



## Identifying usual food choices at meals in overweight and obese study volunteers: implications for dietary advice

Vivienne X. Guan<sup>1,2\*</sup>, Yasmine C. Probst<sup>1,2</sup>, Elizabeth P. Neale<sup>1,2</sup>, Marijka J. Batterham<sup>2,3</sup> and Linda C. Tapsell<sup>1,2</sup>

<sup>1</sup>School of Medicine, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong, NSW 2522, Australia

<sup>2</sup>Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong, NSW 2522, Australia

<sup>3</sup>Statistical Consulting Centre, National Institute for Applied Statistical Research Australia, University of Wollongong, Wollongong, NSW 2522, Australia

(Submitted 8 January 2018 – Final revision received 2 May 2018 – Accepted 22 May 2018 – First published online 17 July 2018)

### Abstract

Understanding food choices made for meals in overweight and obese individuals may aid strategies for weight loss tailored to their eating habits. However, limited studies have explored food choices at meal occasions. The aim of this study was to identify the usual food choices for meals of overweight and obese volunteers for a weight-loss trial. A cross-sectional analysis was performed using screening diet history data from a 12-month weight-loss trial (the HealthTrack study). A descriptive data mining tool, the Apriori algorithm of association rules, was applied to identify food choices at meal occasions using a nested hierarchical food group classification system. Overall, 432 breakfasts, 428 lunches, 432 dinners and 433 others (meals) were identified from the intake data ( $n$  433 participants). A total of 142 items of closely related food clusters were identified at three food group levels. At the first sub-food group level, bread emerged as central to food combinations at lunch, but unprocessed meat appeared for this at dinner. The dinner meal was characterised by more varieties of vegetables and of foods in general. The definitions of food groups played a pivotal role in identifying food choice patterns at main meals. Given the large number of foods available, having an understanding of eating patterns in which key foods drive overall meal content can help translate and develop novel dietary strategies for weight loss at the individual level.

**Key words:** Food choices: Meals: Overweight: Data mining

In recent years, overweight and obesity has continued to grow as a global epidemic<sup>(1,2)</sup>. The prevalence of obesity has more than doubled in adults over the past 30 years<sup>(3)</sup>. In Australia, the percentage of overweight and obese adults increased from 56.3 to 62.8% between 1995 and 2011–2012<sup>(4)</sup>. Overweight and obesity is closely linked to death and chronic diseases (e.g. CVD)<sup>(5)</sup>.

The first-line treatment and management of overweight and obesity are lifestyle interventions including dietary interventions, increased physical activity and support for behavioural change<sup>(6,7)</sup>. Consuming an energy-restricted diet regardless of macronutrient profile is an effective strategy to lose weight<sup>(8)</sup>. However, research has demonstrated that overweight and obese individuals who choose foods that are aligned with a healthy dietary pattern before dietary intervention may experience difficulties complying with energy restrictions compared with those who do not<sup>(9)</sup>. Moreover, in the long term, the adherence to a dietary prescription tends to be poor in the overweight and obese population<sup>(10)</sup>. Approximately half of the participants who were involved in weight-loss trials regained their weight after the intervention<sup>(11)</sup>. The literature suggests that the method of achieving dietary goal for weight loss

may play an important role in adhering to dietary prescription for weight loss<sup>(12)</sup>. Thus, exploring foods usually consumed by overweight and obese individuals may help in developing sustainable strategies for better food choices to weight loss.

Research suggests that there may be value in considering the meals in which foods are consumed. Food choice events are colloquially labelled as meal occasions (e.g. breakfast, lunch, dinner or snacks)<sup>(13)</sup>. During meal occasions, individual foods and/or mixed dishes, which are prepared and/or cooked from individual foods known as ingredients, are eaten<sup>(14,15)</sup>. Thus, meal patterns appear to be closely linked to eating habits<sup>(16)</sup>. Furthermore, the American Heart Association has suggested that translating energy intakes at each meal occasion into food choices is required, when considering meal patterns<sup>(17)</sup>. Therefore, it may indicate that the ultimate translation of dietary recommendations to practice requires identifying the actual foods consumed that make up the meals. Thus, understanding how food choices are made at meal occasions can be a crucial element in developing dietary strategies for better food choices for weight loss, but little research has been conducted in this area.

