterns were evident, i.e., the median for each meal pattern was within the range of the CIs for the other meal patterns.

6.4.2.1 Male Meal Patterns

Median z-scores for selected nutrients and health variables across the male meal patterns are presented in **Figure 6.2**, and confidence intervals are presented in Supplementary Figure 6.1. Median and 95% CIs for daily energy intakes of 2634kcal/d (2387–2887kcal/d) were highest for meal pattern 13, which also had the second lowest median NRF9.3 score of 6.6 (3.14–9.11)., the highest median total sugar intake of 136 g/day (111–159g/day), the second highest median fat intake of 108.7g/d (92.6–121.7g/d), the highest median monounsaturated fat (MUFA) intake of 36.5g/d (33.0–42.3g/d), polyunsaturated fat (PUFA) intake of 24.5g/d (21.5–28.7g/d), and sodium intake of 4226mg/d (3654–4777mg/d) (Supplementary Table 6.6)(**Figure 6.2**). This meal pattern is characterised by breakfast skipping, soft drink and burger or fried meat at lunch and dinner, chocolate or soft drink as snack, and water as a beverage.

The lowest median NRF9.3 score, indicating the lowest nutritional quality, observed in the male data was 5.0 (0.9–9.7) in meal pattern 3 consisting of skipped breakfast and lunch, fried chicken sandwich and chips with soft drink or a meat-based pizza for dinner, chocolate or soft drink for snacks, and water or soft drink as an additional beverage. This meal pattern was also associated with the second highest median sugar intake of 131g/d (99–187g/d), the lowest median PUFA intake 17.6g/d (13.7–26.3g/d), the second lowest median fibre intake of 14.2g/d (10.0–17.3g/d) and sodium intake of 3288mg/d (2545–4066mg/d), and the lowest median calcium intake of 812mg/d (569–1223mg/d) and iron intake of 12.5mg/d (9.8–15.2mg/d).

The highest median NRF9.3 score of 15.2 (12.6–17.3) was for meal pattern 8 consisting of toast- or egg-based breakfast, burger or sandwich at lunch, meat-based pizza or rice- or pasta-based dish at dinner, fruit or chocolate as snacks, and water or alcohol as an additional beverage. This meal pattern had the second highest median fibre intake of 20g/day (18.5–23.3g/day), and moderate intakes of other nutrients relative to the other meal patterns (**Figure 6.2**).

In contrast to the two meal patterns described above with a skipped breakfast (meal patterns 3 and 13), no clear trends in nutrient intakes were observed for the other three meal patterns with a skipped breakfast (meal patterns 4, 5, 12). These three patterns were more likely to include fruit as a snack or not to consume a snack. However, the five patterns with a skipped breakfast had the five lowest NRF9.3 scores compared to the other meal patterns.

No differences were observed among the meal patterns with regard to the health variables assessed.

6.4.2.2 Female Meal Patterns

Median z-scores by meal pattern for nutrient intake and health parameters for females are given in **Figure 6.3**, and confidence intervals are presented in Supplementary Figure 6.2. The highest median energy intakes for females of 2006 kcal/day (1859–2174 kcal/day) were observed for those in meal pattern 2 which is characterised by egg-based breakfast, sandwich or fruit at lunch, fried chicken at lunch, chocolate or fruit as snacks, and water as an additional beverage. Highest median intakes were also observed in this meal pattern for total fat of 89.7 g/day (77.9–100.1 g/day), MUFA of 29.8 g/day (26.4–33.8 g/day), PUFA of 20.8 g/day (18.4–23.8 g/day), and saturated fats of 28.7 g/day (23.4–32.5 g/day), and sodium of 3343 mg/day (3033–3741 mg/day). This pattern was associated with the second highest median intake for calcium of 884.8 mg/day (711–1034 mg/day) (Supplementary Table 6.7).

Those consuming meal pattern 13, had the lowest median NRF9.3 score observed for females of 8.8 (4.5–12.5), and was also associated with the lowest median fibre intake of 12.4 g/day (8.7–16.5 g/day) and iron intake of 9.7 mg/day (6.8–13.6 mg/day), and the second lowest median sodium intakes of 2592 mg/day (1820–3081 mg/day). Meal pattern 13 was characterised by skipped breakfast and lunch, fried chicken or rice-, pasta-, or pizza-based dish at dinner, ice-cream, cheese, or chocolate as snacks, and water as an additional beverage.

In contrast to meal pattern 13, described above, with a skipped breakfast, no clear trends in nutrient intakes were observed for the other four meal patterns with a skipped breakfast (meal patterns 1, 4, 8, and 10). These four patterns were more likely to include fruit at meals or to include rice-based meals as opposed to meals based on fried foods. However, the five patterns with a skipped breakfast had the five lowest NRF9.3 scores compared to the other meal patterns.

The highest median NRF9.3 score among females of 17.7 (14.8–19.9) was observed for meal pattern 3. This pattern also had the second lowest energy intake of 1809 kcal/day (1676–1975 kcal/day), the second lowest saturated fat intake of 22.0 g/day (17.8–25.1 g/day), and the highest fibre intake of 17.3 g/day (14.5–21.0 g/day). This pattern was characterised by toast- or egg-based breakfast, potatoes and chicken or rice-based lunch and dinner, chocolate biscuit or fruit-based snacks, and water as an additional beverage.

No differences were observed among the meal patterns with regard to the health variables assessed.

6.5 Discussion

The current study used a three-step clustering approach to identify meal patterns. This allowed for the examination of both temporal (meal skipping) and content (different combinations of meals) aspects of meal patterns within the same analysis, using a statistical approach that had not previously been used for this purpose. A diverse range of patterns were identified, with up to two meal types being skipped in some meal patterns; breakfast was more commonly skipped than other meal types. Nutrient quality varied among meal patterns, but the health parameters of those consuming them did not.

Other studies have used clustering to identify temporal meal patterns, but not content meal patterns. Both Chau *et al.* [26] and Khanna *et al.* [10] identified meal patterns that were characterised by the consumption of three meals spread across the day, at times which are culturally associated with breakfast, lunch, and dinner in Taiwan and the USA respectively. This three meal per day pattern was associated with the highest nutrient density [26] in the Taiwanese cohort, and with a higher HEI score and higher likelihood to have a BMI in the 20 to 25 kg/m^2 range in the American cohort [11]. Similarly, the current study also identified meal patterns in both the male data (meal patterns 7 and 10) and the female data (meal pattern 12) characterised by the consumption of breakfast, lunch, and dinner, but not of snacks. While those consuming three meals per day patterns in the current study, like the previous studies, had high diet quality scores, there were no clear and consistent differences observed regarding health parameters. The exploratory nature of the current study may have contributed to this. However, this cautious approach was deemed appropriate due to the novelty of the research and the large number of comparisons. In addition to this it is also not clear how consistent content meal patterns are over time. The data used in the current study were from a cross section with each participant completing two 24-hour recalls. Previous research has identified that individuals consume similar temporal meal patterns over a period of five years [27]; however, further research is required to determine what variability exists for content meal patterns over time.

The current study described temporal aspects of meal patterns, but did not examine the exact time at which meals were consumed; however, given that certain meal types are conventionally consumed at certain times in the day, i.e., breakfast in the morning, lunch in the afternoon, and dinner in the evening [3], these meal type descriptions serve as a proxy for timing of intakes. Recent and ongoing work has found that the timing of dietary intake relative to the timing of the circadian clock is thought to impact on metabolic processes [28]. While the research in the field of chrononutrition is not conclusive, with differing results emerging from different study designs, a recent meta-analysis of nine randomised

controlled trials, two of which were conducted in the USA, identified that a greater proportion of daily energy intake being consumed earlier in the day is associated with reduced weight and improved biomarkers of metabolic health compared to a greater proportion of energy intake being consumed later in the day [29]. In the current study, five of the male meal patterns and five of the female meal patterns did not include breakfast, indicating a greater proportion of food may be consumed later in the day. While the nutrient quality, as assessed by NRF9.3 score, of the meal patterns without breakfast were lower than those including breakfast, differences in relation to anthropometry and health parameters were not observed among the meal patterns. This may be due to the way in which meal types were used as a proxy for meal timing in the current study or related to the exploratory or cross-sectional nature of the study. However, using the novel approach applied in this analysis, the results of the current study also provide insight into, not only the timing, but also the content of these patterns. Meal patterns including fruit in meals or snacks or those with main meals based on rice rather than fried foods had higher NRF9.3 scores, despite not consuming breakfast. Further investigation regarding the novel interaction between timing of intakes and content of intakes, as presented in this study, is warranted.

The clearest differences among meal patterns observed in the current study were those for nutrient quality. This may have been influenced by the fact that the nutrients from the nutrient rich foods index (NRF9.3) [15] were used as part of the process to identify food groups and the NRF9.3 score was used as part of the process to identify generic meals. Further research is required to examine other methods that have not yet been used to identify meal patterns and to determine their impact on any differences arising among meal patterns. For example, reduced rank regression could be used to identify meal patterns in such a way that maximises differences in particular health outcomes among those meal patterns [30, 31].

The clustering approach taken in this analysis is novel; however, while clustering has not previously been used to identify content meal patterns, other methods such as PCA [6, 7] and LCA [8] have been used. While it would not be expected to identify identical meal patterns in different populations given the data-driven nature of these methods, in the related area of dietary patterns, there is reasonable agreement of major dietary patterns identified in different studies [30], which warrants comparison of the results of the current study with other studies on meal patterns. The meal patterns identified in previous studies using PCA did not specifically refer to meal skipping. The meal pattern study in an Irish cohort using LCA, on the other hand, did identify patterns with meal skipping, but unlike the current study, no patterns were identified where breakfast skipping was likely [8]. Four of the seven patterns in that study were associated with skipping lunch, compared to two of the

26 patterns (including both males and females) in the current study in which skipping lunch was likely. Another study using PCA was carried out in the same Irish cohort [6] and identified some meal patterns that are similar to those in the current study. For example, some meal patterns in the Irish cohort also feature bread-, cereal-, and egg-based breakfasts, meat and rice or pasta main meals, and sandwich-based light meals. This is in contrast with a study in a Japanese cohort that identified content meal patterns using PCA [7]. In that study, for example, rice was a more common feature of meals, appearing in all but one of the meal patterns. In the current study, however, rice appears in four of thirteen male meal patterns and eleven of thirteen female meal patterns. Seven of the eleven meal patterns in the Japanese cohort were likely to contain vegetables at breakfast whereas this was not the case in either the Irish cohort [6] or the American cohort in the current study. In contrast to the meal patterns in both of those previous cohorts where fried foods and burgers were absent, they appeared frequently in the lunches and dinners identified in the current study. Similar snacks were seen across all studies with confectionary and fruit being identified.

Limited research has examined the relationship between content meal patterns and diet quality or health outcomes. Uzhova *et al.* [8] found a meal pattern in Ireland characterised by a cooked breakfast, skipped light meal, and a protein and carbohydrate main meal were more likely to be associated with higher diastolic blood pressure compared with those consuming a pattern characterised by cereal and/or toast for breakfast, a sandwich light meal, and a protein and carbohydrate or just protein main meal. In that study there were no associations identified between other meal patterns and health outcomes. The current study also identified differences in nutrient quality, but not in health outcomes for various meal patterns; however, given the exploratory nature of this study, further work is required to confirm these findings.

The results presented should be interpreted in the context of the limitations and strengths of the current study. The data used in this study from the NHANES represents a cross section, and therefore does not capture the potential within-person variability in meal patterns such as the potential for meal patterns to change over time. The study design used in the NHANES oversampled population subgroups that were of particular public health interest including Hispanic, Asian, and black people; people with low income; and people aged 80 years or older [13], and this will have influenced the meal patterns identified in the current study. The large size of the dietary intake dataset, however, is a strength in terms of capturing a range of varying meal patterns consumed. The data were limited due to missing health data for a proportion of participants, and this may have increased the uncertainty of the estimates. While the current study provides valuable insights to these

meal patterns the exploratory nature precludes any strong conclusions regarding causation, however, taking that approach has raised important research questions for this field in relation to the need to develop research designs that consider both the temporal and content aspects of meal patterns and to further evaluate the current methods used with regard to their comparability.

6.6 Conclusion

Clustering is a feasible approach to identify combined aspects of temporal and content meal patterns. This exploratory study identified differences among meal patterns in nutrient quality but not in health parameters. Further research is required to refine meal pattern methods and to identify patterns combining multiple meal pattern constructs. Other statistical methods to identify meal patterns should also be considered that can incorporate the health of those consuming meal patterns rather than just the nutrient content of the foods and meals being consumed.

6.7 Tables

Table 6.1: The original food groups from the USA dietary guidelines and subsequent subgroups following *k*-means clustering.

Food Group	Number of Subgroups or Clusters	Description of Subgroups/Clusters
Fruit	2	Fruits 1 Fruits 2
Vegetables	3	Vegetables 1 Vegetables 2 Vegetables 3
Grains	2	Cereals and sweetened instant oats Bread, oats, pasta, or rice
Protein Foods	3	Protein 1 Protein 2 Protein 3
Dairy	2	Cheese, yogurts, and sweet dairy Milk, yogurts, and low-fat dairy
Foods High in Fat, Sugar, or Salt	4	Mixed (sweet and savoury) Savory 1 Savory 2 Sweet
Fats and Oils (not clustered; original subgroups maintained)	3	Animal Fats Margarines Dressings and vegetable oils
Non-Alcoholic Beverages (not clustered; manually grouped)	2	Energy free beverages Energy containing beverages
Alcoholic Beverages (not further subdivided)	1	Alcoholic beverages
Other (not further subdivided)	1	Other

Descriptions of subgroups/clusters were assigned after examination of the foods contained in the group; where no clear pattern was observed, numeric descriptors are given instead.

Table 6.2: The three most commonly consumed meal patterns by males.

Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
4	n = 211	Not consumed	Chicken burger (38%) OR Cheese and chicken sandwich with water (22%)	Fried chicken and water (18%) OR Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise (17%)	Banana/apple/grapes (45%) OR Chocolate (cookie) (41%)	Water (69%) OR Soft drink or coffee or tea (36%)
6	n = 241	Toast (54%) OR Cereal and milk (36%)	Chicken burger (39%) OR Soft drink (cola) and chicken burger (20%)	Fried chicken (39%) OR Soft drink (cola) or iced tea with meat pizza (26%)	Chocolate (cookie) (42%) OR Ice cream or cheese (35%)	Water (64%) OR Soft drink or coffee or tea (42%)
8	n = 309	Toast (53%) OR Scrambled egg or omelette (48%)	Soft drink (cola) with chicken burger and chips (35%) OR Sandwich with luncheon ham and mayonnaise (20%)	Meat pizza (35%) OR Rice/pasta dish with water (16%)	Banana/apple/grapes (60%) OR Chocolate (cookie) (39%)	Water (83%) OR Alcohol (beer) (30%)

The top two generic meals in each meal type and cluster are given with the percentage of people who consumed those meals in that meal pattern given in parentheses.

Table 6.3: The three most commonly consumed meal patterns by females.

Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
7	n = 214	Cereal and milk (55%) OR Toast and water (51%)	Fried egg sandwich and water (62%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes) (21%)	Water, fried chicken, chips (43%) OR Fried chicken and chips (30%)	Chocolate cookie (38%) OR Crackers/popcorn (35%)	Water (82%) OR Soft drink or coffee or tea (28%)
5	n = 237	Scrambled egg or omelette (59%) OR Cereal and milk (32%)	Rice-based dish or pizza with bread (46%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich (19%)	Soft drink (cola) or iced tea with rice-based dish or pizza (30%) OR Fried chicken (24%)	Chocolate cookie (44%) OR Banana/apple/grapes (27%)	Not consumed
9	n = 320	Scrambled egg or omelette (52%) OR Toast and water (52%)	Rice-based dish or pizza with bread (58%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich (32%)	Water with rice/pasta-based dish (35%) OR Soft drink (cola) or iced tea with rice-based dish or pizza (19%)	Chocolate cookie (39%) OR Ice cream / cheese (32%)	Water (84%) OR Soft drink or coffee or tea (38%)

The top two generic meals in each meal type and cluster are given with the percentage of people who consumed those meals in that meal pattern given in parentheses.

Table 6.4: Demographics and anthropometry of study participants.

	Males n = 1432 (48.3%)	**Females** n = 1532 (51.7%)	**Total** n = 2964
Age (y)	53 (35–65)	51 (34–64)	52 (34–64)
Weight (kg)	83.0 (72.2–96.7)	71.6 (60.4–84.6)	77.2 (65.6–91.3)
Height (cm)	173.4 (168.2–178.7)	159.8 (155.3–164.6)	166.1 (159.2–173.7)
BMI (kg/m^2)	27.8 (24.4–31.7)	28.0 (24.0–33.0)	27.9 (24.2–32.3)

Values are given as median with interquartile range in parentheses. Includes participants with plausible reports of energy intake only.

6.8 Figures

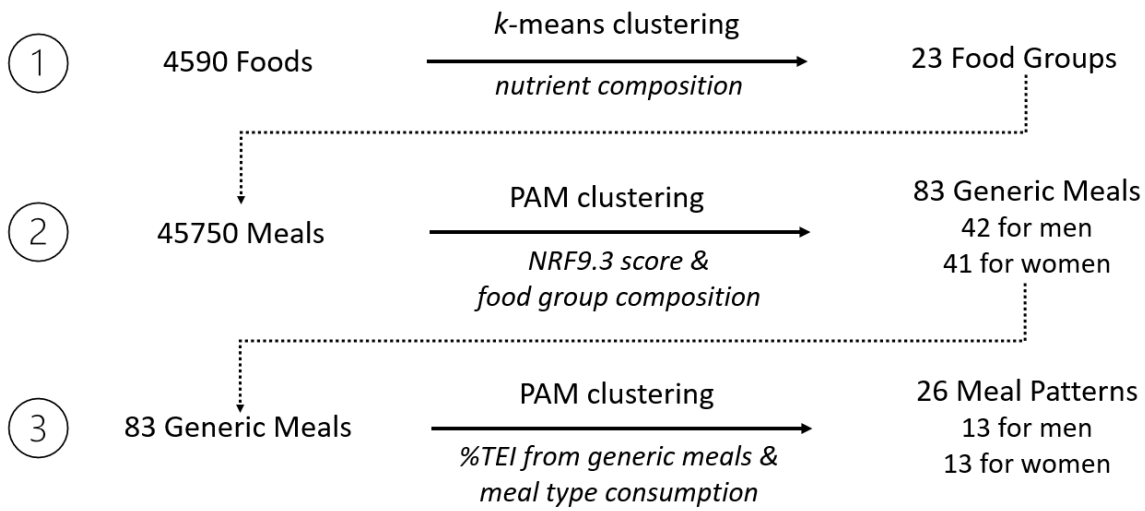


Figure 6.1: Overview of the three steps used in identifying meal patterns.

NRF, nutrient rich foods; PAM, partitioning around the medoids; TEI, total energy intake.

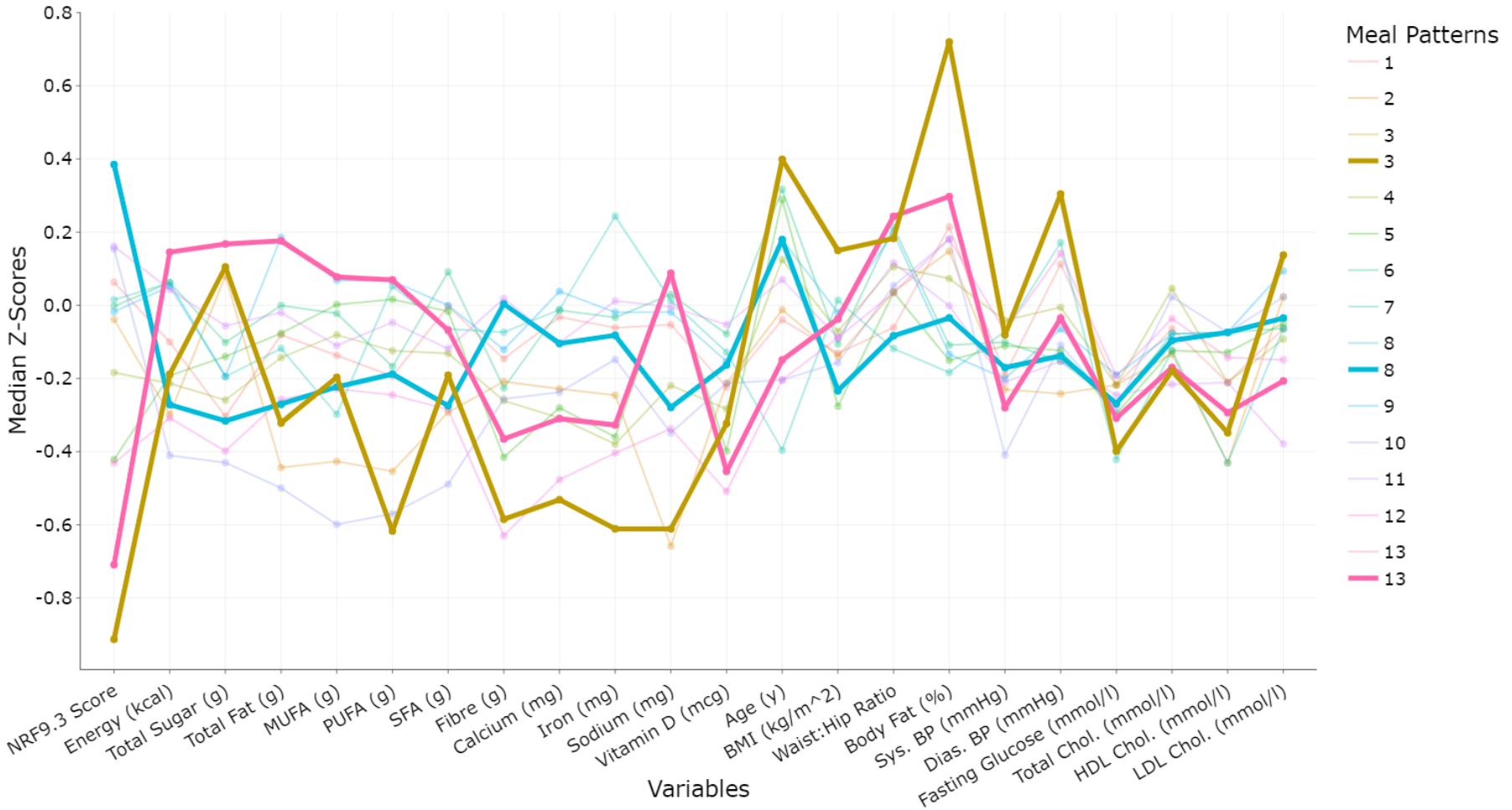


Figure 6.2: Parallel plot of median z-scores for selected nutrients and health parameters within each meal pattern in males.

Meal patterns in which differences among meal patterns for nutrient quality were observed are highlighted.

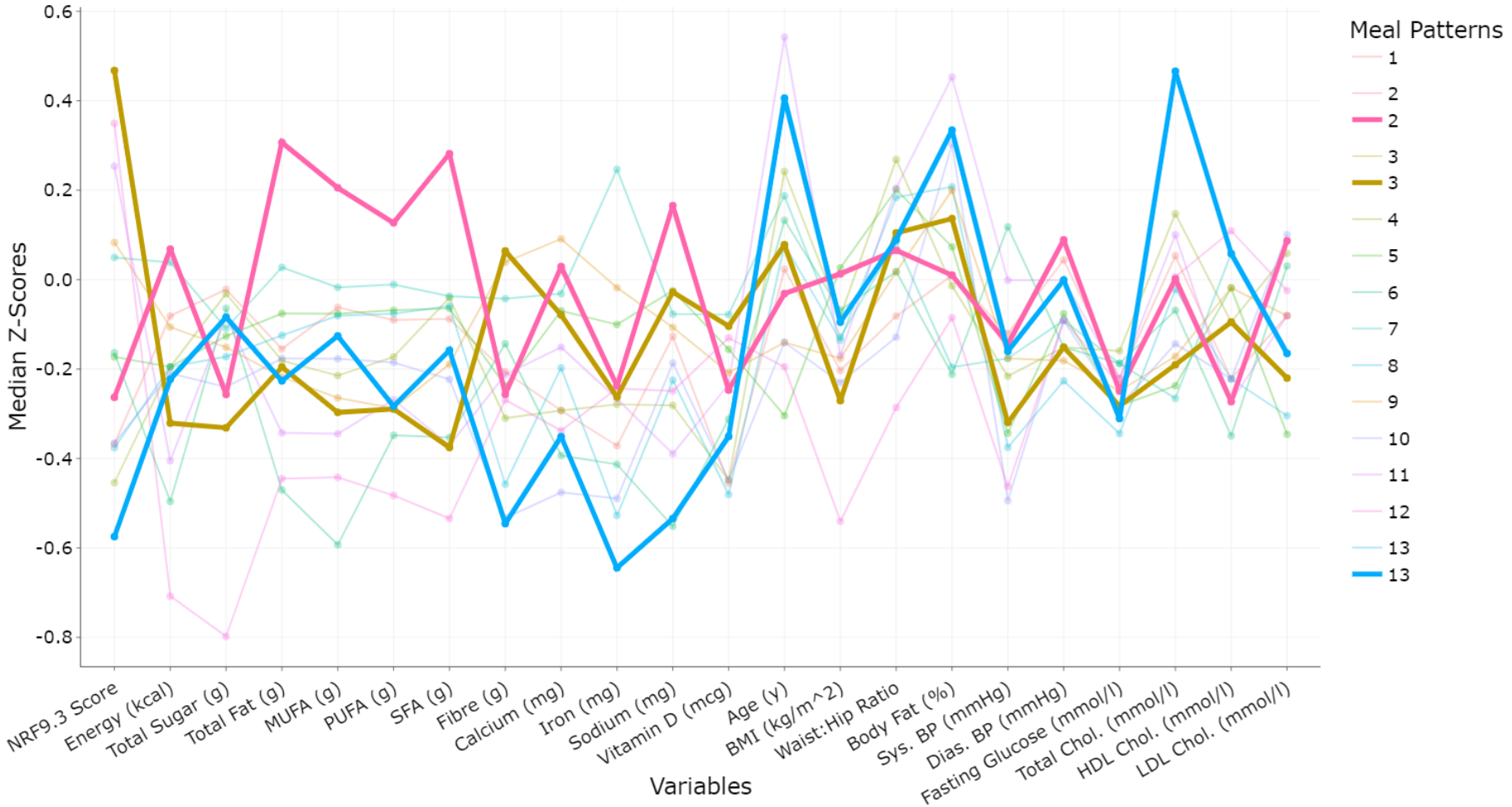


Figure 6.3: Parallel plot of median z-scores for selected nutrients and health parameters within each identified meal pattern in females.

Meal patterns in which differences among meal patterns for nutrient quality were observed are highlighted.

6.9 References

1. O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

2. Leech RM, Worsley A, Timperio A*, et al.* (2015) Understanding Meal Patterns: Definitions, Methodology and Impact on Nutrient Intake and Diet Quality. *Nutr Res Rev*, **28**(1), 1-21.

3. Dietary Guidelines Advisory Committee (2020) Scientific Report of the 2020 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture and the Secretary of Health and Human Services. Washington, DC: *U.S. Department of Agriculture, Agricultural Research Service*.

4. Ministry of Health of Brazil (2015) Dietary Guidelines for the Brazilian Population. Brasília/DF: *Secretariat of Health Care, Primary Health Care Department*.

5. Trang Tran TN, Atas M, Felfernig A*, et al.* (2017) An Overview of Recommender Systems in the Healthy Food Domain. *Journal of Intelligent Information Systems*, **50**(3), 501-526.

6. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

7. Murakami K, Livingstone MBE, Sasaki S*, et al.* (2018) Applying a Meal Coding System to 16-D Weighed Dietary Record Data in the Japanese Context: Towards the Development of Simple Meal-Based Dietary Assessment Tools. *J Nutr Sci*, **7**(e29), 1-10.

8. Uzhova I, Woolhead C, Timon CM*, et al.* (2018) Generic Meal Patterns Identified by Latent Class Analysis: Insights from NANS (National Adult Nutrition Survey). *Nutrients*, **10**(3), 310.

9. O'Hara C, O'Sullivan A, and Gibney ER (2022) A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data. *J Nutr*, **152**(10), 2297-2308.

10. Khanna N, Eicher-Miller HA, Boushey CJ*, et al.* (2011) Temporal Dietary Patterns Using Kernel K-Means Clustering. *IEEE International Symposium on Multimedia*, **2011**, Dana Point, CA, USA, 375-380.

11. Eicher-Miller HA, Khanna N, Boushey CJ*, et al.* (2016) Temporal Dietary Patterns Derived among the Adult Participants of NHANES 1999-2004 Are Associated with Diet Quality. *J Acad Nutr Diet*, **116**(2), 283-291.

12. Shi Z, Riley M, Taylor A*, et al.* (2017) Meal-Specific Food Patterns and the Incidence of Hyperglycemia in a Chinese Adult Population. *Br J Nutr*, **118**(1), 53-59.

13. Centers for Disease Control and Prevention (CDC) (2018) National Health and Nutrition Examination Survey (NHANES) Data 2017-2018. Available at:

https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 (accessed 14 December 2023).

14. U.S. Department of Agriculture (2020) USDA Food and Nutrient Database for Dietary Studies 2017-2018. Available at: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds-download-databases/ (accessed 14 December 2023).

15. Fulgoni VL, Keast DR, and Drewnowski A (2009) Development and Validation of the Nutrient-Rich Foods Index: A Tool to Measure Nutritional Quality of Foods. *J Nutr*, **139**(8), 1549-1554.

16. Charrad M, Ghazzali N, Boiteau V*, et al.* (2014) Nbclust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, **61**(6), 1-36.

17. Kant AK and Graubard BI (2015) 40-Year Trends in Meal and Snack Eating Behaviors of American Adults. *J Acad Nutr Diet*, **115**(1), 50-63.

18. Mathias KC, Almoosawi S, and Karagounis LG (2017) Protein and Energy Intakes Are Skewed toward the Evening among Children and Adolescents in the United States: NHANES 2013–2014. *J Nutr*, **147**(6), 1160-1166.

19. Barrera CM, Moore LV, Perrine CG*, et al.* (2019) Number of Eating Occasions and Source of Foods and Drinks among Young Children in the United States: NHANES, 2009-2014. *Nutrients*, **11**(4), 897.

20. Sullivan VK, Na M, Proctor DN*, et al.* (2021) Consumption of Dried Fruits Is Associated with Greater Intakes of Underconsumed Nutrients, Higher Total Energy Intakes, and Better Diet Quality in US Adults: A Cross-Sectional Analysis of the National Health and Nutrition Examination Survey, 2007-2016. *J Acad Nutr Diet*, **121**(7), 1258-1272.

21. R Core Team (2022) R: A Language and Environment for Statistical Computing, 4.2.2 ed. Vienna, Austria: R Foundation for Statistical Computing.

22. RStudio Team (2022) Rstudio: Integrated Development for R., 2022.07.2+576 ed. Boston, MA, USA: RStudio PBC.

23. Carpenter J and Bithell J (2000) Bootstrap Confidence Intervals: When, Which, What? A Practical Guide for Medical Statisticians. *Stat Med*, **19**(9), 1141-1164.

24. Henry CJ (2005) Basal Metabolic Rate Studies in Humans: Measurement and Development of New Equations. *Public Health Nutr*, **8**(7A), 1133-1152.

25. Black AE (2000) Critical Evaluation of Energy Intake Using the Goldberg Cut-Off for Energy Intake:Basal Metabolic Rate. A Practical Guide to Its Calculation, Use and Limitations. *Int J Obes*, **24**, 1119-1130.

26. Chau CA, Pan WH, and Chen HJ (2017) Employment Status and Temporal Patterns of Energy Intake: Nutrition and Health Survey in Taiwan, 2005–2008. *Public Health Nutr*, **20**(18), 3295-3303.

27. Wilson JE, Blizzard L, Gall SL*, et al.* (2019) An Eating Pattern Characterised by Skipped or Delayed Breakfast Is Associated with Mood Disorders among an Australian Adult Cohort. *Psychol Med*, **50**(16), 2711-2721.

28. Katsi V, Papakonstantinou IP, Soulaidopoulos S*, et al.* (2022) Chrononutrition in Cardiometabolic Health. *J Clin Med*, **11**(2), 296.

29. Young IE, Poobalan A, Steinbeck K*, et al.* (2023) Distribution of Energy Intake across the Day and Weight Loss: A Systematic Review and Meta-Analysis. *Obes Rev*, **24**(3), e13537.

30. Moeller SM, Reedy J, Millen AE*, et al.* (2007) Dietary Patterns: Challenges and Opportunities in Dietary Patterns Research. *J Am Diet Assoc*, **107**(7), 1233-1239.

31. Schulz CA, Oluwagbemigun K, and Nothlings U (2021) Advances in Dietary Pattern Analysis in Nutritional Epidemiology. *Eur J Nutr*, **60**(8), 4115-4130.

6.10 Supplementary Material

Supplementary Table 6.1: The 24 indices used in determining the number of clusters in the clustering steps of identifying meal patterns.

Index	Interpretation
KL	Max value
CH	Max value
Hartigan	Max difference between hierarchy levels of index
CCC	Max value
Scott	Max difference between hierarchy levels of index
Marriot	Max value of second differences between levels of the index
TrCovW	Max difference between hierarchy levels of index
TraceW	Max value of second differences between levels of the index
Friedman	Max difference between hierarchy levels of index
Rubin	Min value of second differences between levels
Cindex	Min value
DB	Min value
Silhouette	Max value
Duda	Smallest number of clusters such that index > critical value
Beale	Number of clusters such that critical value ≥ alpha
Ratkowsky	Max value
Ball	Max difference between hierarchy levels of index
Ptbiserial	Max value
McClain	Min value
Dunn	Max value
Hubert	Graphical method: cluster number where a peak occurred on the second differences plot.
SDindex	Min value
Dindex	Graphical method: cluster number where a peak occurred on the second differences plot.
SDbw	Min value

Adapted from Charrad *et al.* [16].

Supplementary Table 6.2: Demographics for all participants regardless of plausibility of reported energy intakes.

	Males	**Females**	**Total**
	n = 2068 (47.7%)	n = 2264 (52.3%)	n = 4332
Age (y)	52.5 (34–65)	54 (34–64)	52 (34–64)
Weight (kg)	85.1 (73.5–100.4)	74.2 (62.4–89.3)	79.7 (67.4–95.7)
Height (cm)	173.7 (168.5–178.9)	160.0 (155.3–164.6)	166.2 (159.1–173.8)
BMI (kg/m^2)	28.4 (24.9–32.9)	29.0 (24.6–34.8)	28.7 (24.8–33.7)

Values are given as median (interquartile range).

Supplementary Table 6.3: The generic meals identified as part of the process of identifying meal patterns.

	Food Groups	Example Meal
	Males' Breakfast Meals	
1	Protein 1 Non-alcoholic beverages no energy	Scrambled egg or omelette
2	Milk and low-fat dairy Cereals and instant oats	Cereal and milk
3	Non-alcoholic beverages with energy Sweet HFSS	Coffee and sugar
4	Bread, oats, pasta, rice	Toast
	Females' Breakfast Meals	
1	Non-alcoholic beverages with energy Sweet HFSS	Coffee and sugar
2	Protein 1	Scrambled egg or omelette
3	Bread, oats, pasta, rice Non-alcoholic beverages no energy	Toast and water
4	Milk and low-fat dairy Cereals and instant oats	Cereal and milk
	Males' Lunch Meals	
1	Bread, oats, pasta, rice Non-alcoholic beverages no energy	Rice-based dish or pizza with water
2	N/A, see detailed notes in relevant section	Soft drink with pizza
3	Breads, oats, pasta, rice Other	Sandwich with luncheon ham and mayonnaise
4	N/A, see detailed notes in relevant section	Cheese and chicken sandwich with water
5	Protein1 Other	Fried chicken and sauce/soup
6	Non-alcoholic beverages with energy Protein 1	Soft drink (cola) and chicken burger
7	Non-alcoholic beverages with energy Protein 1 Vegetables 1	Soft drink (cola) with chicken burger and chips
8	Protein 1	Chicken burger
9	Non-alcoholic beverages with energy Protein 1 Vegetables 3 Vegetables 1 Other	Soft drink (cola), chicken/cheese burger, lettuce, tomatoes, chips, and mayonnaise
10	Bread, oats, pasta, rice Cheese and sweet dairy Non-alcoholic beverages no energy	Bread, cheese, luncheon ham, tomatoes, lettuce, mayonnaise, and water

	Vegetables 3 Protein 3 Other	

Females' Lunch Meals		
1	Bread, oats, pasta, rice Cheese and sweet dairy Non-alcoholic beverages no energy Vegetables 3 Protein 3 Other	Bread, cheese, luncheon ham, tomatoes, lettuce, mayonnaise, and water
2	Bread, oats, pasta, rice	Rice-based dish or pizza with bread
3	Non-alcoholic beverages with energy Protein 1	Soft drink (cola) and chicken sandwich/burger or fried egg sandwich
4	Vegetables 1 Protein 3	Potatoes and chicken
5	Protein 1 Non-alcoholic beverages no energy	Fried egg sandwich and water
6	Fruits 1	Fruit juice (apple) or raw fruit (apple, banana, grapes)
7	Non-alcoholic beverages with energy Protein 1 Vegetables 3 Vegetables 1 Other	Chicken burger with lettuce, tomatoes, chips, mayonnaise, and soft drink (cola) or iced tea

Males' Dinner Meals		
1	Non-alcoholic beverages with energy Protein 1 Vegetables 1	Soft drink (cola), fried chicken sandwich and chips
2	Non-alcoholic beverages with energy Bread, oats, pasta, rice	Soft drink (cola) or iced tea with meat pizza
3	Protein1 Bread, oats, pasta, rice Other	Chicken, bread or rice, and sauce
4	Cheese and sweet dairy Dressings and vegetable oils Vegetables 3 Vegetables 1 Protein 3	Chicken salad with cheese, dressing, lettuce, tomatoes, and onions
5	Bread, oats, pasta, rice Non-alcoholic beverages no energy	Rice/pasta dish with water
6	Bread, oats, pasta, rice Cheese and sweet dairy Protein 3 Other	Bread, cheese, luncheon ham, mayonnaise
7	Protein 1	Fried chicken and water

	Non-alcoholic beverages no energy	
8	N/A, see detailed notes in relevant section	Soft drink with cookie or cake
9	Bread, oats, pasta, rice Non-alcoholic beverages no energy Vegetables 1 Protein 3 Other	Rice, water, potatoes, chicken, and sauce
10	N/A, see detailed notes in relevant section	Chicken and cheese sandwich
11	Bread, oats, pasta, rice	Meat pizza
12	Protein 1 Vegetables 3 Vegetables 1 Other	Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise
13	Vegetables 1 Protein 3	Potatoes and chicken
14	Protein 2 Fruits 1	Bean-based salad and fruit/fruit juice
15	Protein 1	Fried chicken

Females' Dinner Meals		
1	Non-alcoholic beverages, no energy	Water and soup
2	Non-alcoholic beverages, no energy Dressings and vegetable oils Vegetables 3 Vegetables 1 Protein 3	Water, chicken salad with tomatoes, lettuce, cucumber, onions, and dressing
3	Bread, oats, pasta, rice Non-alcoholic beverages no energy Vegetables 1 Protein 3 Other	Water, rice, potatoes, chicken, and sauce
4	Vegetables 1 Protein 3	Potatoes and chicken
5	Protein 1	Fried chicken
6	Non-alcoholic beverages with energy Bread, oats, pasta, rice	Soft drink (cola) or iced tea with rice-based dish or pizza
7	Non-alcoholic beverages with energy Protein 1 Vegetables 3 Vegetables 1 Other	Soft drink (cola) or iced tea with chicken burger, lettuce, tomatoes, chips, and mayonnaise
8	Protein 1 Vegetables 1	Fried chicken and chips
9	Non-alcoholic beverages no energy Protein 1	Water, fried chicken, chips

	Vegetables 1	
10	Bread, oats, pasta, rice Non-alcoholic beverages no energy	Water with rice/pasta-based dish
11	Bread, oats, pasta, rice	Rice/pasta-based dish or bread or pizza
12	Cheese and sweet dairy	Cheese

Males' Snack Meals		
1	Sweet HFSS	Chocolate (cookie)
2	Savoury HFSS	Crackers or popcorn
3	N/A, see detailed notes in relevant section	Coffee with sugar and cookie
4	Non-alcoholic beverages with energy	Soft drink (cola)
5	Non-alcoholic beverages no energy	Water
6	Fruits 1	Banana/apple/grapes
7	Cheese and sweet dairy	Ice cream or cheese
8	Protein1 Protein2	Tortilla chips and cheese and salted peanuts
9	Sweet HFSS Non-alcoholic beverages no energy	Chocolate cookie and water

Females' Snack Meals		
1	Non-alcoholic beverages with energy	Soft drink (cola)
2	Non-alcoholic beverages with energy Sweet HFSS	Coffee and sugar
3	Fruits 1	Banana/apple/grapes
4	Protein 2	Nuts
5	Sweet HFSS Non-alcoholic beverages no energy	Chocolate cookie and water
6	Non-alcoholic beverages no energy	Water
7	Sweet HFSS	Chocolate cookie
8	Fruits 2	Orange/strawberries
9	Savoury HFSS 2	Crackers/popcorn
10	Protein 1	Tortilla chips and cheese
11	Cheese and sweet dairy	Ice cream / cheese
12	Non-alcoholic beverages no energy Fruits 1	Water and banana/apple/grapes
13	Bread, oats, pasta, rice Non-alcoholic beverages no energy	Bread and water
14	Savoury HFSS1	Crisps

Males' Beverages		
1	All alcoholic beverages	Alcohol (beer)
2	Non-alcoholic beverages with energy	Soft drink or coffee or tea
3	Non-alcoholic beverages with energy Sweet HFSS	Coffee with sugar
4	Non-alcoholic beverages no energy	Water

	Females' Beverages	
1	Non-alcoholic beverages, no energy	Water
2	Non-alcoholic beverages with energy	Soft drink or coffee or tea
3	All alcoholic beverages	Wine or beer
4	Non-alcoholic beverages with energy Sweet HFSS	Coffee with sugar

Supplementary Table 6.4: The meal patterns consumed by males.

Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
1	n = 181	Scrambled egg or omelette (67%) OR Cereal and milk (30%)	Rice-based dish or pizza with water (50%) OR Soft drink with pizza (20%)	Soft drink (cola), fried chicken sandwich and chips (39%) OR Rice/pasta dish with water (18%)	Chocolate (cookie) (38%) OR Banana/apple/grapes (35%)	Water (86%) OR Soft drink or coffee or tea (36%)
2	n = 116	Toast (61%) OR Scrambled egg or omelette (47%)		Soft drink (cola) or iced tea with meat pizza (22%) OR Meat pizza (21%)	Banana/apple/grapes (51%) OR Chocolate (cookie) (40%)	Water (66%) OR Soft drink or coffee or tea (32%)
3	n = 80			Soft drink (cola), fried chicken sandwich and chips (26%) OR Meat pizza (22%)	Chocolate (cookie) (45%) OR Soft drink (cola) (38%)	Water (44%) OR Soft drink or coffee or tea (34%)
4	n = 211		Chicken burger (38%) OR Cheese and chicken sandwich with water (22%)	Fried chicken and water (18%) OR Cheeseburger, lettuce, tomatoes, chips, and Mayonnaise (17%)	Banana/apple/grapes (45%) OR Chocolate (cookie) (41%)	Water (69%) OR Soft drink or coffee or tea (36%)

Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
5	n = 176		Soft drink with pizza (30%) OR Chicken burger (24%)	Fried chicken (26%) OR Soft drink (cola), fried chicken sandwich and chips (16%)	Chocolate (cookie) (40%) OR Soft drink (cola) (34%)	
6	n = 241	Toast (54%) OR Cereal and milk (36%)	Chicken burger (39%) OR Soft drink (cola) and chicken burger (20%)	Fried chicken (39%) OR Soft drink (cola) or iced tea with meat pizza (26%)	Chocolate (cookie) (42%) OR Ice cream or cheese (35%)	Water (64%) OR Soft drink or coffee or tea (42%)
7	n = 73	Toast (47%) OR Scrambled egg or omelette (41%)	Rice-based dish or pizza with water (34%) OR Soft drink (cola) with chicken burger and chips (19%)	Meat pizza (29%) OR Potatoes and chicken (26%)		
8	n = 309	Toast (53%) OR Scrambled egg or omelette (48%)	Soft drink (cola) with chicken burger and chips (35%) OR Sandwich with luncheon ham and mayonnaise (20%)	Meat pizza (35%) OR Rice/pasta dish with water (16%)	Banana/apple/grapes (60%) OR Chocolate (cookie) (39%)	Water (83%) OR Alcohol (beer) (30%)

Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
9	n = 152	Scrambled egg or omelette (71%) OR Cereal and milk (26%)	Chicken burger (35%) OR Soft drink (cola) and chicken burger (19%)	Fried chicken (28%) OR Fried chicken and water (18%)	Banana/apple/grapes (50%) OR Chocolate (cookie) (37%)	
10	n = 140	Scrambled egg or omelette (48%) OR Toast (46%)	Soft drink with pizza (21%) OR Sandwich with luncheon ham and mayonnaise (21%)	Meat pizza (19%) OR Chicken and cheese sandwich (16%)		Water (79%) OR Soft drink or coffee or tea (26%)
11	n = 164	Toast (66%) OR Cereal and milk (30%)	Rice-based dish or pizza with water (28%) OR Soft drink with pizza (20%)	Meat pizza (26%) OR Chicken and cheese sandwich (19%)	Banana/apple/grapes (41%) OR Chocolate (cookie) (34%)	
12	n = 80		Soft drink with pizza (19%) OR Soft drink (cola) and chicken burger (19%)	Fried chicken and water (30%) OR Soft drink (cola), fried chicken sandwich and chips (21%)		Water (55%) OR Soft drink or coffee or tea (30%)

Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
13	n = 145		Soft drink (cola) and chicken burger (49%) OR Soft drink with pizza (34%)	Soft drink (cola), fried chicken sandwich and chips (46%) OR Soft drink (cola) or iced tea with meat pizza (21%)	Chocolate (cookie) (44%) OR Soft drink (cola) (32%)	Water (59%) OR Soft drink or coffee or tea (34%)

The top two generic meals in each meal type and cluster are given with the percentage of people who consumed those meals in that meal pattern in parentheses.

Supplementary Table 6.5: The meal patterns consumed by females.