\* Corresponding author: V. X. Guan, fax +61 02 4221 4844, email [xg885@uowmail.edu.au](mailto:xg885@uowmail.edu.au)

A challenge to exploring food choices at meals is the inter- and intra-individual variation in food consumption. This relates to both the different types of food and the frequency in which they are consumed. Variations in food choices could create inaccessible numbers of food choice combinations, but shortcuts are not the answer. Previous studies suggest that a food combination requires disaggregation into single food groups to provide a more precise intake distribution<sup>(18–21)</sup>. Adding considerations of time intervals (e.g. daily *v.* weekly) while identifying closely related food groups at meals may improve the accuracy of intake estimation.

Advances in data analysis techniques may overcome these challenges for exploring food choice within meals. Meal-based food consumption relationships may be examined through descriptive data mining tools, such as the Apriori algorithm of association rules, which apply a variety of data analysis tools to discover hidden patterns and relationships in a data set<sup>(22)</sup>. The Apriori algorithm has been applied for identifying meal-based food combination patterns in many studies<sup>(20,23–25)</sup>. The Apriori algorithm of association rules is a two-step descriptive method of creating rules to determine associations between items in a data set<sup>(26,27)</sup>. In the first step, the possible combinations of the items are identified based on a specified frequency threshold of combined item sets, referred to as frequent item sets. Second, the identified frequent item sets are used to generate the desired association rules shaped by study interest, with the strength of the identified relationship being an example. The outcomes of the analysis suggest closely related food items at each meal occasion. For example, at the breakfast meal, if bread and butter were reported, milk was also reported. Knowing such a relationship may assist in individualising nutrition counselling for participants, by planning what participants can do to increase the intake frequency of low-energy nutrient-dense foods. It may help determine which food item may be affected by reducing high-energy low-nutrient-dense food intake. The closely related food items may also be suggested to be eaten together to improve dietary adherence to recommendations. Thus, using the Apriori algorithm of association rules may provide better insights into food choices at each meal occasion, and help to translate and develop novel dietetic strategies for weight loss at the individual level. The aim of this study was to explore food choices at meal occasions, reported by a sample of overweight and obese volunteers in a weight-loss trial.

## Method

### Participants

The HealthTrack study was a randomised controlled trial (Australian and New Zealand Clinical Trial Registry ANZCTR N12614000581662) investigating a novel interdisciplinary lifestyle intervention, compared with usual care, on weight loss in overweight and obese adults from the Illawarra region, south of Sydney, Australia. The HealthTrack study was approved by the University of Wollongong/Illawarra Shoalhaven Local Health District Human Research Ethics Committee (HE13/189). Written informed consent from each participant was obtained before data collection. Detailed study protocols are described elsewhere<sup>(28)</sup>.

For the present analyses, screening dietary data from the HealthTrack study before randomisation was analysed. The WHO BMI classifications were applied to determine overweight (BMI of 25.00–29.99 kg/m<sup>2</sup>) and obese (BMI of  $\geq 30.00$  kg/m<sup>2</sup>) participants<sup>(29)</sup>. Dietary intake data were analysed from participants who were overweight and obese at the screening phase.

### Food intake data and food intake data preparation

Self-reported food intake data were obtained from a dietitian-administered diet history interview<sup>(30)</sup> collected between May 2014 and April 2015. The diet history interview protocol was validated previously<sup>(30)</sup>. Participants were asked to recall their usual intake over the past 3-month period during the interview. A food checklist was used to assess commonly omitted food items during the interview. The food intake data were self-reported, reflecting usual weekly consumption recorded on the basis of participant-defined meal occasions, including breakfast, lunch, dinner, beverages, morning tea, afternoon tea, extras, snacks and/or desserts. The intake data were sorted to align with a 7-d equivalent to weekly intake patterns for this study based on reported intake of frequencies. For example, consuming one can of unflavoured tuna (one can weight 70 g as 645.4 kJ) one time per week automatically produces an average daily energy contribution of 92.2 kJ ( $645.4 \text{ kJ}/7 = 92.2 \text{ kJ}$ ). All food intake data were analysed using FoodWorks Professional nutrient analysis software (version 7, 2007; Xyris Software), drawing originally on the AUSNUT 2007 food composition database<sup>(31)</sup>. The analyses did not address data on beverage consumption because of previously identified inconsistencies in the reporting within this data set<sup>(32)</sup>. Thus, 'non-alcoholic beverage' (e.g. tea, coffee, juice, cordial, soft drink or water) and 'alcoholic beverage' (e.g. beers, wines or spirits) were excluded from the analyses. Participant-defined meals other than 'breakfast', 'lunch' and 'dinner', such as morning tea, afternoon tea, desserts, extras and snacks, were all grouped into an other meal. Thus, food intake was grouped into four meal occasions (events): breakfast, lunch, dinner and other meals.

The AUSNUT 2007 food composition database (used in the Australian Health Survey) was the most recent Australian food composition database available when the HealthTrack study commenced. The more recent release of the nested hierarchical food groups of the AUSNUT 2011–2013 food classification system was used for the analyses of this study<sup>(33)</sup>. For this to occur, a matching file was used to translate food items from the AUSNUT 2007 to the AUSNUT 2011–2013 food classification system<sup>(34)</sup>. There are three food group levels in the AUSNUT 2011–2013 food classification system, including the major ( $n = 24$ ), sub-major ( $n = 132$ ) and minor ( $n = 515$ ) food groups<sup>(33,35)</sup>. Examples of food items at each level are presented in the online Supplementary Table S1. At the major food group level, foods are categorised on the basis of the dominant nutrients or ingredients, such as cereal, fruit and vegetable. The sub-major food groups aggregate foods from similar animal/plant species or family, or sharing similar cooking methods and presentations. Detailed and specific characteristics of a food are described at the minor food group level to further differentiate between the food items. For example, a major food group





'cereals and cereal products' is divided into the following sub-major food groups: 'flours and other cereal grains and starches', 'regular breads, and bread rolls', 'breakfast cereals, ready to eat' and so on. 'Grains (other than rice) and grain fractions' and 'rice and rice grain fractions' are included at the minor food group level for 'flours and other cereals grains and starches'. In addition, a total of 106 food groups represent mixed dishes (*n* 12, 21 and 73; at a major, sub-major and minor levels, respectively)<sup>(33)</sup>. Examples of mixed dish food groups include 'pizza, saturated fat ≤5 g/100 g', 'mixed dishes where cereal is the major ingredient' and 'cereal-based products and dishes' at minor, sub-major and major food group levels, respectively.

All the food items listed in the FoodWorks software output were identified using a food group code and food group name for each of the three food grouping levels described above. To prevent duplication, repeated food items at each food group level within meals were removed from the data set to ensure that each food group was only included once for each meal occasion at each food level. For example, four items reported at breakfast – 'bread, from white flour, toasted', 'bread, mixed grain', 'Kellogg's crunchy nut clusters' and 'Kellogg's nutri-grain' – belonged to the same major food group: 'cereals and cereal products'. Thus, only one listing of 'cereals and cereal products' was retained. Further, at the sub-major food group level, two of the four foods belonged to the 'regular breads and bread rolls' sub-major group and two belonged to the 'breakfast cereals, ready to eat' sub-major food group. One food group was retained for each. At the minor food group level, each food item belonged to a different food group; therefore, all the food groups were retained. The food groups at the major, sub-major and minor levels for the above example are shown in Table 1. The frequencies of individual food groups within meals were identified using RStudio, version 1.0.44 (incorporating R, version 3.2.5; The R Foundation for Statistical Computing)<sup>(36)</sup>.

*Statistical analysis*

The present study is a cross-sectional secondary analysis of screening dietary intake data from a 12-month clinical trial for weight loss (the HealthTrack study), with primary results on the randomised group described elsewhere<sup>(37)</sup>.