Meal Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
1	n = 176		Rice-based dish or pizza with bread (49%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes) (22%)	Soft drink (cola) or iced tea with rice-based dish or pizza (30%) OR Soft drink (cola) or iced tea with chicken burger, lettuce, tomatoes, chips, and mayonnaise (23%)	Chocolate cookie (34%) OR Crackers/popcorn (25%)	Water (83%) OR Soft drink or coffee or tea (31%)
2	n = 159	Scrambled egg or omelette (72%) OR Toast and water (29%)	Fried egg sandwich and water (45%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes) (20%)	Fried chicken (67%) OR Soft drink (cola) or iced tea with chicken burger, lettuce, tomatoes, chips, and mayonnaise (26%)	Chocolate cookie (38%) OR Banana/apple/grapes (28%)	Water (79%) OR Soft drink or coffee or tea (28%)
3	n = 210	Toast and water (65%) OR Scrambled egg or omelette (45%)	Potatoes and chicken (41%) OR Rice-based dish or pizza with bread (25%)	Rice/pasta-based dish or bread or pizza (49%) OR Potatoes and chicken (35%)	Chocolate cookie (50%) OR Banana/apple/grapes (37%)	Water (84%) OR Soft drink or coffee or tea (28%)

Meal Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
4	n = 147		Fried egg sandwich and water (52%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich (48%)	Fried chicken (33%) OR Water with rice/pasta-based dish (33%)	Chocolate cookie (37%) OR Banana/apple/grapes (31%)	Water (84%) OR Soft drink or coffee or tea (23%)
5	n = 237	Scrambled egg or omelette (59%) OR Cereal and milk (32%)	Rice-based dish or pizza with bread (46%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich (19%)	Soft drink (cola) or iced tea with rice-based dish or pizza (30%) OR Fried chicken (24%)	Chocolate cookie (44%) OR Banana/apple/grapes (27%)	
6	n = 123	Toast and water (61%) OR Scrambled egg or omelette (59%)		Rice/pasta-based dish or bread or pizza (26%) OR Soft drink (cola) or iced tea with rice-based dish or pizza (25%)	Banana/apple/grapes (42%) OR Chocolate cookie (34%)	Water (70%) OR Soft drink or coffee or tea (24%)

Meal Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
7	n = 214	Cereal and milk (55%) OR Toast and water (51%)	Fried egg sandwich and water (62%) OR Fruit juice (apple) or raw fruit (apple, banana, grapes) (21%)	Water, fried chicken, chips (43%) OR Fried chicken and chips (30%)	Chocolate cookie (38%) OR Crackers/popcorn (35%)	Water (82%) OR Soft drink or coffee or tea (28%)
8	n = 186		Rice-based dish or pizza with bread (38%) OR Fried egg sandwich and water (31%)	Fried chicken (23%) OR Soft drink (cola) or iced tea with rice-based dish or pizza (22%)	Chocolate cookie (41%) OR Banana/apple/grapes (27%)	
9	n = 320	Scrambled egg or omelette (52%) OR Toast and water (52%)	Rice-based dish or pizza with bread (58%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich (32%)	Water with rice/pasta-based dish (35%) OR Soft drink (cola) or iced tea with rice-based dish or pizza (19%)	Chocolate cookie (39%) OR Ice cream / cheese (32%)	Water (84%) OR Soft drink or coffee or tea (38%)
10	n = 118		Rice-based dish or pizza with bread (42%)	Fried chicken and chips (62%) OR	Chocolate cookie (47%) OR	Water (70%) OR

Meal Pattern	Sample Size	Breakfast	Lunch	Dinner	Snacks	Beverages
			OR Fried egg sandwich and water (22%)	Rice/pasta-based dish or bread or pizza (47%)	Banana/apple/grapes (29%)	Soft drink or coffee or tea (33%)
11	n = 194	Toast and water (70%) OR Cereal and milk (32%)	Fried egg sandwich and water (38%) OR Soft drink (cola) and chicken sandwich/burger or fried egg sandwich (32%)	Potatoes and chicken (26%) OR Water with rice/pasta-based dish (22%)	Banana/apple/grapes (32%) OR Chocolate cookie (32%)	
12	n = 106	Toast and water (54%) OR Scrambled egg or omelette (38%)	Fried egg sandwich and water (42%) OR Rice-based dish or pizza with bread (35%)	Water with rice/pasta-based dish (40%) OR Water and soup (28%)		Water (72%) OR Soft drink or coffee or tea (25%)
13	n = 74			Fried chicken (31%) OR Rice/pasta-based dish or bread or pizza (24%)	Ice cream / cheese (31%) OR Chocolate cookie (30%)	Water (53%) OR Soft drink or coffee or tea (20%)

The top two generic meals in each meal type and cluster are given with the percentage of people who consumed those meals in that meal pattern in parentheses.

Supplementary Table 6.6: Daily energy intake, nutrient intakes, and health parameters by meal pattern for males with plausible reports of energy intake.

Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13
Sample Size	144	72	45	149	116	176	36	231	110	79	121	46	107
NRF 9.3	12.6 (10.6–15.5)	11.8 (9.0–15.5)	5.0 (0.8–9.8)	10.7 (8.4–13.4)	8.8 (5.3–10.5)	12.1 (10.0–14.0)	12.3 (7.9–16.0)	15.2 (12.6–17.3)	12.0 (9.5–13.4)	13.4 (8.9–16.5)	13.4 (9.5–15.9)	8.8 (3.5–13.8)	6.6 (3.1–9.1)
Energy (kcal)	2472 (2202–2684)	2341 (1916–2666)	2414 (2080–3059)	2397 (2221–2622)	2411 (2186–2755)	2579 (2277–2780)	2579 (2001–2962)	2359 (2227–2514)	2572 (2202–2821)	2267 (2065–2530)	2566 (2253–2804)	2335 (2037–2728)	2634 (2387–2887)
Sugar (g)	103.5 (85.9–129.2)	129.6 (96.7–153.9)	131.3 (98.9–186.9)	106.5 (89.8–123.0)	114.6 (84.5–136.5)	117.2 (99.7–137.4)	110.9 (73.8–147.7)	102.6 (91.2–115.6)	110.9 (90.9–124.9)	94.8 (64.2–117.1)	120.3 (103.9–139.6)	97.0 (73.8–144.7)	135.5 (110.6–158.9)
Fat (g)	99.6 (90.7–114.2)	86.7 (66.8–105.1)	91.0 (68.3–114.2)	97.4 (86.1–109.4)	99.7 (87.0–116.2)	102.4 (91.4–113.4)	98.3 (78.8–129.3)	92.8 (84.5–101.5)	109.1 (89.1–120.6)	84.7 (70.3–94.3)	101.8 (83.3–115.0)	93.3 (72.7–114.0)	108.7 (92.6–121.7)
MUFA (g)	33.7 (30.3–39.5)	29.8 (23.4–34.8)	32.9 (25.3–43.2)	34.4 (28.4–38.1)	35.5 (28.8–40.6)	35.2 (30.8–39.0)	31.5 (29.1–43.8)	32.5 (29.0–36.0)	36.4 (30.2–42.7)	27.5 (23.5–33.2)	34.0 (28.2–39.9)	32.5 (23.9–38.4)	36.5 (33.0–42.3)

Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13
PUFA (g)	21.9 (19.3–24.9)	19.2 (14.0–24.3)	17.6 (13.7–26.3)	22.6 (19.9–24.7)	24.0 (19.1–27.0)	22.1 (18.9–25.6)	24.4 (17.7–35.9)	21.9 (18.7–24.3)	24.5 (21.7–27.3)	18.1 (14.8–21.3)	23.4 (20.0–26.8)	21.4 (16.7–26.1)	24.5 (21.5–28.7)
SFA (g)	33.1 (29.1–36.2)	29.3 (21.8–34.5)	30.6 (22.6–38.8)	31.4 (25.5–36.8)	33.0 (28.0–37.2)	34.4 (28.7–37.8)	32.3 (23.5–38.4)	29.5 (26.2–32.9)	33.2 (28.6–38.8)	26.6 (22.1–32.6)	31.6 (27.2–36.5)	29.4 (23.2–38.1)	32.3 (29.4–40.5)
Dietary Fiber (g)	18.6 (15.7–22.8)	17.9 (14.1–23.0)	14.2 (10.0–17.3)	17.4 (14.6–21.9)	15.9 (13.0–18.5)	17.8 (15.4–20.4)	19.3 (11.3–23.6)	20.1 (18.5–23.3)	18.8 (15.5–22.8)	17.5 (14.1–23.0)	20.2 (17.2–24.7)	13.7 (10.7–20.0)	16.4 (12.8–18.7)
Calcium (mg)	1043 (852–1197)	952 (709–1179)	812 (569–1224)	916 (778–1077)	928 (781–1150)	1052 (942–1219)	1052 (736–1259)	1010 (882–1138)	1076 (834–1286)	948 (839–1093)	1011 (880–1303)	838 (635–1043)	915 (733–1020)
Iron (mg)	16.6 (14.3–18.7)	15.2 (12.1–17.7)	12.5 (9.8–15.2)	14.2 (12.6–16.0)	14.4 (12.2–16.7)	16.8 (15.0–18.5)	18.8 (13.6–24.6)	16.4 (14.4–17.9)	16.9 (13.8–19.5)	15.9 (13.7–17.2)	17.1 (14.6–19.9)	14.1 (10.4–16.7)	14.6 (12.5–15.9)
Sodium (mg)	4037 (3623–4440)	3225 (2824–3964)	3288 (2545–4066)	3814 (3445–4201)	4218 (3656–4636)	4148 (3665–4437)	4126 (3247–5147)	374 (3459–4112)	4083 (3595–4755)	3640 (3188–4232)	4103 (3656–4515)	3656 (3148–4871)	4226 (3655–4777)
Vitamin D (µg)	4.0 (3.0–5.4)	4.0 (2.3–6.8)	3.5 (1.3–5.0)	3.7 (2.5–4.6)	3.1 (2.4–3.9)	4.7 (3.8–5.3)	4.4 (2.1–7.5)	4.3 (3.2–5.0)	4.3 (2.7–6.4)	4.0 (2.6–5.6)	4.8 (3.3–5.9)	2.5 (1.5–4.2)	2.8 (1.9–4.0)

Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13
Age (years)	50.0 (42.0–57.5)	50.5 (38.0–63.0)	58.0 (34.0–64.0)	53.0 (42.0–60.0)	56.0 (43.0–60.0)	56.5 (47.5–63.0)	43.5 (28.7–60.0)	54.0 (46.0–60.0)	54.0 (45.0–62.0)	47.0 (32.0–62.0)	52.0 (40.0–59.0)	47.0 (34.5–64.8)	48.0 (40.0–60.0)
BMI (kg/m^2)	27.7 (25.1–29.6)	27.6 (25.5–30.5)	29.3 (24.5–32.1)	28.0 (25.9–29.4)	26.8 (25.1–29.3)	27.8 (25.3–29.3)	28.5 (25.3–31.2)	27.1 (25.9–28.8)	28.3 (25.7–29.9)	27.5 (25.0–30.9)	27.9 (25.9–28.9)	27.9 (25.3–30.2)	28.2 (26.3–32.3)
WHR	0.96 (0.94–0.99)	0.97 (0.96–1.00)	0.98 (0.91–1.03)	0.98 (0.95–1.00)	0.97 (0.93–0.99)	0.99 (0.95–1.01)	0.96 (0.87–1.01)	0.96 (0.94–0.99)	0.98 (0.94–1.02)	0.97 (0.94–1.00)	0.98 (0.94–1.00)	0.97 (0.92–1.02)	0.99 (0.96–1.02)
Body Fat (%)	28.1 (26.1–30.9)	27.7 (22.2–29.5)	31.15 (17.86–35.2)	27.25 (24.19–29.7)	25.9 (21.5–28.5)	26.15 (22.72–30.07)	25.7 (20.4–28.8)	26.6 (22.86–29.84)	26 (20.8–30.7)	27.9 (22.6–31.2)	26.8 (23.6–30.3)	27.9 (24.6–30.57)	28.6 (22.35–32.3)
Systolic Blood Pressure (mmHg)	124 (120–128)	123 (111–136)	126 (115–133)	126 (121–132)	125 (118–129)	125 (121–130)	126 (112–136)	124 (120–128)	124 (119–132)	120 (113–131)	123 (118–128)	126 (119–133)	122 (118–128)
Diastolic Blood Pressure (mmHg)	76 (71–79)	72 (67–80)	78.5 (70–83)	75 (71–78)	74 (69–80)	73 (70–76)	77 (65–83)	73.5 (71–76)	74 (69–78)	74 (70–82)	73 (69–77)	77 (71–80)	75 (69–78)
Fasting Glucose (mmol/l)	5.9 (5.6–6.3)	5.9 (5.6–6.4)	5.6 (4.7–9.8)	5.9 (5.5–6.3)	5.8 (5.3–6.0)	5.8 (5.6–6.0)	5.5 (5.2–5.8)	5.8 (5.6–6.0)	6.0 (5.5–6.5)	5.8 (5.3–6.4)	5.9 (5.6–6.2)	6.0 (5.6–6.7)	5.7 (5.4–6.1)

Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13
Total Cholesterol (mmol/l)	4.7 (4.1–5.1)	4.6 (4.1–5.1)	4.5 (4.0–5.1)	4.8 (4.4–5.1)	4.6 (4.1–5.1)	4.7 (4.3–5.0)	4.6 (3.9–5.1)	4.6 (4.4–4.8)	4.7 (4.2–5.0)	4.8 (4.2–5.1)	4.5 (4.0–5.0)	4.7 (3.8–5.4)	4.6 (4.2–5.2)
HDL Cholesterol (mmol/l)	1.2 (1.1–1.3)	1.1 (1.1–1.2)	1.1 (1.0–1.2)	1.2 (1.1–1.3)	1.2 (1.0–1.3)	1.2 (1.1–1.3)	1.1 (0.9–1.4)	1.2 (1.1–1.3)	1.2 (1.1–1.3)	1.2 (1.1–1.3)	1.2 (1.1–1.3)	1.2 (1.0–1.5)	1.3 (1.1–1.3)
LDL Cholesterol (mmol/l)	2.7 (2.2–3.3)	2.8 (1.9–3.5)	2.9 (1.5–3.3)	2.7 (2.1–3.1)	2.8 (2.3–3.2)	2.7 (2.3–3.1)	2.7 (1.9–3.6)	2.8 (2.3–3.1)	2.9 (2.2–3.5)	2.8 (2.2–3.4)	2.5 (2.2–3.0)	2.7 (1.9–3.8)	2.6 (2.3–3.1)

Values given are as median and 95% confidence interval in parentheses.

Supplementary Table 6.7: Daily energy intake, nutrient intakes, and health parameters by meal pattern for females with plausible reports of energy intake.

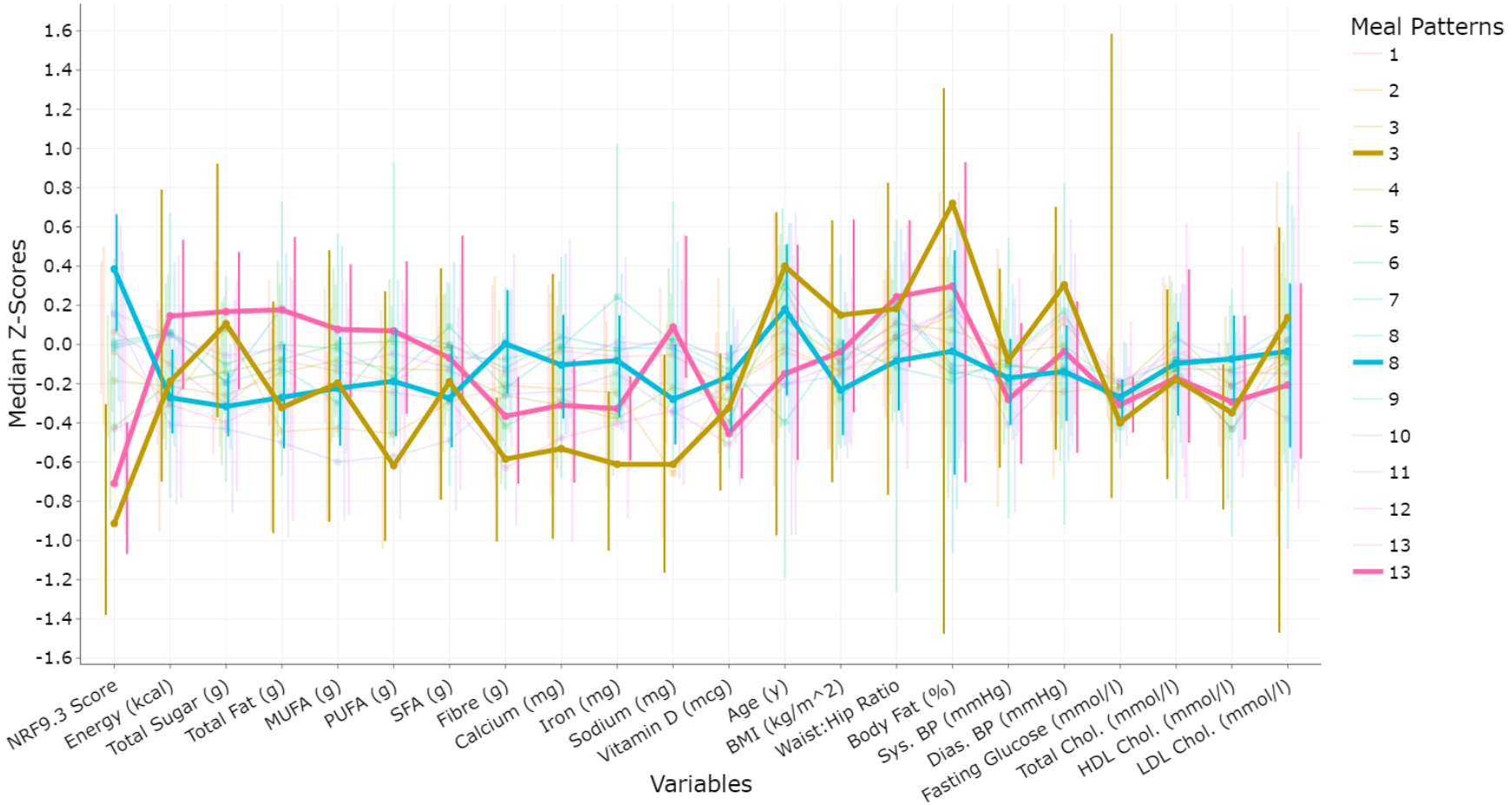
Pattern	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Sample Size	129	124	151	92	169	67	162	113	229	92	124	49	31
NRF 9.3	10.6 (8.3–13.1)	11.5 (9.2–14.1)	17.7 (14.8–19.9)	9.9 (7.2–11.9)	12.3 (9.8–15.1)	12.3 (8.0–17.7)	14.1 (11.7–16.2)	10.5 (7.5–12.9)	14.4 (11.4–17.1)	10.6 (6.4–13.1)	15.9 (14.1–18.8)	16.7 (9.8–21.6)	8.8 (4.5–12.5)
Energy (kcal)	1930 (1742–2154)	2006 (1859–2174)	1809 (1676–1975)	1873 (1700–2182)	1873 (1738–2041)	1721 (1505–1968)	1991 (1764–2140)	1873 (1633–2033)	1918 (1791–2123)	1865 (1712–2090)	1767 (1625–1934)	1614 (1496–2104)	1859 (1416–2092)
Sugar (g)	101.0 (81.7–122.6)	89.6 (75.5–105.0)	86.0 (74.6–104.9)	100.5 (85.2–127.3)	95.9 (77.7–106.3)	98.9 (76.0–113.1)	96.8 (82.9–112.7)	93.7 (76.4–102.0)	94.7 (84.6–110.1)	90.4 (76.6–113.7)	98.0 (79.8–111.0)	63.5 (47.0–109.2)	98.0 (61.3–139.3)
Fat (g)	76.5 (66.0–84.7)	89.7 (77.9–100.1)	75.3 (61.–84.6)	75.8 (66.7–88.5)	78.8 (71.4–87.5)	67.5 (55.3–80.3)	81.7 (73.3–91.3)	77.4 (66.6–88.9)	74.7 (67.1–82.4)	75.9 (66.8–91.5)	71.1 (64.0–77.9)	68.2 (50.9–78.4)	74.5 (54.3–92.0)
MUFA (g)	27.0 (21.4–29.3)	29.8 (26.4–33.8)	24.5 (21.7–28.6)	25.3 (22.3–29.4)	26.8 (24.1–29.7)	21.3 (18.4–27.3)	27.4 (24.6–30.2)	26.8 (23.1–30.7)	24.8 (23.0–27.3)	25.7 (22.4–32.0)	23.9 (21.2–27.2)	22.9 (15.7–27.9)	26.3 (21.9–30.8)

Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13
PUFA (g)	18.9 (16.1–21.2)	20.8 (18.4–23.8)	17.2 (15.4–19.8)	18.2 (16.0–21.6)	19.1 (15.6–21.3)	16.7 (13.1–21.0)	19.6 (16.5–23.5)	19.1 (15.4–21.9)	17.2 (15.1–20.8)	18.1 (15.1–20.9)	17.4 (14.8–18.9)	15.5 (10.6–18.8)	17.3 (12.6–22.6)
SFA (g)	24.9 (20.5–27.7)	28.7 (23.4–32.5)	22.0 (17.8–25.1)	25.4 (21.0–29.3)	25.2 (23.4–28.5)	22.2 (16.3–26.0)	25.4 (22.2–30.0)	25.2 (21.4–27.8)	23.9 (21.8–26.6)	23.5 (21.4–29.2)	22.0 (19.6–26.0)	20.4 (15.4–27.5)	24.2 (17.4–31.4)
Dietary Fiber (g)	15.2 (12.3–18.0)	14.8 (12.4–17.2)	17.3 (14.5–21.0)	14.3 (10.7–16.9)	14.9 (12.8–17.6)	15.7 (10.3–20.2)	16.5 (13.5–18.9)	13.2 (10.9–15.8)	17.1 (14.3–19.0)	12.6 (10.9–15.6)	15.1 (12.7–17.9)	14.7 (11.2–21.5)	12.5 (8.7–16.5)
Calcium (mg)	767 (603–860)	885 (711–1034)	845 (712–911)	767 (644–906)	849 (764–948)	730 (617–965)	863 (783–960)	802 (623–914)	908 (816–992)	699 (570–812)	819 (714–948)	750 (608–899)	745 (464–897)
Iron (mg)	11.1 (9.8–12.5)	11.8 (10.7–13.8)	11.7 (10.6–13.5)	11.6 (9.5–13.8)	12.5 (10.8–14.1)	10.9 (9.0–13.3)	14.3 (12.2–16.6)	10.3 (9.2–12.2)	12.9 (11.2–14.1)	10.5 (9.3–12.2)	11.7 (10.6–13.3)	11.8 (9.9–14.4)	9.7 (6.8–13.6)
Sodium (mg)	3029 (2631–3345)	3343 (3033–3741)	3135 (2738–3367)	2864 (2524–3393)	3139 (2808–3496)	2574 (2072–3018)	3082 (2736–3442)	2924 (2578–3447)	3051 (2683–3299)	2965 (2561–3420)	2747 (2379–3233)	2898 (2489–3733)	2592 (1820–3082)
Vitamin D (µg)	2.4 (1.7–3.1)	3.2 (2.5–4.4)	3.8 (2.7–4.7)	2.4 (1.6–3.5)	3.6 (2.6–4.2)	3.0 (2.0–4.2)	3.9 (2.7–4.6)	2.3 (1.9–3.1)	3.4 (2.8–4.3)	2.4 (1.8–3.9)	3.3 (2.6–4.5)	3.7 (1.7–6.4)	2.8 (1.2–3.9)

Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13
Age (years)	50.0 (40.0–56.0)	49.0 (37.0–58.0)	51.0 (43.0–56.0)	54.0 (46.0–61.5)	44.0 (34.2–54.9)	52.0 (41.0–61.0)	53.0 (41.0–60.0)	51.0 (41.0–60.1)	47.0 (38.0–54.0)	47.0 (35.0–55.0)	59.5 (52.0–63.0)	46.0 (30.0–56.0)	57.0 (22.0–64.6)
BMI (kg/m^2)	27.7 (24.8–29.8)	29.3 (26.6–31.1)	27.2 (25.5–28.9)	28.6 (24.3–31.9)	29.4 (26.2–31.7)	28.7 (24.2–32.3)	28.3 (26.2–31.3)	28.2 (25.8–31.4)	27.9 (25.8–29.5)	27.5 (24.5–30.9)	28.0 (25.4–30.1)	25.2 (21.2–28.3)	28.5 (23.9–33.2)
WHR	0.89 (0.87–0.92)	0.90 (0.88–0.93)	0.91 (0.88–0.93)	0.92 (0.88–0.95)	0.91 (0.88–0.93)	0.90 (0.85–0.94)	0.90 (0.88–0.92)	0.91 (0.87–0.94)	0.90 (0.87–0.92)	0.89 (0.86–0.92)	0.91 (0.87–0.93)	0.88 (0.81–0.91)	0.90 (0.83–0.94)
Body Fat (%)	38.0 (33.6–40.9)	38.0 (34.0–41.4)	38.8 (35.0–41.3)	37.9 (34.4–41.2)	38.4 (33.6–41.7)	36.6 (31.8–42.9)	36.7 (32.3–41.9)	39.2 (35.7–42.2)	39.2 (35.2–41.1)	39.9 (33.7–42.6)	40.8 (35.5–42.5)	37.4 (30–41.4)	40.1 (27.2–45.0)
Systolic Blood Pressure (mmHg)	119 (111–128)	119 (108–125)	115 (109–120)	117 (109–131)	115 (110–126)	124 (114–136)	118 (109–125)	114 (108–123)	118 (112–123)	111 (104–120)	122 (112–130)	112 (105–133)	118 (106–136)
Diastolic Blood Pressure (mmHg)	74 (69–77)	74 (68–78)	71 (67–75)	71 (67–79)	72 (67–76)	71 (65–79)	72 (67–76)	71 (66–75)	71 (68–74)	72 (66–75)	73 (69–75)	72 (66–79)	73 (67–81)
Fasting Glucose (mmol/l)	5.7 (5.3–60)	5.6 (5.3–5.9)	5.6 (5.3–5.9)	5.8 (5.2–6.1)	5.6 (5.3–5.8)	5.7 (5.4–6.2)	5.7 (5.3–6.1)	5.4 (5.2–5.6)	5.6 (5.3–5.9)	5.6 (5.3–5.6)	5.7 (5.2–6.2)	5.6 (5.1–6.4)	5.5 (4.9–5.9)

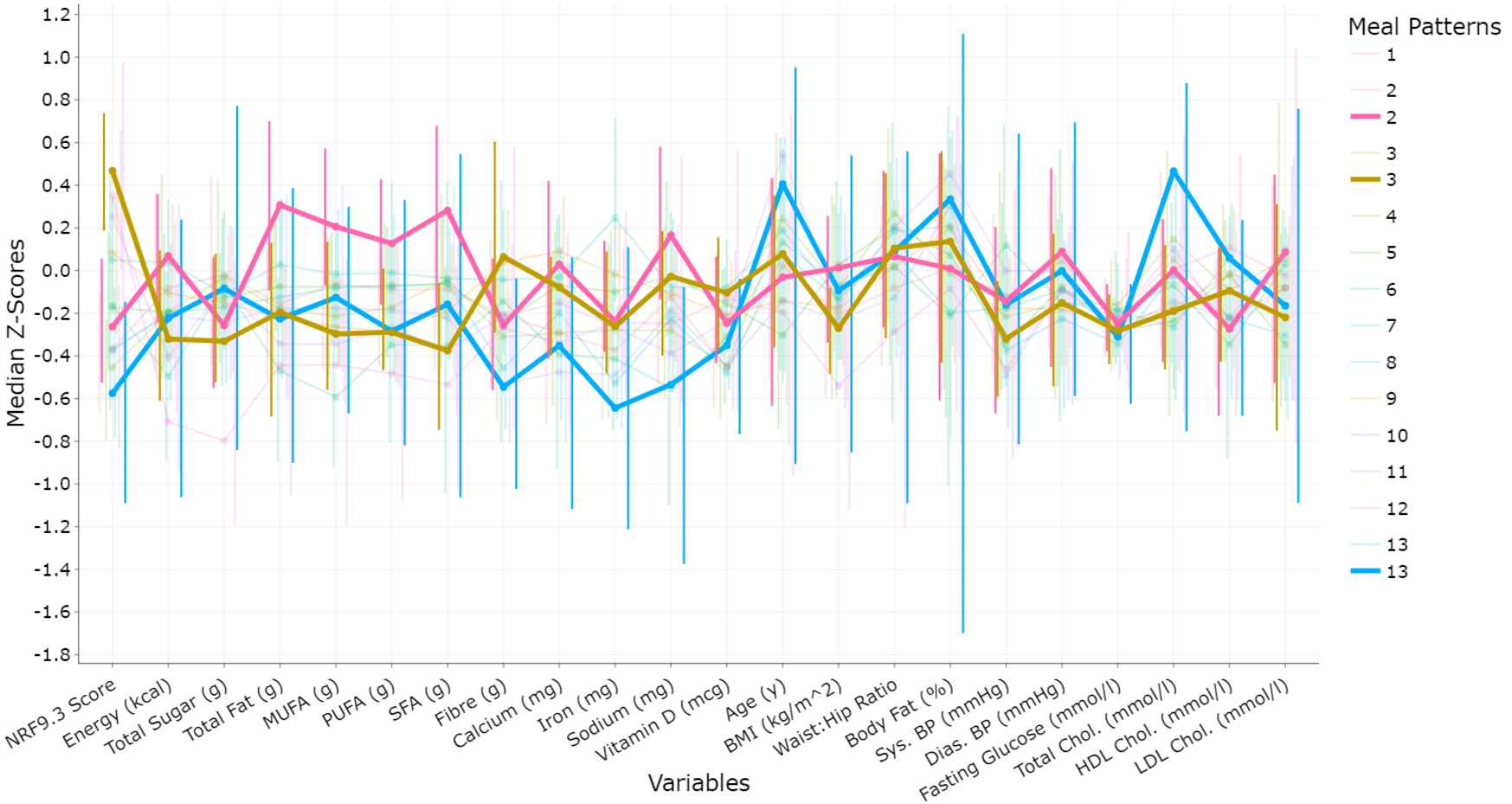
Pattern	1	2	3	4	5	6	7	8	9	10	11	12	13
Total Cholesterol (mmol/l)	5.0 (4.6–5.5)	5.0 (4.5–5.3)	4.8 (4.5–5.1)	5.1 (4.6–5.6)	4.7 (4.3–4.9)	4.9 (4.5–5.5)	4.9 (4.4–5.1)	4.9 (4.3–5.4)	4.8 (4.5–5.1)	4.8 (4.5–5.2)	5.1 (4.6–5.4)	5.0 (4.3–5.6)	5.5 (4.2–5.9)
HDL Cholesterol (mmol/l)	1.4 (1.3–1.5)	1.4 (1.2–1.6)	1.5 (1.3–1.6)	1.5 (1.3–1.6)	1.5 (1.3–1.6)	1.4 (1.1–1.5)	1.5 (1.3–1.6)	1.4 (1.3–1.6)	1.5 (1.3–1.6)	1.4 (1.2–1.6)	1.4 (1.3–1.5)	1.6 (1.3–1.7)	1.5 (1.2–1.6)
LDL Cholesterol (mmol/l)	2.8 (2.3–3.4)	2.9 (2.2–3.3)	2.6 (2.2–3.1)	2.9 (2.5–3.6)	2.5 (2.2–3.0)	2.9 (2.4–3.1)	2.7 (2.3–3.0)	2.6 (2.2–3.0)	2.8 (2.4–3.1)	2.9 (2.3–3.3)	2.8 (2.3–3.3)	2.8 (2.1–3.8)	2.7 (1.8–3.5)

Values given are as median and 95% confidence interval in parentheses.



Supplementary Figure 6.1: Parallel plot of median z-scores and 95% confidence intervals for selected nutrients and health parameters within each meal pattern in males.

Meal patterns in which differences among meal patterns for nutrient quality were observed are highlighted.



Supplementary Figure 6.2: Parallel plot of median z-scores and 95% confidence intervals for selected nutrients and health parameters within each meal pattern in females.

Meal patterns in which differences among meal patterns for nutrient quality were observed are highlighted.

CHAPTER 7

Dietary Intake Assessment Using a Novel Generic Meal-Based Recall and a 24-Hour Recall: A Comparison Study

A version of this chapter has been published as:

O'Hara C and Gibney ER (2024) Dietary Intake Assessment Using a Novel Generic Meal-Based Recall and a 24-Hour Recall: A Comparison Study. *J Med Internet Res*, **26**, e48817.

7.1 Abstract

Dietary intake assessment is an integral part of addressing sub-optimal dietary intakes. Intakes of whole meals can be reported in a way that is less burdensome than reporting individual foods, yet no meal-based dietary recalls have been developed. The objective of this study was to develop a novel meal-based method of dietary intake assessment and compare its ability to estimate nutrient intakes with a web-based 24-hour recall (24HR). Participants completed a web-based generic meal-based recall. This involved, for each meal type (breakfast, light meal, main meal, snack, beverage), choosing, from a selection of meal images, the ones that most represented their intakes during the previous day. Meal images were based on generic meals from previous research that were representative of actual meal intakes in Ireland. Participants also completed a web-based 24HR. Both methods were completed on the same day, three hours apart. In a crossover design, participants were randomised as to which method they completed first. Two weeks later, participants repeated the process in the reverse order. The study was completed by 161 participants. For the 26 nutrient variables compared, the median percentage difference between the two methods was 7.4% with *p* values ranging from <0.001 to 0.965, and effect sizes for the differences were small for 21 variables, moderate for three variables, and large for two variables. Correlation coefficients were statistically significant ($p < 0.05$) for 21 of the 26 variables. Statistically significant correlations ranged from 0.16 to 0.45 with a median correlation of 0.32. In conclusion, a generic meal-based method of dietary intake assessment provides comparable estimates of nutrient intake to a web-based 24-hour recall, but with varying levels of agreement among nutrients. Further work is required to refine the generic recall across a range of nutrients and to consider the user experience of the generic recall.

7.2 Introduction

Well-established causal relationships exist between dietary intakes and health [1]. Accurate dietary intake assessment is required to identify suboptimal intakes and to devise interventions to address these [2]. Existing food-based methods of dietary intake assessment can be time-consuming and burdensome for individuals to complete [3]. In addition, not all methods provide information such as the timing of meals, the different foods that are eaten in combination as part of these meals, or the combinations of different meals over a day, i.e., the food frequency questionnaire instead focuses on mean daily food and nutrient intakes [4].

The time and effort required for individuals to complete existing methods of dietary intake assessment such as food frequency questionnaires, 24-hour recalls (24HR), and food diaries may limit adherence to and engagement with such methods, which are often used on web- and mobile-based personalised nutrition platforms [3]. Indeed, a survey of 2382 adults across Europe found that ease of use was the most important feature for participants when choosing a nutrition or diet application [5]. Digital versions of the 24HR and food diary require that the user text-search for a food and then select from a list of results the food that they consumed. This process is then repeated for each food in the meal and each meal in the day [6]. Intakes of whole meals can be recorded in a way that is less burdensome than recording individual foods, providing a potentially lower-burden method for dietary intake assessment [3]. For example, instead of text searching for individual foods, as is required in the 24HR and food diary, the user could be presented with images of whole meals and choose the image most similar to their meal [3]. The use of meal-based methods may also be preferential in personalised nutrition because people tend to perceive their dietary intakes in terms of the meals they have consumed rather than their daily intakes of nutrients or foods, so recording dietary intakes and providing dietary advice in this manner may be more intuitive [7, 8].

While the number of studies examining meal patterns has increased in recent years [4], only three [9-11] have developed a meal-based method of dietary intake assessment rather than using existing food-based methods. Englund-Ögge *et al.* [9] used a method whereby participants reported how often they consumed various meal types (breakfast, morning snack, lunch, afternoon snack, dinner, evening snack, supper, and night meal). Wilson *et al.* [10] had a similar approach, but instead of using meal types divided the day into time periods and asked participants to report for each one whether they ate nothing, a snack, a small meal, or a large meal and whether they drank nothing, alcohol, water, or something else. These approaches to meal-based dietary assessment are simple to complete and

provide qualitative information on meal types and their timing. They do not, however, provide the qualitative detail necessary to identify the different combinations of foods being consumed in meals and the combinations of those meals over a day or the quantitative detail required to estimate the nutrient intakes arising from those consumptions. Murakami *et al.* [11] developed an approach that involves participants reporting the frequency of consumption of combinations of food groups and foods at specified meal types (breakfast, morning snack, lunch, afternoon snack, dinner, night snack), and has been designed for use in the Japanese population. This approach does allow for the identification of meal patterns and nutrient intakes, but still requires individuals to report intakes at the food level. None of those studies, however, allow for the reporting of meal portion sizes or capture information from the previous 24 hours in the form of a recall.

Several tools have been developed that use image recognition software for dietary intake assessment [12]. However, these methods remain food-based rather than meal-based. The software first segments a meal image into its constituent foods, and then provides a suggested match for each food in the image. The user must then confirm that the suggested foods are correct. For any missing or incorrect foods, the user must text-search for the correct food and add this to their record [12, 13]. No study has been identified that allows individuals to record their dietary intakes at the meal level by reporting intakes of whole meals rather than individual foods or food groups.

The aim of this study was to develop a novel meal-based method of dietary intake assessment that would allow individuals to report their intakes of whole meals rather than reporting the individual foods that make up those meals, and to compare this method with a web-based 24HR.

7.3 Methods

7.3.1 Recruitment

The Human Research Ethics Committee, University College Dublin granted ethical approval to conduct this study (LS-21-64-OHara-Gibney). The target sample size was 160 participants based on a previous review of comparisons between digital and paper-based 24-hour recalls which found a range of sample sizes from 53 to 167 [14]. There were no studies on meal-based methods of dietary assessment on which to base the sample size.

Recruitment was carried out using local radio, local newspaper, posters, social media, and word of mouth. Researchers directed potential participants to a webpage containing full

details of the study requirements and contact details of the researchers for further queries if required. After reading the study information potential participants could indicate whether they had read and understood the material and agree or disagree to proceed to the online screening questionnaire to determine their eligibility to take part in the study. Those who were eligible then completed an online consent form to provide electronic informed consent. People were eligible if they were over 18 years of age, were fluent in English, had regular access to the internet, and were not current or former students of a degree in nutrition or dietetics (**Figure 7.1**).

7.3.2 Study Design

Participants who were deemed eligible to take part and provided consent, were contacted by email providing details on the next steps of the study and the links to the two web-based dietary intake assessment tools (the 24HR and the generic meal-based recall). A crossover design was used with regard to the order in which participants completed the recalls. Participants were randomised to complete one or other of the two methods first, and then complete the second method at least 3 hours later on the same day. Participants were also randomised as to whether they would recall a weekday or a weekend day. Two weeks after having completed the first set of recalls, participants completed the recalls again in the reverse order to which they were completed on the first occasion, followed by the completion of the evaluation questionnaire (**Figure 7.1**).

7.3.3 Overview of the Generic Meal-Based Dietary Recall

The meal-based dietary intake assessment method was administered using Qualtrics® XM, a web-based platform for questionnaires. Prior to completing the dietary intake assessment, participants provided information on their sex, age, weight, and height via an online questionnaire. Participants were then asked to select the meal types they consumed on the previous day from the following list: breakfast, morning snack, lunch, afternoon snack, evening meal, evening snack, and beverage (for beverages consumed alone without food). For each meal type selected, participants were presented with a series of images of generic meals that are associated with that meal type and asked to choose the meal image that was most similar to the meal that they consumed (**Figure 7.2**). For each selected meal, participants were asked to choose from three different images of that meal, each representing a different portion size, and then asked to select whether the chosen image was smaller than the chosen portion, the same size as the portion chosen, or larger than the portion chosen (**Figure 7.2**). For beverages, participants were asked to choose from three different images of that beverage representing three different portion sizes and then

asked to select the number of those portion sizes that they consumed. For each meal type, participants could choose the option that none of the images presented were representative of their intake for that meal type, e.g., “none of the above options are similar to what I ate for breakfast”. If participants selected this option, a box appeared to enter a free text description of what they consumed for that meal. This allowed the researchers to determine whether a suitable generic meal could have been chosen or whether there was in fact no matching generic meal.

7.3.4 Development of the Meal-Based Dietary Recall

The process of identifying the generic meals that were presented to participants as images has been described in detail elsewhere [15]. In brief, data from the Irish National Adult Nutrition Survey (NANS) (2008–2010) [16] were used. This is a representative dataset on dietary intakes of 1500 adults in Ireland collected using four-day weighed food diaries. The meals reported were categorised into the following meal types: breakfasts, light meals, main meals, snacks, and beverages. A nutrient profiling score, the Nutrient Rich Foods Index (NRF9.3) [17], was calculated for each meal. Within each meal type, meals were grouped/clustered using partitioning around the medoids (PAM) clustering to identify groups of meals that had similar NRF9.3 scores and food group composition. These groups were defined as generic meals.

Of the 27,336 individual meals consumed by the participants in the NANS research, these were condensed to 63 generic meals; 31 of these were consumed during the week and 32 during the weekend. Of the 63 generic meals that were identified, there was overlap between weekday and weekend meals. That is, some weekday meals were the same as the weekend meals. When these duplicates were removed, participants were presented with 43 meal images: 5 breakfasts, 5 snacks (repeated for morning, afternoon, and evening snacks), 10 lunches, 19 dinners, and 4 beverages.

The nutrient content of a given generic meal was defined as the mean nutrient content of the individual meals that made up that generic meal per 100g. Each generic meal was assigned seven portion sizes by ordering each of the individual meals by weight and dividing the meals into sevenths based on septile values for meal weight. The median weights for each seventh were assigned as the generic portion size for that meal. The nutrient composition for each of the portion sizes was calculated using the meal weight for that portion and the generic meal nutrient composition [15]. Within the meal intake assessment tool presented in the current study, the second, fourth, and sixth portion size were used as the three portion size images shown to participants, with the options asking

whether the image chosen was smaller, the same, or larger than that consumed allowing participants to be categorised as consuming the first, third, fifth, or seventh portion size for a given meal.

7.3.5 24-Hour Recall

Participants completed their 24-hour recalls (24HR) using a validated web-based self-administered 24HR tool called Foodbook24 which follows the multi-pass recall method [18-20]. Participants first chose the meal types that they consumed from the following list: breakfast, morning snack, lunch, afternoon snack, evening meal, and evening snack with the option to add additional snacks. For each of the selected meal types, participants added the foods and beverages they consumed at those meal types by text searching from the food list using a search bar. Portion size was then reported based on the number of the food or beverage items consumed and/or from portion size photographs as appropriate. Participants were then presented with the list of foods they had recorded for review, before being presented with a list of commonly forgotten foods. The food list contained food composition data from McCance and Widdowson's The Composition of Foods Integrated Dataset 2021 with some additions relevant to dietary intakes in Ireland. The development of Foodbook24 and its food list is described in detail elsewhere [18, 20].

7.3.6 Statistical Analysis

All analysis was carried out using R version 4.2.2 [21] in the RStudio integrated development environment (version 2022.07.2+576) [22]. Data from the 24HRs were used to identify participants likely to be mis-reporters of energy intake, based on a ratio of estimated energy intake to basal metabolic rate (EI:BMR) using the BMR equations from Henry [23]. Based on the Goldberg equations [24], an EI:BMR less than 0.96 was deemed indicative of under-reporting, with a ratio greater than 2.49 indicative of over-reporting. The analysis presented in the current study includes all participants, given the negligible differences observed when mis-reporters were removed; the analysis of the smaller dataset with mis-reporters excluded is given in the Supplementary Tables 7.1–7.4.

The Shapiro-Wilk test was used to determine whether the differences in the variables between methods were normally distributed and confirmed with visual inspection of histograms. Wilcoxon's signed rank test was carried out to compare nutrient intake estimates arising from the web-based 24HR with estimates arising from the generic meals-based recall. Wilcoxon's effect size *r* was calculated. An effect size of ≥ 0.1 and < 0.3 was considered small, ≥ 0.3 and <0.5 was considered moderate, and ≥ 0.5 was considered large

[25]. Bland–Altman analysis was also carried out whereby the mean difference between the 2 datasets and the limits of agreement (mean difference ± 1.96 SD) for each nutrient were calculated. The correlation of nutrient intakes between the two methods was assessed using Spearman rank correlation coefficients. A correlation coefficient of <0.20 was considered poor correlation, a coefficient of ≥0.20 and <0.50 was considered acceptable correlation, and a coefficient of ≥0.5 was considered good correlation [26].

Cross-classification of quartiles was performed for all nutrients assessed. That is, nutrient intakes from both methods were divided into quartiles to determine the proportion of participants who remained in the same quartile for both methods (exact agreement), the proportion who were classified in the same or adjacent quartiles (exact + adjacent), the proportion who were classified 2 quartiles apart (disagreement), and the proportion who were classified 3 quartiles apart (extreme disagreement). Participants were also classified, separately for both methods, according to nutrient-based dietary guidelines [27-29]. For example, they were classified as to whether their nutrient intakes were low, adequate, or high according to those guidelines. The nutrients assessed were those deemed to be of public health relevance and included protein, carbohydrate, fat, monounsaturated fat, polyunsaturated fat, saturated fat, salt, dietary fibre, calcium, iron, folate, thiamine, riboflavin, and vitamin C. The proportion of individuals who were classified into the same category based on both methods was calculated for each nutrient.

7.3.7 Evaluation Questionnaire

Participants also completed an evaluation questionnaire at the end of the study which was administered via Qualtrics® XM. Participants were asked whether the meals presented in the meal-based dietary intake assessment tool were representative of what they consume, to rate the clarity and ease of use of the tool, and whether they preferred that tool to the food-based tool.

7.4 Results

7.4.1 Study Sample

A total of 161 participants completed both methods of dietary intake assessment at two timepoints. Most participants were female (81.4%), and the median and interquartile range for age was 54 (39–63) years, and median BMI was 25.3 (22.5–28.9) kg/m^2 (**Table 7.1**).

7.4.2 Daily Nutrient Intakes

For the 26 variables compared, the percentage difference between the meal-based and 24HR methods ranged from 0.0% to 46.7%, with the median percentage difference being 7.4% and with the generic method providing a higher estimate than the 24HR for 18 nutrients. *P* values for the differences between the two methods ranged from <0.001 to 0.965, with 14 of 26 comparisons reaching statistical significance ($P < 0.05$). Effect sizes for the differences were small for 21 variables, moderate for two variables (folate in µg and sodium in mg), and large for two variables (polyunsaturated fat in g and as %TEI) (**Table 7.2**).

Comparing differences using the Bland–Altman analysis, the mean differences between the two methods for macronutrients were close to zero, whereas those for the micronutrients were larger. The limits of agreement tended to be wide for most nutrients. The analysis identified 17 nutrients for which ≥95% of participants fell within the LOA. The proportion of individuals that fell within the LOA ranged from 92.5% for polyunsaturated fats (%TEI) to 98.1% for dietary fibre (g). Bland-Altman plots for energy and the macronutrients are presented in **Figure 7.3**; values from the Bland-Altman analysis for the remaining nutrients are given in **Table 7.3**.

Correlation coefficients were statistically significant ($P < 0.05$) for 21 of the 26 variables. Those where no significant correlation was identified were total fat (%TEI), monounsaturated fat (%TEI), polyunsaturated fat (%TEI), vitamin D (µg), and sodium (mg). For those where a correlation was identified, they ranged from 0.16 for saturated fat as %TEI to 0.45 for sugar in grams with a median correlation of 0.32 (**Table 7.4**).