In the present analyses, the Apriori algorithm of association rules was applied to examine the food choices at meal (events) at each food group level<sup>(26,27)</sup>, using RStudio, version 1.0.44 (incorporating R, version 3.2.5)<sup>(36)</sup>. In the first step of the algorithm, a frequency threshold was used to determine the frequent item sets. This is referred to as support and represents

the percentage of the records containing identified frequent item sets<sup>(26,27)</sup>. In the second step, the support and confidence are used to determine the association rules containing both precursor and consequent items, which reflect the strength of an identified rule<sup>(26,27)</sup>. The association rules are presented to indicate that if the precursor food items are reported the consequent food items are also reported. The confidence of a rule is the percentage of records containing both precursor and consequent items and is calculated as the percentage of the records containing both items divided by the percentage of the records only containing precursor items<sup>(22)</sup>. The confidence indicates how likely the consequent item is presented in the identified association rules<sup>(26,27)</sup>. Higher values of support and confidence imply a stronger relationship for the identified association rule. Another variable used to select the desired rule is referred to as lift, which is used to assess the dependency between precursor and consequent items, and is determined by the confidence of a rule divided by the percentage of records only containing the consequent item<sup>(22)</sup>. A lift >1 indicates that precursor and consequent items are more likely to depend on each other. An example of the two-step algorithm is as follows: in a data set of 100 food choice records, where 80% of records contained both breakfast cereal and milk (*n* 80), 75% of records contained banana (*n* 75) and 65% of the records contained the combination of breakfast cereal, milk and banana (*n* 65), the support of the frequent item set (breakfast cereal, milk and banana) would be 0.65. The confidence of the association rule (e.g. if breakfast cereal and milk are reported, then banana is also reported) is 0.81 (0.65 ÷ 0.8). This indicates that 81% of the times that a participant reports having breakfast cereal and milk, a banana is also reported. The lift of the rule is 1.08 (0.81 ÷ 0.75), which suggests that reporting intake of breakfast cereal and milk depends on the reporting intake of consuming a banana.

In the analyses reported here, the threshold of the possible food group combinations at events (meals) was set as the proportion of participants in the HealthTrack study who were overweight<sup>(38,39)</sup>. In other words, the proportion of the possible combinations of food groups at meals was required to be greater than the proportion of overweight participants in the study. For example, if a total of 20% of participants in the study were overweight, at least 20% of participants in the study would need to report a specific combination of food groups for the food combination to be reported in the present analysis. The percentage of times that a participant reported consuming closely related food groups at meals was determined by the default value for the Apriori algorithm within the R software (0.80)<sup>(36)</sup>.

**Table 1.** Exemplar reported food items at the major, sub-major and minor levels based on food groups of the 2011–2013 Australian Health Survey food classification system<sup>(33)</sup>

Reported weekly consumption	Major food group level	Sub-major food group level	Minor food group level
Bread, from white flour, toasted	Cereals and cereal products	Regular breads, and bread rolls	Breads, and bread rolls, white, mandatorily fortified
Bread, mixed grain	Cereals and cereal products	Regular breads, and bread rolls	Breads, and bread rolls, mixed grain, mandatorily fortified
Kellogg's crunchy nut clusters	Cereals and cereal products	Breakfast cereals, ready to eat	Breakfast cereal, maize based, fortified
Kellogg's nutri-grain	Cereals and cereal products	Breakfast cereals, ready to eat	Breakfast cereal, mixed grain, fortified, sugars >20 g/100 g

Dependency between food groups in the identified food combinations was also assessed<sup>(36)</sup>, which was calculated as follows:

the percentage of times closely related food groups were reported ÷ percentage of records containing related food groups > 1.

Because of food consumption variability, at each event, many closely related food groups may be found in the data set. Thus, to minimise unnecessary complexity, redundant closely related food groups were removed<sup>(40)</sup>. They were determined by comparing closely related food groups at events<sup>(22,40)</sup>. For example, two closely related food groups at breakfast were generated, which contained the following:

- (1) 81% of the times that a participant reported having ready-to-eat breakfast cereal and dairy milk, it was also reported with tropical fruit, and
- (2) 32% of the time that a participant reported having ready-to-eat breakfast cereal, dairy milk and nuts, tropical fruit was also reported.