7.4.3 Categorisation of Daily Nutrient Intakes

Cross-classification of quartiles are also presented in **Table 7.4**. The proportion of individuals remaining in the same quartile ranged from 23% for total fat (%TEI) to 39% for carbohydrates (g) with a median of 32%. Three nutrients had extreme disagreement for ≤5% of participants (protein in both grams and %TEI, and sugars as %TEI), 19 had extreme disagreement for between >5% and ≤10% of participants, and three had extreme disagreement for >10% of participants (monounsaturated and polyunsaturated fat as %TEI, and vitamin D in micrograms). When participants were classified according to nutrient-based guidelines (e.g., low, adequate, or high) for the 14 nutrients, the proportion that were classified into the same category by both methods ranged from 52.8% for total fat (%TEI) to 84.5% for protein (g/kg bodyweight) (**Table 7.5**). The median proportion classified correctly among the 14 nutrients was 70.5%.

7.4.4 Participant Evaluation Questionnaire

The majority of participants either somewhat agreed or agreed in relation to the generic recall that the instructions provided were clear and easy to understand (91%), that the portion sizes in the tool were largely representative of their portion sizes (85%), and that overall, the tool was easy to use (81%). However, a majority reported that the meal images were not representative of what they actually consumed (55%). This was anticipated, and for each meal in the generic recall, participants were given the option to select that no meal image presented to them was similar to what they had consumed. This option was chosen for 36.0% (n = 683) of the total of 1895 meals consumed, broken down across the meal types as follows: 17.4% (n = 119) of breakfasts, 24.7% (n =169) of light meals, 20.4% (n = 139) of main meals, 29.3% (n = 200) of snacks, and 8.2% (n = 56) of beverages. When these choices were reviewed by the researchers, comparing participants' text descriptions of their meals with the possible options from generic meal images, it was determined that 593 of the 683 meals were correctly recorded by participants as not having a matching generic meal, whereas 90 meals could have been matched to one of the generic meal images. While a majority reported that the ease of use of the meal-based tool was worse or somewhat worse than the 24HR (66%), 39% of participants reported that they would consider using a similar tool to the generic recall again.

7.5 Discussion

The current study reports on a novel meal-based recall that allows individuals to report their intakes of whole meals rather than reporting individual food or food group intakes for each meal, as is necessary in commonly used traditional dietary intake assessment methods. Estimated nutrient intakes from the generic meal-based method were comparable with those from the 24HR for some, but not all nutrients. Comparisons between the methods were more similar at the group level compared to the individual level. Participants found the meal-based method understandable and easy to use. Previous studies have identified generic meals that exist in national dietary survey data [15, 30-32], but the current study is the first to use images of those generic meals as a novel method of dietary intake assessment.

In previous studies on generic meals, comparisons were made between estimated nutrient intakes from the original data and from the generic data. Agreement between the original and generic data was stronger in those studies than the agreement between the two methods of dietary assessment presented in the current study; however, this is expected because those studies compared intakes arising from generic meals with intakes arising from the original data from which those generic meals were derived [15, 30-32]. In the current

study, comparisons were made on data collected using two different methods, therefore the generic meal images presented to participants were not influenced by those participants' food intakes. Agreement between the methods in the current study varied depending on the nutrient in question. In general, percentage differences and effect sizes for differences between the two methods were small for most nutrients. Certain nutrients showed poorer agreement than others including total fat, polyunsaturated fats, monounsaturated fats, and vitamin D. A number of features of the generic method may have given rise to differences in these fat-soluble nutrients. In the food-based 24HR, participants could specify the types of fats that they added to foods; however, in the generic method that was not the case as participants had to choose between pre-defined generic meals. Some of these nutrients are found in relatively high concentrations in relatively few foods, which may also give rise to differences between a food-based and a generic method. These trends have also been observed previously when comparing FFQs with 24HRs and food records, as noted by Cui *et al.* [33] in a meta-analysis of 130 such studies. In the case of polyunsaturated fats, the difference between the median intakes was considerably larger than other nutrients. This was identified as having arisen from food composition differences between the generic data and the 24HR data. The generic meals used in the generic recall and their nutritional composition were derived from data from the National Adult Nutrition Survey (NANS) in Ireland [15, 16]. The composition data in that survey in turn came from a variety of sources including food packaging, industry information, published papers and various food composition tables including those from the UK, Finland, and Australia that were published between 2002 and 2010 [34]. The composition data used in the 24HR came from the 2021 publication of McCance and Widdowson's Composition of Foods Integrated Dataset (CoFID) [35]. On further examination of polyunsaturated fat content of individual foods from both NANS and CoFID, a number of foods were identified in NANS that had considerably higher values for polyunsaturated fat than those corresponding foods in CoFID in a way that was not evident for other nutrients. This was observed across a range of foods, but it was most pronounced in foods that had a value of zero for polyunsaturated fat content in CoFID (56% of foods consumed by participants) and a value greater than zero in NANS (20% of foods present in generic meals consumed by participants). Examples include raw tomatoes, lettuce, and carrots in which polyunsaturated fat content from NANS data ranges from 0.2 to 0.4g per 100g.

While other studies have not examined the potential for dietary assessment based on individuals' reporting of whole meal intakes, comparisons can be drawn, however, with other related methods. Murakami *et al.* [36] developed a meal-focused method called the Food Combination Questionnaire (FCQ). In this method participants reported, for each meal

type (breakfast, lunch, dinner, snacks), what staple foods they consumed, what accompanying foods they ate with those staple foods, and how frequently per week they consumed them during the preceding month. Agreement for estimates of nutrient intake between that method and a 4-day weighed food record were similar to agreement reported in the current study for the comparison between the generic recall and the 24HR. The median value of the statistically significant correlation coefficients was 0.35 compared with 0.31 in the current study. In that study however, the FCQ estimated lower intakes for most nutrients compared to a food record, whereas in the current study the generic recall estimated higher intakes for most nutrients compared to a 24HR.

Another approach to dietary intake assessment reported by Katz *et al.* [37] involves participants reporting their intakes at the dietary pattern level instead of the meal, food group, or food levels. In this format participants are presented with two images at a time. Each image contains multiple different foods and meals representing different dietary patterns based on the Healthy Eating Index, and the participant must choose which image is more representative of their dietary intakes. This process of choosing between two different diet pattern images is repeated until a "best fit" is identified for the individual [37]. Comparison of estimated nutrient intakes arising from this method with those from three separate 24HRs resulted in a median correlation of 0.30 for correlations that were statistically significant [38]. Like the current study, that diet pattern approach also tended to estimate higher nutrient intakes compared to the 24HR. Similar trends were also observed in the Bland-Altman analysis with low bias, or systematic error, for the macronutrients (i.e., mean differences were close to zero), but slightly greater bias for the micronutrients. Wide limits of agreement were also observed in both studies indicating considerable random error. Random error can be reduced through repeated measures [39]; therefore, conducting more than the two generic recalls carried out in the current study may mitigate against the random error observed here. Another trend that is evident from the Bland-Altman plots for macronutrients expressed as %TEI, is that the generic recall gives higher estimates of intake than the 24HR when intakes are low and gives lower estimates of intake when intakes are high. This arises from the narrowing of the distribution of intakes (i.e., reduced variance) in the generic method where participants can only choose from a small number of generic meals compared to the practically limitless different combinations of foods that participants can choose from the 24HR. The use of portion sizes mitigates against this trend for values expressed in absolute intakes, e.g., grams of macronutrient intake. However, expressing those values relative to energy intake effectively cancels out the increased variance introduced by portion size giving rise to the trends observed in the Bland-Altman plots for values given as %TEI but not for those given as absolute intakes.

A version of the FCQ described above has been used to provided dietary advice in a pilot study by Murakami *et al.* [40]. Similarly, the dietary pattern-based method of dietary assessment described above has also been incorporated into a commercially available tool that provides nutrition recommendations to users [41], although the recommendations aspect of this tool has not been described in the scientific literature. It has also been proposed that a meal-based method of dietary intake assessment could provide a lower burden alternative to gathering dietary data for personalised nutrition advice [3]. The current study presents the first attempt at implementing such a proposal. Its performance in estimating nutrient intakes is similar to other approaches that focus on meals or dietary patterns. However, these methods tend to have poorer agreement compared to that between FFQs and 24HRs where median correlation for the various nutrients assessed of 0.42 has been reported [33] or compared to the agreement between 24HRs and diet records where median correlations in the range of 0.45 and 0.50 have been reported [42, 43]. Further work is required to appraise the performance of these methods, including the generic recall, against objective biomarkers of dietary intake.

Many of the comparisons between the two methods in the current study are based on their ability to provide point estimates of nutrient intakes, with agreement for point estimates of nutrient intake typically stronger at the group level compared to the individual level. This trend has also been observed in comparisons of other methods of dietary intake assessment, for example between the 24HR and diet record [42], between the FFQ and diet record [44], and between the FFQ and 24HR [45]. In the current study, this is to be expected given that there are fewer generic meals than foods for participants to choose, i.e., the variance has been intentionally reduced, but reduced in a way that is systematic and consistent across meals [15]. This trend has been observed not just in comparison studies of different methods of dietary assessment, but also in studies which have used generic meals as a method for secondary analysis of dietary data that have already been collected [15, 30-32]. Approaches to personalised nutrition facilitated by technology, however, do not rely solely on point estimates of nutrient intake. Instead, individuals are categorised into ranges of nutrient intakes (e.g., low, adequate, or high), allowing room for error in those point estimates [40, 46]. In the Food4Me study on personalised nutrition, for example, participants were categorised in this way and dietary advice reports were tailored depending on the participants' categories for various nutrients [40, 46]. The current study has shown that individuals can be classified according to nutrient-based dietary guidelines using the generic recall. Similarly, this approach can be used to rank individuals based on their nutrient intakes, with values from the cross-classification of quartiles being comparable to

those observed in studies comparing FFQs with 24HRs where median exact + adjacent agreement has been reported between 66% and 86.1% [47-49].

Of promise with regard to the use of the generic recall in dietary assessment in personalised nutrition is that the majority of participants reported in the evaluation questionnaire that the recall was easy to use. Ease of use has previously been reported as an important factor in the choice of nutrition or diet mobile applications among a cohort of 2382 adults in Europe [5]. On the other hand, however, two thirds of participants reported that the ease of use of the 24HR was better than that of the meal-based approach. This is understandable given that the web-based 24HR used in the current study was developed as a stand-alone platform in collaboration with software developers [20], whereas the generic recall was a concept or pilot tool implemented by the authors using a commercially available questionnaire platform, and not specifically designed for user experience. Future user evaluations and collaboration with software development professionals could further enhance the user experience of the generic recall, including incorporation of meal image recognition.

To date much research has been carried out on image-based food recognition using computer vision as a means of reducing the burden of data input for food-based methods of dietary intake assessment [13]. These image recognition tools, however, are food-based, insofar as when an individual takes a photo of their meal, the software segments the image and provides a suggested match for each of the individual foods that make up the meal [13]. The user must then confirm that each of the suggested foods are correct. For any missing or incorrect foods, the user must text-search for the correct food and add this to their record [12]. It is possible that a meal-based approach could be taken to image recognition in dietary assessment, removing the need to identify individual foods. Instead, the software would classify a whole meal image as one of the generic meals used in the current study. Further work is required, however, to determine the feasibility of such an approach for meal-based image recognition in nutrition.

A number of limitations exist in the current study. The generic meal images used are based on dietary intake data from the 2008–2010 national adult nutrition survey [16]. While, at the time of writing, these were the most recently published intake data in Ireland, the generic meals do not account for any changes in food composition or meal intakes that may have occurred since that time. This may account for the findings in the evaluation questionnaire of the current study, with 55% of participants reporting that the generic meals were not representative of their intakes and shows the need to ensure that generic meals are revised using most recent data. This may be of particular importance if the tool was to be used with

a younger cohort of participants than those who participated in the current study. The current study also used a convenience sample rather than one that is representative of the population in Ireland. This may result in the recruitment of participants who are more likely to have an interest in nutrition and health. The unbalanced nature of the demographics of the study participants precludes subgroup analysis in relation to demographic factors.

The two types of recall used in the current study are impacted by measurement error and only provide an estimation of true intakes. The current study aims to compare these two methods and, therefore, the reported statistics should be interpreted as representing the relationship between the generic recall and the 24-hour recall and not the relationship between the generic recall and true intakes. A comparison study was deemed more appropriate as an initial indication of the generic recall's strengths and weaknesses prior to considering more labour intensive and costly objective measures of comparison such as the feeding studies or biomarkers of dietary intake.

The strengths of the current study include its sample size of 161 participants. The randomisation of participants to which recall method they would complete first, the reversal of the order of completion in their second set of recalls, and the two-week washout period mitigates against any learning effect that participants may have experienced after completing the first recall. The comparison method used, Foodbook24, is a validated web-based 24HR [20]. The current study also captured dietary intakes on both weekend and weekdays, accounting for the potential differences in eating habits that occur between the two [50-52].

7.6 Conclusion

A generic meal-based method of dietary intake assessment provides comparable estimates of nutrient intake to a web-based 24HR. The agreement ranges among nutrients from weak to moderate, with better agreement at the group level than the individual level. Further work is required to improve this method of dietary assessment considering the number of recalls required and the use of more recent dietary intake data to define the generic meals. Future work will determine the feasibility of taking a meal-based approach to image recognition in dietary intake assessment.

7.7 Tables

Table 7.1: Participant demographics and anthropometry.

	Female	**Male**	**Total**
	n = 131 (81.4%)	n = 30 (18.6%)	n = 161 (100%)
Age	54 (41 – 63)	54 (37 – 63)	54 (39–63)
Weight (kg)	67.9 (60.5 – 75.6)	80.8 (75.6 – 94.0)	70 (62–80.8)
Height (m)	1.63 (1.59 – 1.68)	1.8 (1.75 – 1.83)	1.65 (1.60–1.74)
BMI (kg/m^2)	25.2 (22.5 – 28.8)	26.0 (22.9 – 28.8)	25.3 (22.5–28.9)

Variables are given as median (25^{th} percentile – 75^{th} percentile). Weight and height were self-reported by participants and BMI subsequently calculated from those self-reported values.

Table 7.2: Median (interquartile range) daily nutrient intakes estimated from a web-based 24-hour recall and novel generic meal-based recall.

	Recall Method		% Difference	*P* value	Effect Size (*r*)	Effect Size (Magnitude)
	24HR	Generic				
Energy (kcal)	1569.9 (1318.9–2005.7)	1715.3 (1440.9–1935.3)	9.3	0.176	0.107	small
Protein (g)	65.4 (51.4–86.8)	72.7 (63.2–82.1)	11.2	0.685	0.032	small
Protein (%TEI)	16.6 (13.8–20.1)	16.8 (15.8–18.3)	0.8	0.445	0.060	small
Carbohydrate (g)	192 (140.5–243.3)	202.8 (170.3–240)	5.6	0.041	0.162	small
Carbohydrate (%TEI)	44.5 (39.4–50.4)	45.5 (43.7–47.4)	2.3	0.150	0.113	small
Sugars (g)	71.8 (49.5–97.7)	84.8 (64.3–102.4)	18.1	0.001	0.263	small
Sugars (%TEI)	16.6 (12.8–21.4)	18.6 (16.8–21)	12.2	0.002	0.249	small
Dietary Fibre (g)	17.2 (12.3–21.7)	16.7 (14.6–19.8)	-3.4	0.237	0.093	small
Total Fat (g)	61.6 (47.1–77.7)	63.1 (52.8–73.5)	2.4	0.887	0.011	small
Total Fat (%TEI)	34.7 (29.4–40.2)	33.8 (31.8–35.4)	-2.5	0.033	0.169	small
Saturated Fat (g)	22 (15.7–30)	25.1 (21.3–29.3)	14.2	0.010	0.202	small
Saturated Fat (%TEI)	12.3 (9.3–15)	13.5 (12.9–14.2)	9.8	0.005	0.220	small
Monounsaturated Fat (g)	22.1 (15.4–29.3)	22.7 (19–26.3)	2.7	0.965	0.004	small
Monounsaturated Fat (%TEI)	12.3 (10.4–14.7)	12.1 (11.2–12.8)	-1.5	0.333	0.076	small
Polyunsaturated Fat (g)	7.6 (5.7–9.7)	11.1 (9.2–12.8)	46.7	0.000	0.598	large
Polyunsaturated Fat (%TEI)	4.2 (3.4–5)	5.9 (5.5–6.2)	41.0	0.000	0.723	large
Vitamin D (µg)	2.2 (0.8–4.2)	2.3 (1.9–2.6)	7.6	0.153	0.113	small
Folate (µg)	215.1 (173.4–279.8)	198.1 (165.5–223.4)	-7.9	0.000	0.334	moderate
Vitamin C (mg)	72.3 (32.4–123.3)	67.1 (55.1–80.2)	-7.2	0.002	0.244	small

	Recall Method		% Difference	*P* value	Effect Size (*r*)	Effect Size (Magnitude)
	24HR	Generic				
Calcium (mg)	686 (509.4–884.2)	796.4 (670.3–944)	16.1	0.002	0.250	small
Iron (mg)	9.9 (7.7–12.6)	9.9 (8.1–12)	0.0	0.132	0.119	small
Potassium (mg)	2865.5 (2095.8–3462.8)	2716.2 (2200.3–3066.4)	-5.2	0.010	0.204	small
Sodium (mg)	1685.5 (1157–2142.6)	2161 (1853.1–2424.3)	28.2	0.000	0.398	moderate

P values were derived using Wilcoxon's Signed Rank Test, with $P < 0.05$ indicating statistical significance. An effect size of ≥ 0.1 and < 0.3 was considered small, ≥ 0.3 and <0.5 was considered moderate, and ≥ 0.5 was considered large [25].

Table 7.3: Bland-Altman analysis of nutrient intake estimates arising from a web-based 24-hour recall and the generic recall.

	Mean Difference	Lower LOA	Upper LOA	Percentage within LOA
Sugars (g)	-6.2	-89.6	77.1	94.4
Sugars (%TEI)	-1.3	-14.6	12.0	96.3
Dietary Fibre (g)	1.4	-16.0	18.8	98.1
Saturated Fat (g)	-1.7	-26.0	22.6	94.4
Saturated Fat (%TEI)	-0.8	-9.3	7.7	96.3
Monounsaturated fat (g)	0.8	-22.3	24.0	97.5
Monounsaturated Fat (%TEI)	0.4	-6.4	7.2	96.3
Polyunsaturated fat (g)	-3.0	-11.9	6.0	93.8
Polyunsaturated Fat (%TEI)	-1.5	-4.9	1.8	92.5
Vitamin D (µg)	0.8	-5.3	6.8	95.7
Folate (µg)	43.9	-176.6	264.4	94.4
Vitamin C (mg)	26.3	-124.3	177.0	95.7
Calcium (mg)	-54.2	-765.0	656.6	96.3
Iron (mg)	0.6	-7.9	9.2	94.4
Potassium (mg)	252.2	-1733.8	2238.2	93.2
Sodium (mg)	-347.6	-2266.3	1571.2	95.0

Differences are given as values from the 24HR minus values from the generic recall. %TEI, % Total Energy Intake.

Table 7.4: Comparison of daily nutrient intakes estimated using a web-based 24-hour recall and the novel meal-based method, based on correlation and cross-classification of quartiles.

	Correlation		Cross-Classification of Quartiles			
	Spearman Coefficient	*P*	Exact Agreement (%)	Exact + Adjacent (%)	Disagreement (%)	Extreme Disagreement (%)
Energy (kcal)	0.34	0.000	35.4	71.4	21.1	7.5
Protein (g)	0.26	0.001	34.8	69.6	25.5	5.0
Protein (%TEI)	0.43	0.000	36.6	80.1	16.1	3.7
Carbohydrate (g)	0.42	0.000	38.5	78.9	15.5	5.6
Carbohydrate (%TEI)	0.33	0.000	31.7	75.8	17.4	6.8
Sugars (g)	0.45	0.000	32.9	78.3	16.1	5.6
Sugars (%TEI)	0.40	0.000	34.2	76.4	19.9	3.7
Dietary Fibre (g)	0.43	0.000	35.4	77.6	16.1	6.2
Total Fat (g)	0.24	0.002	32.9	70.2	22.4	7.5
Total Fat (%TEI)	0.14	0.074	23.0	63.4	28.6	8.1
Saturated Fat (g)	0.20	0.011	32.9	69.6	23.6	6.8
Saturated Fat (%TEI)	0.16	0.045	25.5	68.3	23.6	8.1
Monounsaturated fat (g)	0.20	0.013	32.9	69.6	22.4	8.1
Monounsaturated Fat (%TEI)	0.02	0.753	26.1	60.2	25.5	14.3
Polyunsaturated fat (g)	0.18	0.023	32.3	67.1	25.5	7.5
Polyunsaturated Fat (%TEI)	0.13	0.112	29.2	69.6	19.9	10.6
Vitamin D (µg)	0.14	0.071	31.1	64.6	21.7	13.7

	Correlation		Cross-Classification of Quartiles			
	Spearman Coefficient	*P*	Exact Agreement (%)	Exact + Adjacent (%)	Disagreement (%)	Extreme Disagreement (%)
Folate (µg)	0.31	0.000	31.7	70.8	21.1	8.1
Vitamin C (mg)	0.34	0.000	26.1	68.9	21.7	9.3
Calcium (mg)	0.26	0.001	35.4	70.8	21.1	8.1
Iron (mg)	0.32	0.000	28.6	65.8	28.0	6.2
Potassium (mg)	0.39	0.000	26.1	65.8	28.0	6.2
Sodium (mg)	0.12	0.120	28.6	72.0	20.5	7.5

Table 7.5: Percentage of participants classified to the same category when their mean daily nutrient intakes estimated from both the 24-hour recall and the novel meal-based recall were categorised according to nutrient-based guidelines.

Nutrient	Possible categories for classification of individual nutrient intakes	% classified to the same category
Protein (g/kg BW)	Low, adequate, and high	84.5
Carbohydrate (%TEI)	Low, adequate, and high	62.7
Total Fat (%TEI)	Low, adequate, and high	52.8
Monounsaturated Fat (%TEI)	Low, adequate, and high	79.5
Polyunsaturated Fat (%TEI)	Low, adequate, and high	59.6
Saturated Fat (%TEI)	Adequate and high	70.8
Salt (g)	Adequate and high	61.5
Dietary Fibre (g)	Low and adequate	75.8
Calcium (mg)	Low, adequate, and high	57.1
Iron (mg)	Low, adequate, and high	83.9
Folate (µg)	Low, adequate, and high	82.6
Thiamin (mg)	Low and adequate	77.0
Riboflavin (mg)	Low and adequate	70.2
Vitamin C (mg)	Low, adequate, and high	65.8

7.8 Figures

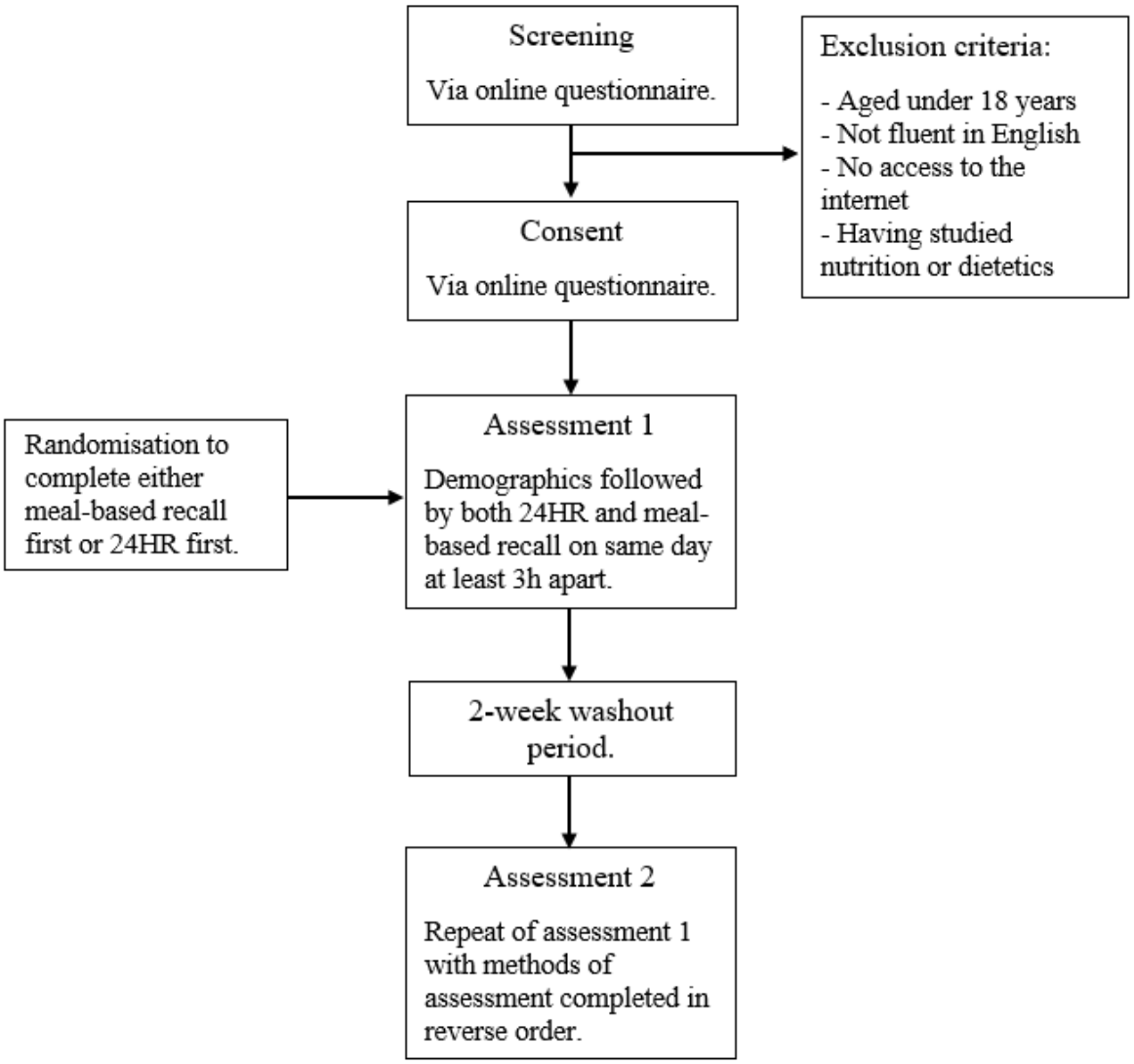


Figure 7.1: Flow diagram of participants' journey through the study.

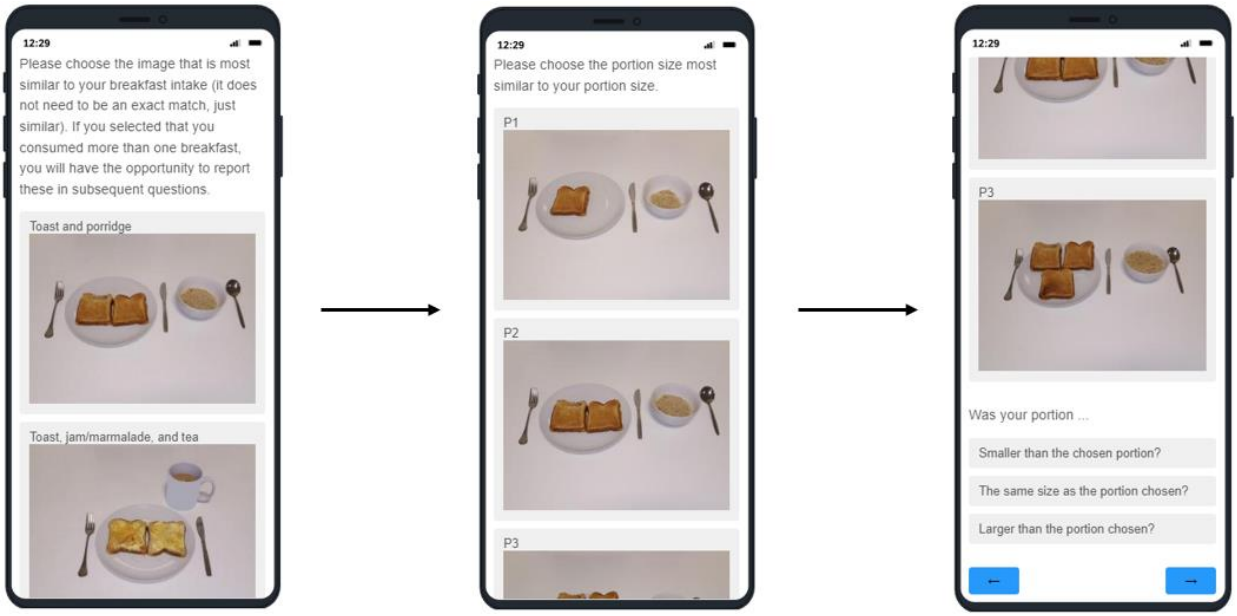


Figure 7.2: An example of the user interface of the generic meal-based recall.

Participants are asked to choose the generic breakfast image that most represents their breakfast intake on the previous day, to specify the portion size that they consumed for their chosen breakfast meal, and to answer a follow up portion size question.

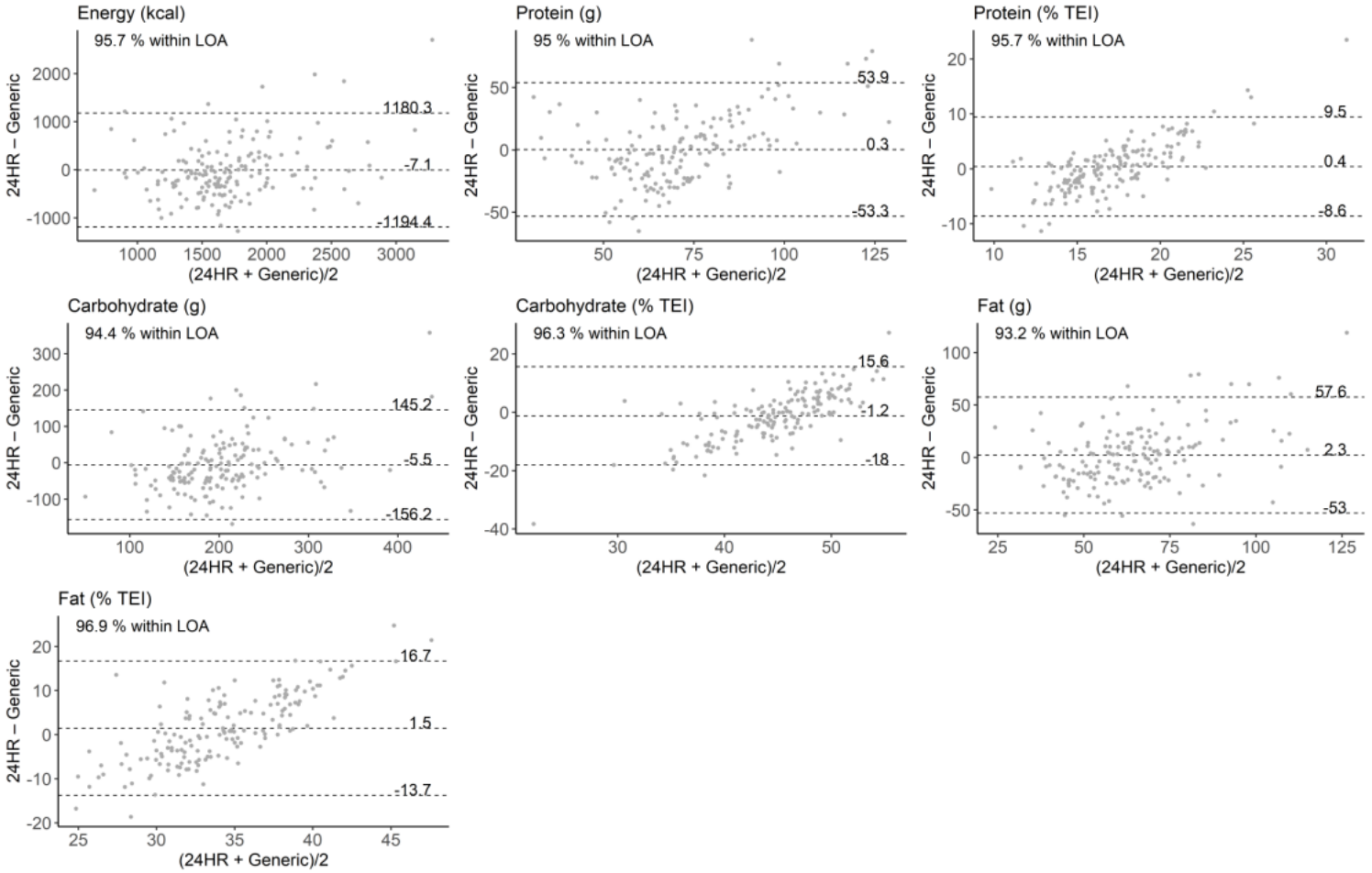


Figure 7.3: Bland-Altman plots of energy and macronutrient intake estimates.

The middle dashed line, and associated number, represents the mean difference in daily intakes between the web-based 24HR and the generic recall. The upper and lower dashed lines, and associated numbers, represent the upper and lower LOA, respectively. Each point represents an individual participant (n = 161). LOA: limits of agreement. TEI: total energy intake.

7.9 References

1. Afshin A, Sur PJ, Fay KA*, et al.* (2019) Health Effects of Dietary Risks in 195 Countries, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *The Lancet*, **393**(10184), 1958-1972.

2. Welch A (2019) Dietary Assessment. In *Manual of Dietetic Practice*, 6th ed., pp. 64-69 [J Gandy, editor]. Hoboken, NJ, USA: John Wiley & Sons, Inc.

3. Gibney MJ and Walsh MC (2013) The Future Direction of Personalised Nutrition: My Diet, My Phenotype, My Genes. *Proc Nutr Soc*, **72**(2), 219-225.

4. O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

5. Vasiloglou MF, Christodoulidis S, Reber E*, et al.* (2021) Perspectives and Preferences of Adult Smartphone Users Regarding Nutrition and Diet Apps: Web-Based Survey Study. *JMIR Mhealth Uhealth*, **9**(7), e27885.

6. 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**(1), 55.

7. Camargo AM, Botelho AM, Dean M*, et al.* (2020) Meal Planning by High and Low Health Conscious Individuals During a Simulated Shop in the Supermarket: A Mixed Methods Study. *Appetite*, **144**, 104468.

8. Bisogni CA, Falk LW, Madore E*, et al.* (2007) Dimensions of Everyday Eating and Drinking Episodes. *Appetite*, **48**(2), 218-231.

9. Englund-Ögge L, Birgisdottir BE, Sengpiel V*, et al.* (2017) Meal Frequency Patterns and Glycemic Properties of Maternal Diet in Relation to Preterm Delivery: Results from a Large Prospective Cohort Study. *PLoS One*, **12**(3), 1-19.

10. Wilson JE, Blizzard L, Gall SL*, et al.* (2019) An Eating Pattern Characterised by Skipped or Delayed Breakfast Is Associated with Mood Disorders among an Australian Adult Cohort. *Psychol Med*, **50**(16), 2711-2721.

11. Murakami K, Shinozaki N, McCaffrey TA*, et al.* (2021) Data-Driven Development of the Meal-Based Diet History Questionnaire for Japanese Adults. *Br J Nutr*, **126**(7), 1056-1064.

12. Boushey CJ, Spoden M, Zhu FM*, et al.* (2017) New Mobile Methods for Dietary Assessment: Review of Image-Assisted and Image-Based Dietary Assessment Methods. *Proc Nutr Soc*, **76**(3), 283-294.

13. Dalakleidi KV, Papadelli M, Kapolos I*, et al.* (2022) Applying Image-Based Food-Recognition Systems on Dietary Assessment: A Systematic Review. *Adv Nutr*, **13**(6), 2590-2619.

14. Timon CM, van den Barg R, Blain RJ*, et al.* (2016) A Review of the Design and Validation of Web- and Computer-Based 24-H Dietary Recall Tools. *Nutr Res Rev*, **29**(2), 268-280.

15. O'Hara C, O'Sullivan A, and Gibney ER (2022) A Clustering Approach to Meal-Based Analysis of Dietary Intakes Applied to Population and Individual Data. *J Nutr*, **152**(10), 2297-2308.

16. Irish Universities Nutrition Alliance (2011) National Adult Nutrition Survey (2008–2010). Available at: https://www.iuna.net/surveyreports (accessed 14 December 2023).

17. Fulgoni VL, Keast DR, and Drewnowski A (2009) Development and Validation of the Nutrient-Rich Foods Index: A Tool to Measure Nutritional Quality of Foods. *J Nutr*, **139**(8), 1549-1554.

18. Evans K, Hennessy A, Walton J*, et al.* (2017) Development and Evaluation of a Concise Food List for Use in a Web-Based 24-H Dietary Recall Tool. *J Nutr Sci*, **6**, e46.

19. Timon C, Evans K, Kehoe L*, et al.* (2017) Comparison of a Web-Based 24-H Dietary Recall Tool (Foodbook24) to an Interviewer-Led 24-H Dietary Recall. *Nutrients*, **9**(5), 425.

20. Timon CM, Blain RJ, McNulty B*, et al.* (2017) The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. *J Med Internet Res*, **19**(5), e158.

21. R Core Team (2022) R: A Language and Environment for Statistical Computing, 4.2.2 ed. Vienna, Austria: R Foundation for Statistical Computing.

22. RStudio Team (2022) Rstudio: Integrated Development for R., 2022.07.2+576 ed. Boston, MA, USA: RStudio PBC.

23. Henry CJ (2005) Basal Metabolic Rate Studies in Humans: Measurement and Development of New Equations. *Public Health Nutr*, **8**(7A), 1133-1152.

24. Black AE (2000) Critical Evaluation of Energy Intake Using the Goldberg Cut-Off for Energy Intake:Basal Metabolic Rate. A Practical Guide to Its Calculation, Use and Limitations. *Int J Obes*, **24**, 1119-1130.

25. Cohen J (1992) A Power Primer. *Psychol Bull*, **112**(1), 115–159.

26. Lombard MJ, Steyn NP, Charlton KE*, et al.* (2015) Application and Interpretation of Multiple Statistical Tests to Evaluate Validity of Dietary Intake Assessment Methods. *Nutr J*, **14**, 40.

27. Institute of Medicine (2005) Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids. Washington, DC, USA: The National Academies Press.

28. World Health Organization (2007) Protein and Amino Acid Requirements in Human Nutrition, *WHO Technical Report Series.* Geneva, Switzerland: WHO Press.

29. Institute of Medicine (2000) Dietary Reference Intakes: Applications in Dietary Assessment. Washington, DC, USA: The National Academies Press.

30. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

31. Uzhova I, Woolhead C, Timon CM*, et al.* (2018) Generic Meal Patterns Identified by Latent Class Analysis: Insights from NANS (National Adult Nutrition Survey). *Nutrients*, **10**(3), 310.

32. Murakami K, Livingstone MBE, Sasaki S*, et al.* (2018) Applying a Meal Coding System to 16-D Weighed Dietary Record Data in the Japanese Context: Towards the Development of Simple Meal-Based Dietary Assessment Tools. *J Nutr Sci*, **7**(e29), 1-10.

33. Cui Q, Xia Y, Wu Q*, et al.* (2021) Validity of the Food Frequency Questionnaire for Adults in Nutritional Epidemiological Studies: A Systematic Review and Meta-Analysis. *Crit Rev Food Sci Nutr*, **63**(12), 1670-1688.

34. Li K, McNulty BA, Tiernery AM*, et al.* (2016) Dietary Fat Intakes in Irish Adults in 2011: How Much Has Changed in 10 Years? *Br J Nutr*, **115**(10), 1798-1809.

35. Public Health England (2021) Mccance and Widdowson's the Composition of Foods Integrated Dataset (Cofid). Available at: https://www.gov.uk/government/publications/composition-of-foods-integrated-dataset-cofid (accessed 3rd April 2023).

36. Murakami K, Shinozaki N, Livingstone MBE*, et al.* (2023) Food Combination Questionnaire for Japanese: Relative Validity Regarding Food and Nutrient Intake and Overall Diet Quality against the 4-Day Weighed Dietary Record. *Journal of Nutritional Science*, **12**, e22.

37. Katz DL, Rhee LQ, Katz CS*, et al.* (2020) Dietary Assessment Can Be Based on Pattern Recognition Rather Than Recall. *Med Hypotheses*, **140**(February), 109644-109644.

38. Turner-McGrievy G, Hutto B, Bernhart JA*, et al.* (2022) Comparison of the Diet ID Platform to the Automated Self-Administered 24-Hour (ASA24) Dietary Assessment Tool for Assessment of Dietary Intake. *J Am Nutr Assoc*, **41**(4), 360-382.

39. Subar AF, Freedman LS, Tooze JA*, et al.* (2015) Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr*, **145**(12), 2639-2645.

40. Murakami K, Shinozaki N, Masayasu S*, et al.* (2021) Web-Based Personalized Nutrition System for Delivering Dietary Feedback Based on Behavior Change Techniques: Development and Pilot Study among Dietitians. *Nutrients*, **13**(10), 3391.

41. Diet ID (2023) The Evidence-Based Nutrition Wellbeing Toolkit. Available at: www.dietid.com/our-solutions (accessed 15th March 2023).

42. Frankenfeld CL, Poudrier JK, Waters NM*, et al.* (2012) Dietary Intake Measured from a Self-Administered, Online 24-Hour Recall System Compared with 4-Day Diet Records in an Adult US Population. *J Acad Nutr Diet*, **112**(10), 1642-1647.

43. De Keyzer W, Huybrechts I, De Vriendt V*, et al.* (2011) Repeated 24-Hour Recalls Versus Dietary Records for Estimating Nutrient Intakes in a National Food Consumption Survey. *Food Nutr Res*, **55**, 7307.

44. Araujo MC, Yokoo EM, and Pereira RA (2010) Validation and Calibration of a Semiquantitative Food Frequency Questionnaire Designed for Adolescents. *J Am Diet Assoc*, **110**(8), 1170-1177.

45. Dehghan M, del Cerro S, Zhang X*, et al.* (2012) Validation of a Semi-Quantitative Food Frequency Questionnaire for Argentinean Adults. *PLoS One*, **7**(5), e37958.

46. Celis-Morales C, Livingstone KM, Marsaux CF*, et al.* (2017) Effect of Personalized Nutrition on Health-Related Behaviour Change: Evidence from the Food4Me European Randomized Controlled Trial. *Int J Epidemiol*, **46**(2), 578-588.

47. Ye Q, Hong X, Wang Z*, et al.* (2016) Reproducibility and Validity of an FFQ Developed for Adults in Nanjing, China. *Br J Nutr*, **115**(5), 887-894.

48. Zack RM, Irema K, Kazonda P*, et al.* (2018) Validity of an FFQ to Measure Nutrient and Food Intakes in Tanzania. *Public Health Nutr*, **21**(12), 2211-2220.

49. Liu L, Wang PP, Roebothan B*, et al.* (2013) Assessing the Validity of a Self-Administered Food-Frequency Questionnaire (FFQ) in the Adult Population of Newfoundland and Labrador, Canada. *Nutr J*, **12**(49).

50. Haines PS, Hama MY, Guilkey DK*, et al.* (2003) Weekend Eating in the United States Is Linked with Greater Energy, Fat, and Alcohol Intake. *Obes Res*, **11**(8), 945-949.

51. Racette SB, Weiss EP, Schechtman KB*, et al.* (2008) Influence of Weekend Lifestyle Patterns on Body Weight. *Obesity*, **16**(8), 1826-1830.

52. An R (2016) Weekend-Weekday Differences in Diet among U.S. Adults, 2003–2012. *Ann Epidemiol*, **26**(1), 57-65.

7.10 Supplementary Material

Supplementary Table 7.1: Daily nutrient intakes estimated using a web-based 24-hour recall and the novel generic meal-based recall (n = 108), with implausible reporters of energy intake removed.

	Recall Method		% Difference	*P* value	Effect Size (*r*)	Effect Size (Magnitude)
	24HR Median (IQR)	Generic Median (IQR)				
Energy (kcal)	1807.7 (1513.1–2148.7)	1744.9 (1508.8–1963.5)	-3.5	0.080	0.168	small
Protein (g)	75.4 (60.4–93.2)	72.9 (65.2–83)	-3.3	0.027	0.213	small
Protein (%TEI)	16.3 (13.8–19.6)	16.4 (15.5–17.9)	0.3	0.707	0.036	small
Carbohydrate (g)	209.4 (181–264.2)	209.9 (180.8–244.1)	0.2	0.470	0.070	small
Carbohydrate (%TEI)	45.4 (39.4–50.4)	45.4 (43.7–47.4)	0.0	0.165	0.134	small
Sugars (g)	84.6 (62.6–110.5)	89 (70–104.5)	5.2	0.327	0.094	small
Sugars (%TEI)	17 (13.5–21.9)	19.1 (17.5–21.7)	12.7	0.015	0.235	small
Dietary Fibre (g)	18.3 (14.2–23.5)	16.8 (15–19.9)	-8.1	0.002	0.293	small
Total Fat (g)	70.8 (57.9–83.6)	63.5 (54.5–75.7)	-10.3	0.005	0.271	small
Total Fat (%TEI)	34.5 (30–39.6)	33.6 (31.4–35.4)	-2.7	0.036	0.202	small
Saturated Fat (g)	24.9 (19–33.7)	25.3 (21.9–30.5)	1.5	0.923	0.009	small
Saturated Fat (%TEI)	12.4 (9.5–15.3)	13.4 (12.9–14.1)	7.9	0.134	0.145	small
Monounsaturated Fat (g)	25.3 (20–31.1)	23.1 (19.7–27)	-8.9	0.005	0.268	small
Monounsaturated Fat (%TEI)	12.4 (10.6–14.6)	12.1 (11.2–12.8)	-2.5	0.099	0.159	small
Polyunsaturated Fat (g)	8.6 (6.5–10.2)	11.4 (9.4–12.9)	32.9	0.000	0.497	moderate
Polyunsaturated Fat (%TEI)	4.1 (3.4–4.9)	5.9 (5.3–6.2)	41.7	0.000	0.766	large
Vitamin D (µg)	2.2 (0.8–4.2)	2.4 (2–2.6)	6.1	0.180	0.129	small
Vitamin E (mg)	8.3 (6.5–11)	8.2 (6.7–9.3)	-1.6	0.036	0.201	small
Folate (µg)	234.2 (187.7–289)	200.4 (174.4–226.6)	-14.4	0.000	0.477	moderate
Vitamin C (mg)	89.1 (46.4–147.3)	67.8 (56.1–80.9)	-23.9	0.000	0.413	moderate

	Recall Method		% Difference	*P* value	Effect Size (*r*)	Effect Size (Magnitude)
	24HR Median (IQR)	Generic Median (IQR)				
Calcium (mg)	766.5 (597.7–965.3)	822.1 (714.6–948.5)	7.3	0.233	0.115	small
Magnesium (mg)	298.9 (253.4–358.1)	258.3 (218.8–285.9)	-13.6	0.000	0.571	large
Phosphorous (mg)	1272.5 (1019.3–1493.3)	1218.3 (1079.5–1365.6)	-4.3	0.037	0.201	small
Iron (mg)	10.9 (9.1–13.6)	10.2 (8.4–12.1)	-6.6	0.001	0.327	moderate
Potassium (mg)	3109.9 (2464.4–3726.7)	2763.5 (2343.9–3089.4)	-11.1	0.000	0.435	moderate
Sodium (mg)	1917.9 (1502.3–2381.7)	2162.1 (1909–2444.7)	12.7	0.029	0.210	small

P values were derived using Wilcoxon's Signed Rank Test, with $P < 0.05$ indicating statistical significance. An effect size of ≥ 0.1 and < 0.3 was considered small, ≥ 0.3 and <0.5 was considered moderate, and ≥ 0.5 was considered large. Energy misreporters were defined as those with an energy intake to basal metabolic rate ratio of less than 0.96 or greater than 2.49.