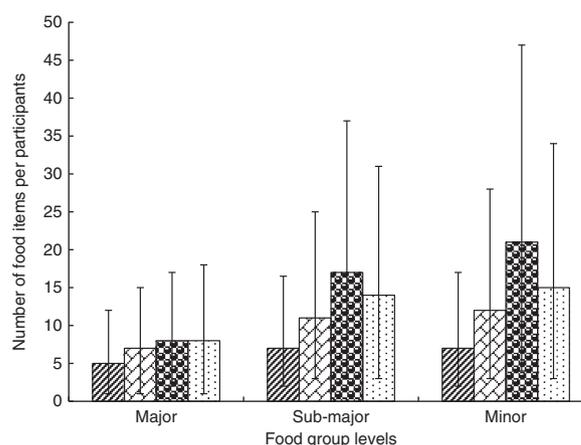
The second combination was based on the first combination but with a much lower percentage of occurrences. Thus, the second combination was considered redundant and removed from further analysis. Subsequently, the major food group combinations for this data set were retained. In addition, to prevent removal of relevant food group combinations, those combinations comprising a high number of food groups reported by at least half of the sample in the study were scrutinised.

## Results

Data for 433 participants screened for the HealthTrack study were analysed (116 male and 317 female) (Table 2). Within the analysed sample, 32% of the participants (*n* 128, twenty-seven male and 111 female) were overweight. Thus, the threshold of the possible food group combinations at events was set at 0.32.

A total of 432 records contained breakfast meal entries and 428 and 432 records contained lunch and dinner meal entries, respectively. A total of 433 records contained other meals. The numbers of meal occasions are presented in Table 2. Data for one participant included reported intake data as a 'main meal' only, which was unable to be differentiated into meal types. A total of thirteen participants reported skipping a meal (nine breakfasts and four lunches). Meal-based energy and macro-nutrient intakes are provided in the online Supplementary Table S2.

Overall, participants reported more food groups for the dinner meal than the breakfast, lunch and other meals at the sub-major and minor food group levels, Fig. 1. At the major food group level, the number of reported food groups was the same for the dinner meal and the other meal (*n* 8). There were slightly more reported food groups at the minor food group level than those at the sub-major food group level for the number of food items per participant (median: 17 (interquartile range (IQR) 14–20) *v.* 21 (17–26)), online Supplementary Table S1.



**Fig. 1.** Number of reported food items per participant per meal occasion. Values are medians and interquartile ranges represented by vertical bars. ▨, Breakfast; ▩, lunch; ▤, dinner; ▦, other meals.

**Table 2.** Participant characteristics of the data set (Mean values and standard deviations; medians and interquartile ranges (IQR))

Characteristics	Total ( <i>n</i> 433)		Male ( <i>n</i> 116, 27%)		Female ( <i>n</i> 317, 73%)	
	Mean	SD	Mean	SD	Mean	SD
Age (years)	43	8.1	43	7.7	43	8.2
Height (m)	1.7	0.1	1.8	0.1	1.6	0.1
Weight (kg)	92.5	15.5	105.2	14.1	87.8	13.2
BMI (kg/m <sup>2</sup> )	32.8	4.2	33.2	4.0	32.6	4.3
Waist circumference (cm)	104.0	11.7	111.8	10.1	101.1	10.9
	Median	IQR	Median	IQR	Median	IQR
Number of meal occasions*						
Breakfast meal occasions*	1	1–1	1	1–1	1	1–1
Lunch meal occasions*	1	1–1	1	1–1	1	1–1
Dinner meal occasions*	1	1–1	1†		1	1–1
Other meal occasions*	1	1–2	1	1–2	1	1–2
Total meal occasions*	4	4–5	4	4–5	4	4–5

\* Excluded beverage intakes.

† The number of dinner meal occasions of male was constant.

### Breakfast meal occasion

At the major food group level, the highest proportion of reported food groups at the breakfast meal was 'cereals and cereal products' (94%) (online Supplementary Table S1). A total of seven items of closely related food groups were identified, with 76% of the participants reporting the combination of 'milk products and dishes' and 'cereal and cereal products'. Overall, 96% of the time that 'milk products and dishes' were reported, 'cereal and cereal products' was also reported.

At the sub-major food group level, the highest proportion of reported food groups was 'regular breads and bread rolls' (74%) (online Supplementary Table S1). There were two closely related food groupings. A total of 42% of the participants reported having a combination of 'regular bread and bread roll' and 'eggs', and 92% of the time the 'eggs' food group was reported 'regular bread and bread roll' was also reported. There were also 35% of participants reporting a combination of 'dairy milk', 'regular bread and bread roll' and 'breakfast cereal, ready to eat', and 82% of the time that 'regular bread and bread roll' and 'breakfast cereal, ready to eat' were reported 'dairy milk' was also reported.

At the minor food group level, the highest proportion of reported food groups was 'egg, chicken' (45%) (online Supplementary Table S1). 'Milk, cow, fluid, reduced fat, <2 g/100 g' was the most frequently reported food group for milk ( $n$  113). The proportions of reported mixed grain, wholemeal and white bread were similar for 25, 25 and 24% of the participants, respectively. There were no closely related food groups identified at the minor food group level for the breakfast meal.

### Lunch meal occasion

At lunch, the highest proportion of reported food groups at the major food group level was 'vegetable products and dishes' (91%) (online Supplementary Table S1). In addition, 71% of the participants reported the combination of 'cereal and cereal products', 'meat, poultry and game products and dishes' and 'vegetable products and dishes', including 51% of the participants also having 'milk products and dishes'. A total of thirteen closely related food groupings were identified. Half of the participants reported having the combination of 'fish and seafood products and dishes' and 'vegetable product and dishes', and 95% of the time that 'fish and seafood products and dishes' was reported 'vegetable product and dishes' was also reported. In addition, half of the participants reported that 'savory sauces and condiment' was combined with either 'cereals and cereal products' or 'vegetable products or dishes', and when 'savory sauces and condiment' was reported 'cereal and cereal products' or 'vegetable products or dishes' was also reported (94 and 96% of the time, respectively).

At the sub-major food group level, the highest proportion of reported food groups was 'regular breads and bread rolls' (76%) (online Supplementary Table S1). There were four closely related food groupings identified at the lunch meal. Approximately half of the participants reported having 'regular bread and bread roll' combined with either 'cheese' or 'leaf and stalk vegetables' or 'processed meat' in food combinations

(51, 49 and 44%, respectively). Furthermore, 92% of the time when 'processed meat' was reported, 'regular bread and bread roll' was also reported, and 'cheese' also at 85% of the time. 'Leaf and stalk vegetables' was reported with 'regular bread and bread roll' of 81% of the time.