Supplementary Table 7.2: Bland-Altman analysis of nutrient intake estimates arising from a web-based 24-hour recall and the generic recall (n = 108), with implausible reporters of energy intake removed.

	Mean Difference	Lower LOA	Upper LOA	Percentage within LOA
Energy (kcal)	123.2	-892.8	1139.3	96.3
Protein (g)	6.6	-45.8	59.1	95.4
Protein (%TEI)	0.3	-7.6	8.1	95.4
Carbohydrate (g)	9.0	-120.1	138.1	93.5
Carbohydrate (%TEI)	-1.2	-15.4	13.1	95.4
Sugars (g)	-0.6	-81.8	80.5	93.5
Sugars (%TEI)	-1.2	-13.3	10.9	95.4
Dietary Fibre (g)	2.4	-11.5	16.3	95.4
Total Fat (g)	8.5	-44.5	61.5	93.5
Total Fat (%TEI)	1.7	-12.5	15.9	97.2
Saturated Fat (g)	0.8	-24.5	26.0	94.4
Saturated Fat (%TEI)	-0.6	-8.9	7.8	97.2
Monounsaturated fat (g)	3.2	-17.3	23.7	95.4
Monounsaturated Fat (%TEI)	0.6	-5.5	6.7	95.4
Polyunsaturated fat (g)	-2.3	-10.7	6.1	93.5
Polyunsaturated Fat (%TEI)	-1.5	-4.3	1.2	93.5
Vitamin D (µg)	0.8	-5.4	7.0	96.3
Vitamin E (mg)	1.0	-7.1	9.0	95.4
Folate (µg)	56.9	-159.7	273.4	93.5
Vitamin C (mg)	38.0	-120.5	196.5	95.4
Calcium (mg)	0.9	-766.4	768.2	95.4
Magnesium (mg)	58.5	-115.0	232.0	93.5
Phosphorous (mg)	107.8	-761.3	976.9	96.3
Iron (mg)	1.4	-6.0	8.8	93.5
Potassium (mg)	486.2	-1475.8	2448.2	94.4
Sodium (mg)	-124.0	-2048.0	1800.0	92.6

Differences are given as values from the 24HR minus values from the generic recall. Energy misreporters were defined as those with an energy intake to basal metabolic rate ratio of less than 0.96 or greater than 2.49. %TEI: % Total Energy Intake.

Supplementary Table 7.3: Comparison of daily nutrient intakes estimated using a web-based 24-hour recall and the novel meal-based method, based on correlation and cross-classification of quartiles (n = 108), with implausible reporters of energy intake removed.

	Correlation		Cross-Classification of Quartiles			
	Spearman Coefficient	*P* Value	Exact Agreement (%)	Exact Agreement + Adjacent (%)	Disagreement (%)	Extreme Disagreement (%)
Energy (kcal)	0.35	0.000	38.0	74.1	19.4	6.5
Protein (g)	0.21	0.027	32.4	63.9	28.7	7.4
Protein (%TEI)	0.42	0.000	30.6	77.8	19.4	2.8
Carbohydrate (g)	0.45	0.000	34.3	78.7	15.7	5.6
Carbohydrate (%TEI)	0.47	0.000	38.0	82.4	12.0	5.6
Sugars (g)	0.40	0.000	38.9	77.8	16.7	5.6
Sugars (%TEI)	0.43	0.000	34.3	77.8	19.4	2.8
Dietary Fibre (g)	0.47	0.000	35.2	79.6	13.0	7.4
Total Fat (g)	0.20	0.035	33.3	72.2	18.5	9.3
Total Fat (%TEI)	0.17	0.087	26.9	66.7	25.0	8.3
Saturated Fat (g)	0.19	0.050	31.5	70.4	20.4	9.3
Saturated Fat (%TEI)	0.14	0.139	25.0	65.7	26.9	7.4
Monounsaturated fat (g)	0.18	0.066	30.6	70.4	23.1	6.5
Monounsaturated Fat (%TEI)	0.09	0.356	27.8	64.8	24.1	11.1
Polyunsaturated fat (g)	0.13	0.175	26.9	64.8	25.0	10.2
Polyunsaturated Fat (%TEI)	0.22	0.023	28.7	73.1	17.6	9.3
Vitamin D (µg)	0.11	0.256	29.6	68.5	14.8	16.7
Vitamin E (mg)	0.27	0.006	29.6	72.2	20.4	7.4
Folate (µg)	0.30	0.002	23.1	68.5	23.1	8.3
Vitamin C (mg)	0.26	0.006	27.8	64.8	25.9	9.3

	Correlation		Cross-Classification of Quartiles			
	Spearman Coefficient	*P* Value	Exact Agreement (%)	Exact Agreement + Adjacent (%)	Disagreement (%)	Extreme Disagreement (%)
Calcium (mg)	0.19	0.044	31.5	66.7	22.2	11.1
Magnesium (mg)	0.31	0.001	34.3	72.2	19.4	8.3
Phosphorous (mg)	0.24	0.014	25.0	71.3	23.1	5.6
Iron (mg)	0.27	0.005	24.1	66.7	25.9	7.4
Potassium (mg)	0.34	0.000	25.9	67.6	22.2	10.2
Sodium (mg)	0.12	0.202	24.1	60.2	27.8	12.0

Energy misreporters were defined as those with an energy intake to basal metabolic rate ratio of less than 0.96 or greater than 2.49.

Supplementary Table 7.4: Percentage of participants classified to the same category when their mean daily nutrient intakes estimated from both the 24-hour recall and the novel meal-based recall were categorised according to nutrient-based guidelines (n = 108), with implausible reporters of energy intake removed.

Nutrient	Possible categories for classification of individual nutrient intakes	% classified to the same category
Protein (g/kg BW)	Low, adequate, and high	89.8
Carbohydrate (%TEI)	Low, adequate, and high	68.5
Total Fat (%TEI)	Low, adequate, and high	54.6
Monounsaturated Fat (%TEI)	Low, adequate, and high	81.5
Polyunsaturated Fat (%TEI)	Low, adequate, and high	66.7
Saturated Fat (%TEI)	Adequate and high	71.3
Salt (g)	Adequate and high	76.9
Dietary Fibre (g)	Low and adequate	73.1
Calcium (mg)	Low, adequate, and high	55.6
Iron (mg)	Low, adequate, and high	88.0
Folate (µg)	Low, adequate, and high	81.5
Thiamin (mg)	Low and adequate	80.6
Riboflavin (mg)	Low and adequate	78.7
Vitamin C (mg)	Low, adequate, and high	64.8

Energy misreporters were defined as those with an energy intake to basal metabolic rate ratio of less than 0.96 or greater than 2.49.

CHAPTER 8

General Discussion

8.1 Thesis Summary and Key Findings

This thesis examined a meal-based approach to analysing dietary intake data and to dietary intake assessment. The literature review described in Chapter 3 identified three main constructs in the field of meal pattern analysis: temporal patterns referring to the timing and distribution of meals, content meal patterns referring to the combinations of foods within a meal and the combinations of meals throughout a day, and context meal patterns referring to activities external to the meal such as eating alone or in company or engaging in other activities during meals. A number of gaps in the literature were identified within the review in Chapter 3 including the absence of portion size information in meal pattern research, a lack of studies comparing different methods used in meal pattern analysis, the limited research incorporating more than one meal pattern construct in a single study, and the lack of a meal-based approach to dietary intake assessment.

The work presented within this thesis primarily focused on content meal patterns while also considering some aspects of temporal meal patterns, and the approach taken to examine meal patterns can be described in three broad steps, namely: 1) food groups, 2) generic meals, and 3) meal patterns. Methods for meal pattern analysis previously used in other studies were adapted and refined to address some of the research gaps that were identified, including the lack of a data-driven approach to defining food groups, portion sizes not being considered, the absence of comparisons among the different methods used for meal pattern analysis, limited investigation of the relationship between meal patterns and health, and the absence of a meal-based method of dietary intake assessment.

In Chapter 4, a new approach was taken to identify food groups and generic meals. In many areas of nutrition research, including meal patterns and dietary patterns, the process of assigning food groups is often vaguely described referring to the grouping of foods based on some combination of nutrient composition, culinary use, and frequency of consumption [(1)]. To overcome this in the current thesis, a data-driven clustering approach was used to identify food groups in a systematic and reproducible manner. While the approach taken in this thesis was based on the existing diet quality score NRF9.3 [(2)], this approach is adaptable for different research questions because the nutrients used as input variables to group the foods can be changed as required. This was demonstrated within this thesis as the same approach was successfully applied to two different datasets in Chapters 4 and 5. The lack of use of portion sizes when identifying generic meals in previous research [(3, 4)] was also considered in this thesis. Portion size was included in the work presented here by basing generic meal portion sizes on the actual portion sizes of the meals that they represented. Having addressed these issues in the first two steps of meal pattern analysis

(food groups and generic meals), estimates of nutrient intake arising from the generic meal-based data were then compared with estimates from the original food-based data. Differences in estimates of nutrient intake between the two methods were small or negligible, demonstrating that the generic meals identified were representative of population intakes. The work presented here also demonstrated for the first time that individuals could be classified according to nutrient guidelines (low, adequate, or high intakes) using the meal-based data, suggesting that this approach may be useful in the provision of dietary advice.

The next step in the process for meal pattern analysis is determining the meal patterns themselves, i.e., combinations of generic meals that are commonly consumed. This is an area that remains a big gap in nutrition research. One of the major research issues was that, despite different methods being used in different studies to identify meal patterns, no study has directly compared these methods using a single dataset; as such, the impact of the statistical approach on the outcome of the analysis was unknown. Chapter 5 addressed this gap in the literature, by comparing clustering, principal component analysis, and latent class analysis approaches on the same dataset. The findings show that these different methods produce varying numbers of meal patterns and limited similarities in the content of the meal patterns identified by the different methods. Clustering and LCA both identified meal patterns in which meal skipping was a feature, but PCA did not. Overall, this demonstrates that the choice of statistical method influences the findings, and thus caution in the interpretation and comparison of studies is needed.

Ultimately, it is necessary to consider the link between meal patterns and health; however, only one study to date has investigated the role of content meal patterns in health, with limited findings [4]. Within this thesis Chapter 6 examined the link between meal patterns and health by describing the meal patterns identified in a cohort in the USA and exploring the relationships between those patterns and health. While the results showed clear associations between content meal patterns and dietary quality, this was not observed for health parameters. No clear trends were observed between the number of meals per day and diet quality or health in the male data, but in the female data a pattern consisting of three meals per day had the second highest diet quality (NRF9.3 score). Five of the male meal patterns and five of the female meal patterns did not include breakfast, indicating a greater proportion of food being consumed later in the day. While the nutrient quality, as assessed by NRF9.3 score, of the meal patterns without breakfast were lower than those including breakfast, differences in anthropometry and health parameters were not observed among the meal patterns. Using the novel approach applied in this analysis, the results also provide insight into the content of these patterns. For meal patterns characterised by a

skipped breakfast, those that included fruit in meals or snacks or those with main meals based on rice rather than fried foods had a higher nutrient quality.

The final chapter of the thesis describes the development of a meal-based dietary intake recall tool. It had previously been suggested that such a tool could be used in personalised nutrition to provide a method for people to record what they eat and drink at the meal level [5]; however, the work described in this thesis is the first to address and develop this concept and compare it with an established method. When compared with the 24-hour recall, estimated nutrient intakes from the generic meal-based recall were comparable for some, but not all nutrients. Comparisons between the methods were more similar at the group level compared to the individual level. Participants found the meal-based method understandable and easy to use.

8.2 Future Directions

While the work presented here addressed the gaps identified, it has also given rise to a number of research questions that require further investigation relating to meal patterns and health, meal-based dietary intake assessment, and meal-based dietary recommendations.

8.2.1 Meal Patterns and Health

While the link between meal patterns and health was examined in this thesis, overall, there is limited research conducted on the impact of content meal patterns on health. Scientific reports published in the development of recent dietary guidelines have highlighted the limited research in this area and the need for further work examining relationships between meal patterns and health to inform those guidelines [6, 7]. Because people intuitively conceptualise their dietary intakes as a collection of meals rather than a collection of foods [8], this work can also inform research on personalised nutrition recommender systems to provide recommendations for meal intakes as opposed to food intakes [9]. However, research on meal patterns and health is required to provide the evidence upon which meal-based nutrition recommendations can be made [6, 7].

While some studies have examined the relationship between health and specific meals or meal types with health [10, 11], only one study has examined the relationship between health and meal patterns (i.e., the combinations of multiple meals) [4]. Using data from Ireland, those adhering to the meal pattern “cooked breakfast, skipped light meal, and protein-carbohydrate main meal” were likely to have higher diastolic blood pressure compared with those consuming the meal pattern “cereal and/or toast for breakfast, sandwich for light meal

and protein-carbohydrate or just protein main meal" [4]. In the current thesis a novel exploratory approach was taken to examine the link between meal patterns and health, and while certain differences were identified among meal patterns in relation to nutrient quality, no differences were observed for health parameters. Currently studies investigating the relationship between content meal patterns and health are limited and this area requires further research.

The area of temporal meal patterns has shown more consistent results with greater similarities among studies and populations [12]. This is understandable given that there is less complexity in temporal meal patterns than in the content meal patterns considered in this thesis, related to the use of food groups and the identification of generic meals. Studies examining temporal patterns suggest that meal patterns characterised by three meals per day have a greater association with diet quality and health compared to those characterised by skipped meals or grazing [12-14]. These findings are also in keeping with the broader area of chrononutrition in which the timing of dietary intake relative to the timing of the circadian clock is thought to impact on metabolic processes [15]. As the area of chrononutrition grows, consideration should also be given to a variety of methodological approaches to determining meal patterns to fully explore the different aspects of meal patterns and health. Further research is required in all constructs of meal pattern research to assess the impact of the subjective decisions that are required as part of the otherwise data-driven approach to meal patterns as well as trialling other statistical approaches that have not yet been used in meal pattern research such as reduced rank regression, treelet transform analysis, Gaussian graphical models, and random forest with classification tree analysis. In particular, reduced rank regression can identify meal patterns based on the health outcomes of those consuming those patterns [1].

The inclusion of more than one construct of meal patterns in the same analysis is limited. Riou *et al.* [16], for example, examined temporal and context patterns and Uzhova *et al.* [4] also identified content meal patterns that included aspects of meal skipping as was also considered in the current thesis. Given the complexity of content meal patterns, no studies have attempted to incorporate all three constructs into the same study. Researchers have, however, begun to examine meal patterns alongside other lifestyle behaviours. For example, Lin *et al.* [17] have identified combinations of temporal meal patterns and temporal physical activity patterns. This is a continuation of the trend of nutrition research becoming more holistic and analysing a wider variety of variables within individual studies [12, 18]. The abundance of nutrition and health data that are now available as well as the emergence of advanced analytical techniques and artificial intelligence continues to drive high-dimensional research examining multiple interlinked factors. It is likely that this research will

not only incorporate patterns of dietary intake but other lifestyle factors such as exercise, sleep, stress, and others. This will allow for the identification of interactions between different aspects of health behaviours in both health and disease [19, 20].

8.2.2 Meal-Based Dietary Intake Assessment

While the current thesis demonstrated significant potential for the use of meal-based assessment in measuring dietary intake, it also identified a number of limitations which would be important to consider further. The meal-based recall in the current thesis was developed without the input of those with user experience or software development expertise. This is in contrast to web-based methods of 24-hour recall (e.g., foodbook24, myfood24, intake24) in which this expertise is incorporated into the development of such tools [21-23]. Further work is required to consider these aspects of the meal-based recall. The comparison study carried out for the current thesis provided a comparison between the meal-based method and 24-hour recall and demonstrated its potential use in dietary intake assessment studies. This should be considered, however, as a first step in the development of this novel method of dietary intake assessment. Future research could consider the comparability of the meal-based method with other methods of dietary intake assessment as well as with objective methods such as feeding studies and biomarkers of dietary intake. Future work on the user experience will likely provide scope to improve user engagement and use.

Furthermore, the meal-based method of dietary intake assessment described in this thesis could also have a role in image recognition in dietary intake assessment. As described previously, existing image recognition software can automatically identify foods and portions sizes from individuals' photographs of their dietary intake [24-27]. The complex process that facilitates this can be generally described in four broad steps. Firstly, the software segments the image to differentiate non-food from food items and to differentiate among the different food items present. Secondly, the individual food segments are then identified or classified as one of the foods from a database of known foods. Once each food is identified, the portion size of each food is then estimated by volume. Finally, the nutritional content of the meal is estimated by combining the information in the previous steps with the relevant data from a food composition database [28, 29].

However, accurately identifying foods from images using these steps is challenging because of the vast variety of foods that can be contained in any given meal. This is further complicated by the fact that some foods or ingredients may be hidden on a plate (e.g., under other foods or sauce) and that the same food can have different appearances in

different contexts (e.g., boiled and poached eggs). Notwithstanding these challenges, the accuracy of in correctly classifying food images has been reported to range from 54.7% to 93.0% [30], depending on the inputs used for training the image recognition algorithm, the settings used within the algorithm, and the number of different classes (foods) that the algorithm is expected to identify. The most accurate methods of image recognition require large training datasets, are computationally intensive, and have long running times for training [29]. Thus, simpler approaches need to be considered, and this is where meal-based approaches could apply. To date, all image-based dietary assessment methods have aimed to imitate the already existing paper-based methods by being food-based. That is, they aim to segment a meal and identify the individual food components of that meal. Consideration has not previously been given to meal-based dietary intake assessment in general or its potential role in image-based dietary intake assessment. For example, it may be possible for image recognition software to classify a meal image as a whole meal rather than a collection of individual foods. Building on the work presented here, one potential way to do this is to classify a given meal image as one of the generic meals defined in the current thesis (demonstrated to be usable in assessing intakes). In taking this approach, one is reducing the near infinite possibilities for combinations of foods that could exist in a meal, to one of the 63 generic meals identified in this thesis. Future research is required, however, to determine the feasibility of such an approach.

8.2.3 Meal-Based Dietary Recommendations

Nutrition recommender systems provide automated recommendations for foods or recipes based on information about an individual's health and dietary preferences. To date these systems have focused on taking a food-based approach [9, 31]. However, future research could consider the possibility of incorporating these systems in a meal-based framework whereby the recommendations could be informed by meal-based dietary intake assessment. This could take the form of people completing the meal-based dietary intake assessment and being categorised to one of the meal patterns identified in the current thesis. Based on the nutrient quality of the baseline meal pattern, recommendations may be made to move the individual from a lower nutrient quality meal pattern to a higher nutrient quality meal pattern. These recommendations could be facilitated by a recipe database that is based on the generic meals from the meal patterns with higher nutrient quality. For example, it could be recommended to change from one breakfast or lunch to another while also providing information on an appropriate recipe, retail store, or food outlet to support that change. However, further work is needed both on the feasibility of such an approach and its impact on behaviour change compared to existing approaches.

8.3 Impact

The meal-based approach described in the current thesis can influence how we interpret, collect, and use dietary intake data in a way that could have wide reaching impacts in the field of human nutrition and health. The chapters in this thesis relating to the identification of the generic meals and the meal-based dietary assessment tool have been registered with UCD's technology transfer office, NovaUCD, in the form of invention disclosures. Other impacts are discussed in the following subsections in relation to three key areas: nutrition research methods, public health, and personalised nutrition. In the area of nutrition research methods, an impact has already been seen with other research groups incorporating the meal-based methods described in the current thesis into their research, as discussed below. For the other two key areas, public health and personalised nutrition, the potential impact (as opposed to actual impact) is discussed below.

8.3.1 Nutrition Research Methods

One of the benefits of the generic meals framework is its flexibility to be adapted for use in a variety of datasets and to answer a variety of different research questions in a range of different areas of human nutrition research. This has given rise to a number of collaborations with other research groups.

One area of impact for the generic meals framework is the identification of characteristic meals in any given dataset. The World Health Organization (WHO), for example, have raised concerns regarding the increasing use of meal delivery applications despite the limited information available regarding the nutritional quality of the meals available on these applications [32]. The limited research that has been conducted in this area indicates that these applications disproportionately increase access to meals of lower nutrient quality compared to those with higher nutrient quality [33-35]. To address this issue and to more broadly understand the nutrient quality of meals available from takeaway restaurants, the WHO is conducting research to characterise the nutritional profile of takeaway meals available for sale across the European region (Holly Rippin, personal communication, 30th January 2023). While this information is available from the large multi-national fast-food chains, there is no information available from the small independent operators. One solution to this problem that has been discussed with the WHO research team is to use the generic meals process to specifically examine takeaway meals in national dietary survey data to provide generic nutritional information for generic takeaway meals. This method could provide a useful characterisation of the nutrient quality of those meals in the absence of specific information from the individual food business operators.

Building on from this, determining average nutritional content of common meals, could benefit analysis of national dietary surveys. This approach has been discussed with the MRC epidemiology unit in Cambridge, who run the National Diet and Nutrition Survey (NDNS) in the UK. Dietary data within NDNS are now collected via an automated web-based 24-hour recall called Intake24 [36]. Intake24 contains a comprehensive food list for people to report the foods they eat. However, one area that is currently limited is the ability for individuals to report whole meals; people will sometimes search for whole meals in the search bar rather than individual foods (Birdem Amoutzopoulos, personal communication, 6th November 2023). The generic meals framework described within thesis has been discussed with the research team for the NDNS as a potential option to provide a meal list to complement the existing food list. This could also work for other similar technologies and web-based intake assessment tools.

Finally, the use of the generic meals framework is also being considered as a method within nutrition and sustainability research. Human health and planetary health are intrinsically linked, and environmental sustainability is increasingly being considered and included in national dietary guidelines [37, 38]. It stands to reason, therefore, that in addition to the nutritional and health aspects of meal patterns, the sustainability aspects of meal patterns should also be investigated. Preliminary work has begun on this topic in collaboration with researchers at Teagasc (the Irish state agency for food and agriculture development) to investigate the CO_2 equivalents associated with different generic meals and meal patterns. In the current thesis, generic meals and meal patterns were identified using nutrients of public health importance as input variables; future work could also consider the use of sustainability metrics such as greenhouse gas emissions, land use, or water use etc. as input variables to the meal pattern process.

8.3.2 Public Health

A food-based approach to public health nutrition guidance is the primary approach to provision of dietary guidelines in most countries [39]; however, some guidance in the USA and Brazil now acknowledge the potential for meal-based information to complement what is already known about nutrients, foods, and dietary patterns by referring to the context in which meals are consumed, the frequency or number of meals consumed, and the various combinations of foods that can be used to make meals [6, 7]. However, the need for further research is highlighted to provide a stronger evidence base to support such guidelines [6]. The current thesis adds to this evidence, but also raises important questions that need to be addressed prior to their broad inclusion in dietary guidelines.

More recently, public health bodies have begun to provide access to mobile-based applications to assist with their public health nutrition remit. For example, in the UK, the NHS provide access to a food scanner app [40] which allows consumers to scan barcodes of foods during shopping or purchase to get suggestions for swaps to healthier foods. In the USA, the US Department of Agriculture provide the MyPlate app with which users can complete a food frequency questionnaire and receive personalised advice based on their responses relative to the food-based dietary guidelines [41]. Self-monitoring of behaviour is a well-established behaviour change technique in the fields of psychology [42] and nutrition [43]. This monitoring can be achieved via the various methods available for dietary intake assessment and is a feature of the two applications described above. While the more burdensome methods of dietary intake assessment may have a role to play in habit-forming during initial engagement with nutrition technologies, allowing users to change from one dietary assessment method to another may aid in maintaining such engagement over a prolonged period [44]. In this context, the availability of multiple forms of dietary intake assessment could form part of a suite of options available to a user depending on individual preference. It is plausible that the further development of the meal-based recall, as described above, could make it a candidate as a method for dietary intake monitoring available to influence behaviour change via public health nutrition applications.

8.3.3 Personalised Nutrition

Personalised nutrition is an approach to nutrition research and practice that aims to tailor dietary advice according to an individual's existing diet and lifestyle, their phenotypic profile, their genetic profile, or a combination of those factors [5, 45]. This is in contrast to, for example, government dietary guidelines which are typically targeted more broadly at national populations or certain population subgroups based on characteristics such as sex or age [39].

Personalised nutrition advice at the diet level is provided to individuals following dietary intake assessment, usually using 24-hour recalls, food frequency questionnaires, or diet records. Typically, people are categorised as low, adequate, or high consumers of various nutrients or food groups based on their dietary intake assessment, and dietary advice is provided accordingly [46-51]. For personalised nutrition based on phenotype, a broad range of measures are used to tailor advice from physical measures such as waist circumference and weight to biochemical measurements such as blood lipids [5, 52]. Personalised nutrition based on genetics considers specific genetic variations that give rise to varying phenotypes that can be influenced by diet [53]. For example, people with two copies of a variant (rs9939609) of the fat mass and obesity-associated (FTO) gene have an increased risk of

obesity [54], while individuals who have both a low-activity variant of the 5,10-methylene tetrahydrofolate reductase (MTHFR) gene and a low intake of folate have a higher risk of cardiovascular disease [55].

Despite the well-established dietary, phenotypic, and genetic influences on health, the efficacy of personalised interventions based on these variables varies. A recent systematic review of 11 randomised controlled trials summarised the effectiveness of personalised nutrition interventions in individuals without chronic disease. The outcome measures varied among studies and included measures such as anthropometry, dietary intake, physical activity, biochemistry, and psychological measures. Eight of the 11 studies noted improvements in the outcome measures from personalised nutrition. However, only one study identified an improvement based on genetic information over and above the improvements from dietary or phenotype information [45]. One aspect that is highlighted in this and other personalised nutrition research is that personalised nutrition interventions that incorporate behaviour change theory and techniques are more effective than those that do not [45, 56-58].

Dietary intake assessment is an essential element of providing dietary advice that is personalised to individuals' existing dietary intakes [59]. As discussed above, it is also a method of self-monitoring of behaviour which is a well-established behaviour change technique [42, 43]. While more research and development is required to further develop the meal-based dietary intake assessment and a meal-based recommender system, it is conceivable that they could be applied to existing personalised nutrition frameworks to influence behaviour change. There are numerous companies that provide digital methods of dietary intake assessment and dietary advice. The large datasets of meal images, health information, and dietary intakes available to these companies would lend themselves well to the development of meal-based image recognition, the investigation of the relationship between meal patterns and health, and the development of meal-based recommender systems. It is also conceivable that a meal-based framework for personalised nutrition could be integrated into the wider digital ecosystem, providing meal-based recommendations that are linked to online platforms for food shopping and delivery from supermarkets or takeaway restaurants.

8.4 Strengths and Limitations

One of the major strengths of this thesis is that the data-driven approach that was developed provides a clear framework that can be applied to or adapted for other datasets. Our approach was successfully applied across two large dietary intake datasets from

national nutrition surveys (NANS and NHANES). However, it is important to note that whilst using a statistically driven approach is beneficial, it also means that, because patterns derived using data-driven methods are based only on the underlying data, they may not be reproducible within populations over time or between populations. This would need to be considered when comparing the findings within this thesis to future work using this approach [1, 60]. Furthermore, even when a data-driven approach is used, there is a small requirement for some subjective decisions on the part of the researcher; for example, in relation to the approach chosen for food groups, the way in which the number of meal patterns are chosen, or the specific variables that are chosen as input variables for the various statistical techniques. However, the rationale for these decisions have been clearly provided in the methods chapter and in the methods sections of each original research chapter. Furthermore, this level of subjective decision making is significantly reduced compared to traditional approaches, where individual decisions are often made at each step of processing of food intake data [60, 61].

In the current thesis, a nationally representative sample was to identify generic meals that are representative of the meals consumed in Ireland. However, at the time of analysis, the most recently available survey that was available was from 2008 to 2010 [62] which may not still be reflective of current day intakes. A more recent survey (National Adult Nutrition Survey 2 (NANS2)), is ongoing, and the work described in the current thesis could repeated on those data for comparison or to examine changes in eating habits across the two datasets. Similarly, data from the USA were used in the current thesis, specifically the NHANES cohort. These data provided a large sample size on which to identify a range of meal patterns, as per NANS, but also allowed exploration of relationships with health outcomes. However, given the complex sampling process used in the NHANES [63] and the complex multi-step process used in the current thesis to identify meal patterns, it was not feasible to adjust the NHANES data using weightings to provide a nationally representative sample. Instead, that work was carried out on the sample in which certain demographics were deliberately oversampled [63]. So, while it is a limitation of the thesis that those meal patterns do not necessarily reflect the meal patterns that exist at the population level in the USA, that work does provide a true reflection of meal patterns in that sample and a true comparison of different methods for meal pattern analysis.

The dietary intake data used in this thesis are derived from subjective methods of dietary intake assessment (food records and 24-hour recalls). As discussed in detail in the introduction to the thesis, it is well established that these data are subject to a certain degree of error and bias [64]. This was mitigated against in the current thesis by using data collected using structured and validated processes, i.e., the 4-day weighed food records from the

NANS in Ireland and the two non-consecutive 24-hour recalls in NHANES. Two non-consecutive 24-hour recalls were also used in Chapter 7 as a comparison method for the newly developed generic recall. Other strengths of that comparison study included a large sample size of 161 at the higher end of sample sizes in similar comparison studies [65, 66], randomisation as to whether the generic recall or food-based recall would be completed first, a two-week washout period, and the structured inclusion of both weekend days and weekdays. However, within the resources available for the current thesis it was only possible to use a convenience sample rather than one that is representative of the population as a whole.

8.5 Conclusion

This thesis developed and demonstrated a reproducible approach to determining meal patterns in dietary intake data. It has shown that a clustering approach can be used to identify characteristic, or generic meals, in a dietary intake dataset. The process for identifying generic meals described in this thesis provides similar estimates of nutritional intake to standard food-based approaches. Meal patterns, i.e., the different combinations of generic meals consumed over the course of a day, can be identified using different statistical methods. However, differences arise in the meal patterns that are identified using different methods and their relative strengths and weaknesses should be considered when choosing a method. Some associations between meal patterns and diet quality were identified. These were limited but warrant further investigation. A meal-based method of dietary intake assessment is a feasible method of collecting dietary intake data, showing moderate agreement with the 24-hour recall. Further research is required to further develop this tool and improve its accuracy across a range of nutrients.

8.6 References

1. Schulz CA, Oluwagbemigun K, and Nothlings U (2021) Advances in Dietary Pattern Analysis in Nutritional Epidemiology. *Eur J Nutr*, **60**(8), 4115-4130.

2. Fulgoni VL, Keast DR, and Drewnowski A (2009) Development and Validation of the Nutrient-Rich Foods Index: A Tool to Measure Nutritional Quality of Foods. *J Nutr*, **139**(8), 1549-1554.

3. Woolhead C, Gibney MJ, Walsh MC*, et al.* (2015) A Generic Coding Approach for the Examination of Meal Patterns. *Am J Clin Nutr*, **102**(2), 316-323.

4. Uzhova I, Woolhead C, Timon CM*, et al.* (2018) Generic Meal Patterns Identified by Latent Class Analysis: Insights from NANS (National Adult Nutrition Survey). *Nutrients*, **10**(3), 310.

5. Gibney MJ and Walsh MC (2013) The Future Direction of Personalised Nutrition: My Diet, My Phenotype, My Genes. *Proc Nutr Soc*, **72**(2), 219-225.

6. Dietary Guidelines Advisory Committee (2020) Scientific Report of the 2020 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Agriculture and the Secretary of Health and Human Services. Washington, DC: *U.S. Department of Agriculture, Agricultural Research Service*.

7. Ministry of Health of Brazil (2015) Dietary Guidelines for the Brazilian Population. Brasília/DF: *Secretariat of Health Care, Primary Health Care Department*.

8. Bisogni CA, Falk LW, Madore E*, et al.* (2007) Dimensions of Everyday Eating and Drinking Episodes. *Appetite*, **48**(2), 218-231.

9. Trang Tran TN, Atas M, Felfernig A*, et al.* (2017) An Overview of Recommender Systems in the Healthy Food Domain. *Journal of Intelligent Information Systems*, **50**(3), 501-526.

10. Shi Z, Riley M, Taylor A*, et al.* (2017) Meal-Specific Food Patterns and the Incidence of Hyperglycemia in a Chinese Adult Population. *Br J Nutr*, **118**(1), 53-59.

11. Shinozaki N, Murakami K, Asakura K*, et al.* (2021) Identification of Dish-Based Dietary Patterns for Breakfast, Lunch, and Dinner and Their Diet Quality in Japanese Adults. *Nutrients*, **13**(1), 67.

12. O'Hara C and Gibney ER (2021) Meal Pattern Analysis in Nutritional Science: Recent Methods and Findings. *Adv Nutr*, **12**(4), 1365-1378.

13. Aqeel MM, Guo J, Lin L*, et al.* (2020) Temporal Dietary Patterns Are Associated with Obesity in US Adults. *J Nutr*, **150**(12), 3259-3268.

14. Lin L, Guo J, Li Y*, et al.* (2022) The Discovery of Data-Driven Temporal Dietary Patterns and a Validation of Their Description Using Energy and Time Cut-Offs. *Nutrients*, **14**(17), 3483.

15. Katsi V, Papakonstantinou IP, Soulaidopoulos S*, et al.* (2022) Chrononutrition in Cardiometabolic Health. *J Clin Med*, **11**(2), 296.

16. Riou J, Lefèvre T, Parizot I*, et al.* (2015) Is There Still a French Eating Model? A Taxonomy of Eating Behaviors in Adults Living in the Paris Metropolitan Area in 2010. *PLoS One*, **10**(3), 1-18.

17. Lin L, Guo J, Aqeel MM*, et al.* (2022) Joint Temporal Dietary and Physical Activity Patterns: Associations with Health Status Indicators and Chronic Diseases. *Am J Clin Nutr*, **115**(2), 456-470.

18. Mozaffarian D, Rosenberg I, and Uauy R (2018) History of Modern Nutrition Science-Implications for Current Research, Dietary Guidelines, and Food Policy. *BMJ*, **361**, k2392.

19. Bond A, McCay K, and Lal S (2023) Artificial Intelligence & Clinical Nutrition: What the Future Might Have in Store. *Clin Nutr ESPEN*, **57**, 542-549.

20. Miyazawa T, Hiratsuka Y, Toda M*, et al.* (2022) Artificial Intelligence in Food Science and Nutrition: A Narrative Review. *Nutr Rev*, **80**(12), 2288-2300.

21. Timon CM, Blain RJ, McNulty B*, et al.* (2017) The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. *J Med Internet Res*, **19**(5), e158.

22. 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**(1), 233-240.

23. Simpson E, Bradley J, Poliakov I*, et al.* (2017) Iterative Development of an Online Dietary Recall Tool: Intake24. *Nutrients*, **9**(2), 118.

24. Boushey CJ, Spoden M, Zhu FM*, et al.* (2017) New Mobile Methods for Dietary Assessment: Review of Image-Assisted and Image-Based Dietary Assessment Methods. *Proc Nutr Soc*, **76**(3), 283-294.

25. Zhu F, Bosch M, Khanna N*, et al.* (2015) Multiple Hypotheses Image Segmentation and Classification with Application to Dietary Assessment. *IEEE J Biomed Health Inform*, **19**(1), 377-388.

26. Kong F, He H, Raynor HA*, et al.* (2015) Dietcam: Multi-View Regular Shape Food Recognition with a Camera Phone. *Pervasive and Mobile Computing*, **19**(May), 108-121.

27. Wang Y, He Y, Boushey CJ*, et al.* (2018) Context Based Image Analysis with Application in Dietary Assessment and Evaluation. *Multimed Tools Appl*, **77**(15), 19769-19794.

28. Konstantakopoulos FS, Georga EI, and Fotiadis DI (2023) A Review of Image-Based Food Recognition and Volume Estimation Artificial Intelligence Systems. *IEEE Rev Biomed Eng*(Online ahead of print).

29. Dalakleidi KV, Papadelli M, Kapolos I*, et al.* (2022) Applying Image-Based Food-Recognition Systems on Dietary Assessment: A Systematic Review. *Adv Nutr*, **13**(6), 2590-2619.

30. Lo FPW, Sun Y, Qiu J*, et al.* (2020) Image-Based Food Classification and Volume Estimation for Dietary Assessment: A Review. *IEEE J Biomed Health Inform*, **24**(7), 1926-1939.

31. Kirk D, Kok E, Tufano M*, et al.* (2022) Machine Learning in Nutrition Research. *Adv Nutr*, **13**(6), 2573–2589.

32. World Health Organization (2021) Slide to Order: A Food Systems Approach to Meal Delivery Apps. Copenhagen, Denmark: *WHO European Office for the Prevention and Control of Noncommunicable Diseases*.

33. Horta PM, Souza JPM, Rocha LL*, et al.* (2021) Digital Food Environment of a Brazilian Metropolis: Food Availability and Marketing Strategies Used by Delivery Apps. *Public Health Nutr*, **24**(3), 544-548.

34. Norriss D, Crossin R, Curl A*, et al.* (2022) Food Outlet Access and the Healthiness of Food Available 'On-Demand' Via Meal Delivery Apps in New Zealand. *Nutrients*, **14**(20), 4228.

35. Poelman MP, Thornton L, and Zenk SN (2020) A Cross-Sectional Comparison of Meal Delivery Options in Three International Cities. *Eur J Clin Nutr*, **74**(10), 1465-1473.

36. Public Health England (2021) Evaluation of Changes in the Dietary Methodology in the National Diet and Nutrition Survey Rolling Programme from Year 12 (2019 to 2020). London, England: *Public Health England*.

37. James-Martin G, Baird DL, Hendrie GA*, et al.* (2022) Environmental Sustainability in National Food-Based Dietary Guidelines: A Global Review. *Lancet Planet Health*, **6**(12), e977-e986.

38. Food and Agriculture Organization of the United Nations and Food Climate Research Network (2016) Plates, Pyramids, Planet. Developments in National Healthy and Sustainable Dietary Guidelines: A State of Play Assessment. Rome, Italy: *Nutrition and Food Systems Division, Food and Agriculture Organization of the United Nations*.

39. Herforth A, Arimond M, Álvarez-Sánchez C*, et al.* (2019) A Global Review of Food-Based Dietary Guidelines. *Adv Nutr*, **10**(4), 590-605.

40. National Health Service (2023) Download the Free Nhs Food Scanner App. Available at: https://www.nhs.uk/healthier-families/food-facts/nhs-food-scanner-app/ (accessed 29th November 2023).

41. U.S. Department of Agriculture (2023) Start Simple with Myplate App. Available at: https://www.myplate.gov/resources/tools/startsimple-myplate-app (accessed 29th November 2023).

42. Michie S, Richardson M, Johnston M*, et al.* (2013) The Behavior Change Technique Taxonomy (V1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. *Ann Behav Med*, **46**(1), 81-95.

43. Berry R, Kassavou A, and Sutton S (2021) Does Self-Monitoring Diet and Physical Activity Behaviors Using Digital Technology Support Adults with Obesity or Overweight to Lose Weight? A Systematic Literature Review with Meta-Analysis. *Obes Rev*, e13306.

44. Turner-McGrievy GM, Yang C-H, Monroe C*, et al.* (2021) Is Burden Always Bad? Emerging Low-Burden Approaches to Mobile Dietary Self-Monitoring and the Role Burden Plays with Engagement. *Journal of Technology in Behavioral Science*, **6**(3), 447-455.

45. Jinnette R, Narita A, Manning B*, et al.* (2020) Does Personalized Nutrition Advice Improve Dietary Intake in Healthy Adults? A Systematic Review of Randomized Controlled Trials. *Adv Nutr*, **12**((3)), 657-669.

46. Wright JL, Sherriff JL, Dhaliwal SS*, et al.* (2011) Tailored, Iterative, Printed Dietary Feedback Is as Effective as Group Education in Improving Dietary Behaviours: Results from a Randomised Control Trial in Middle-Aged Adults with Cardiovascular Risk Factors. *Int J Behav Nutr Phys Act*, **8**, 43.

47. Forster H, Walsh MC, O'Donovan CB*, et al.* (2016) A Dietary Feedback System for the Delivery of Consistent Personalized Dietary Advice in the Web-Based Multicenter Food4Me Study. *J Med Internet Res*, **18**(6), e150.

48. Kroeze W, Oenema A, Campbell M*, et al.* (2008) The Efficacy of Web-Based and Print-Delivered Computer-Tailored Interventions to Reduce Fat Intake: Results of a Randomized, Controlled Trial. *J Nutr Educ Behav*, **40**(4), 226-236.

49. Robroek SJ, Bredt FJ, and Burdorf A (2007) The (Cost-)Effectiveness of an Individually Tailored Long-Term Worksite Health Promotion Programme on Physical Activity and Nutrition: Design of a Pragmatic Cluster Randomised Controlled Trial. *BMC Public Health*, **7**, 259.

50. Xu Z, Geng J, Zhang S*, et al.* (2020) A Mobile-Based Intervention for Dietary Behavior and Physical Activity Change in Individuals at High Risk for Type 2 Diabetes Mellitus: Randomized Controlled Trial. *JMIR Mhealth Uhealth*, **8**(11), e19869.

51. Rollo ME, Haslam RL, and Collins CE (2020) Impact on Dietary Intake of Two Levels of Technology-Assisted Personalized Nutrition: A Randomized Trial. *Nutrients*, **12**(11), 3334.

52. Schoeneck M and Iggman D (2021) The Effects of Foods on Ldl Cholesterol Levels: A Systematic Review of the Accumulated Evidence from Systematic Reviews and Meta-Analyses of Randomized Controlled Trials. *Nutr Metab Cardiovasc Dis*, **31**(5), 1325-1338.

53. Ferguson LR, De Caterina R, Gorman U*, et al.* (2016) Guide and Position of the International Society of Nutrigenetics/Nutrigenomics on Personalised Nutrition: Part 1 - Fields of Precision Nutrition. *J Nutrigenet Nutrigenomics*, **9**(1), 12-27.

54. Frayling TM, Timpson NJ, Weedon MN*, et al.* (2007) A Common Variant in the Fto Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. *Science*, **316**(5826), 889-894.

55. Liew SC and Gupta ED (2015) Methylenetetrahydrofolate Reductase (Mthfr) C677t Polymorphism: Epidemiology, Metabolism and the Associated Diseases. *Eur J Med Genet*, **58**(1), 1-10.

56. Lara J, Evans EH, O'Brien N*, et al.* (2014) Association of Behaviour Change Techniques with Effectiveness of Dietary Interventions among Adults of Retirement Age: A Systematic Review and Meta-Analysis of Randomised Controlled Trials. *BMC Med*, **12**(177), 1741-7015.

57. Celis-Morales C, Lara J, and Mathers JC (2015) Personalising Nutritional Guidance for More Effective Behaviour Change. *Proc Nutr Soc*, **74**(2), 130-138.

58. Dijksterhuis GB, Bouwman EP, and Taufik D (2021) Personalized Nutrition Advice: Preferred Ways of Receiving Information Related to Psychological Characteristics. *Front Psychol*, **12**, 575465.

59. Welch A (2019) Dietary Assessment. In *Manual of Dietetic Practice*, 6th ed., pp. 64-69 [J Gandy, editor]. Hoboken, NJ, USA: John Wiley & Sons, Inc.

60. Moeller SM, Reedy J, Millen AE*, et al.* (2007) Dietary Patterns: Challenges and Opportunities in Dietary Patterns Research. *J Am Diet Assoc*, **107**(7), 1233-1239.

61. Hu F (2002) Dietary Pattern Analysis: A New Direction in Nutritional Epidemiology. *Curr Opin Lipidol*, **13**(1), 3-9.

62. Irish Universities Nutrition Alliance (2011) National Adult Nutrition Survey (2008–2010). Available at: https://www.iuna.net/surveyreports (accessed 14 December 2023).

63. Centers for Disease Control and Prevention (CDC) (2018) National Health and Nutrition Examination Survey (NHANES) Data 2017-2018. Available at: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 (accessed 14 December 2023).

64. Subar AF, Freedman LS, Tooze JA*, et al.* (2015) Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr*, **145**(12), 2639-2645.

65. Timon CM, van den Barg R, Blain RJ*, et al.* (2016) A Review of the Design and Validation of Web- and Computer-Based 24-H Dietary Recall Tools. *Nutr Res Rev*, **29**(2), 268-280.

66. Murai U, Tajima R, Matsumoto M*, et al.* (2023) Validation of Dietary Intake Estimated by Web-Based Dietary Assessment Methods and Usability Using Dietary Records or 24-H Dietary Recalls: A Scoping Review. *Nutrients*, **15**(8), 1816.

APPENDICES

Appendix 1: The recruitment poster used for Chapter 7.

Insight
SFI RESEARCH CENTRE FOR DATA ANALYTICS

UCD DUBLIN
UCD Institute of Food and Health

WOULD YOU LIKE FEEDBACK ON YOUR DIETARY INTAKES?

Researchers at the Insight Centre for Data Analytics and the UCD Institute of Food and Health are looking for volunteers to compare two different online tools to track their diet.

Feedback on dietary intakes is provided to participants

Find out more or sign up by scanning the QR code:

or visit

mealstudy.ie

Appendix 2: The screening questionnaire used for Chapter 7.

Please provide the following details

First Name

Surname

Email Address

Do you live in Ireland?

- ○ Yes
- ○ No

What is your age (you must be aged 18 or older to take part in this study)?

Have you previously or are you currently studying nutrition or dietetics at degree level?

- ○ Yes
- ○ No

Appendix 3. The evaluation questionnaire used in Chapter 7.

Please answer the following questions in relation to the meal-based tool (i.e. the tool that presented you with a number of photos of meals and you were asked to choose the one that was most similar to your intake).

	Agree	Somewhat Agree	Somewhat Disagree	Disagree
The meals that appeared in the tool were largely representative of the meals that I consume.	○	○	○	○
The portion sizes in the tool were largely representative of the portion sizes that I consume.	○	○	○	○
The instructions provided were clear and easy to understand.	○	○	○	○
Overall, the tool was easy to use.	○	○	○	○

Please add any additional comments regarding the above questions here

How would you describe the ease of use of the meal-based tool compared with the alternative food-based tool (Foodbook24)?

- ○ Better
- ○ Somewhat better
- ○ Somewhat worse
- ○ Worse

Please add any additional comments regarding the above questions here

Do you think you would consider using a tool similar to the meal-based tool in the future to record your dietary intakes?

- ○ Yes
- ○ No

Please describe why.

Are there any features that you particularly like about the meal-based tool?

Are there any features that you particularly dislike about the meal-based tool?

Please add any other final comments that you have here.