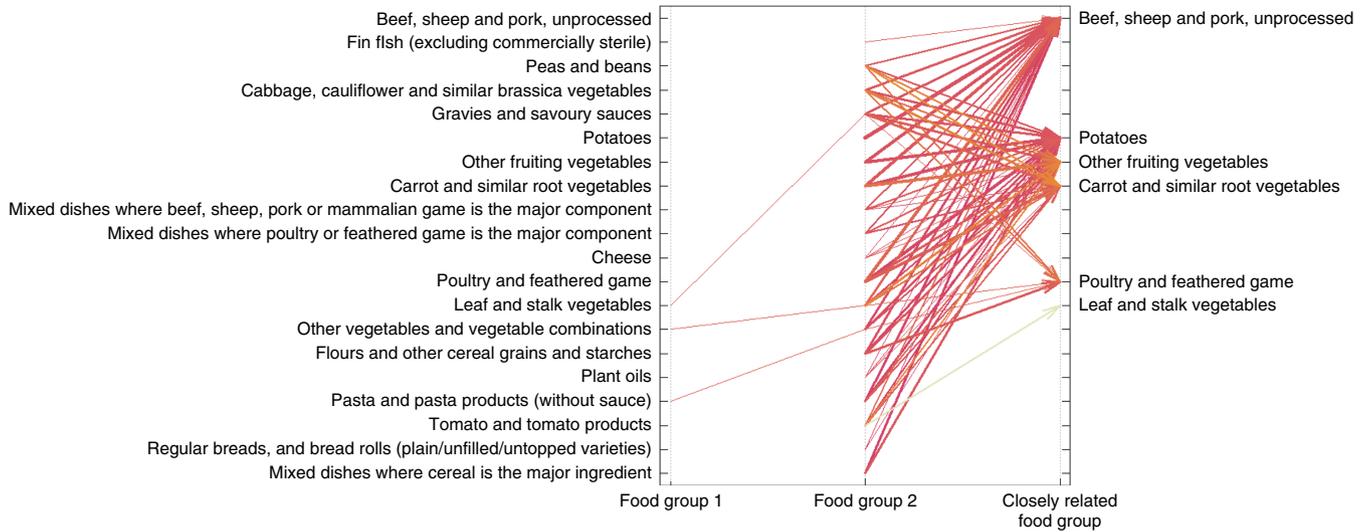
At the minor food group level, the highest proportion of reported food groups was 'leaf vegetables' (59%) (online Supplementary Table S1). Almost half (45%) of the participants reported 'cheese, hard cheese ripened style'. Approximately 40% of the participants reported having 'chicken', 'ham' or 'packed fin fish' and 34% of the participants reported having white bread. There was two closely related food groupings identified related to 'leaf vegetables', with 44% of the participants reporting the combination of 'tomato' and 'leaf vegetables', and the combination of 'other fruiting vegetables' and 'leaf vegetables' by 40% of participants.

### Dinner meal occasion

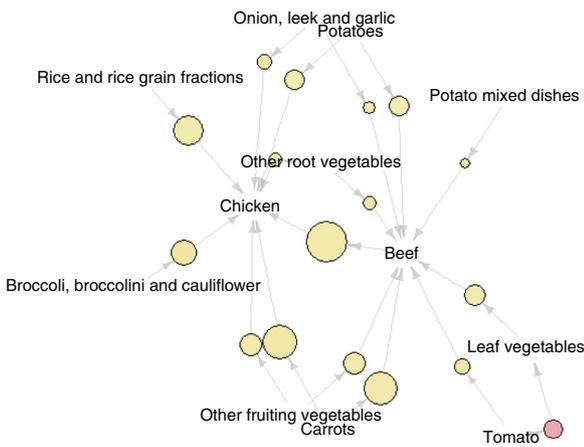
At dinner, the highest proportion of reported food groups at major food group level was 'vegetable products and dishes' (99%) (online Supplementary Table S1). The combination of 'cereal-based products and dishes', 'cereals and cereal products', 'meat, poultry and game products and dishes', 'savory sauces and condiments' and 'vegetable products and dishes' was reported by 52% of the participants. Moreover, although 90% of the participants reported having the combination of 'cereals and cereal products', 'meat, poultry and game products and dishes' and 'vegetable products and dishes', half of the participants reported the additional food group or either 'fish and seafood products and dishes' or 'milk products and dishes' in the combination to form a combination comprising four food groups. In addition, more than 72% of the participants reported having the combination of 'cereal and cereal product', 'savory sauces and condiments' and 'vegetables products and dishes' or the combination of 'meat, poultry and game products and dishes', 'savory sauces and condiments' and 'vegetables products and dishes'. There were fifteen closely related food groupings identified. If either 'cereal-based products and dishes' or 'cereal and cereal products' or 'fats and oils' or 'fish and seafood products and dishes' or 'meat, poultry and game products and dishes' or 'milk products and dishes' or 'savory sauces and condiments' was reported, then 'vegetable products and dishes' was also reported 100% of the time.

At the sub-major food group level, the food group reported most frequently at dinner was 'beef, sheep and pork, unprocessed' (84%) (online Supplementary Table S1). Half of the participants reported having the combination of 'beef, sheep and pork, unprocessed' and 'carrot and similar root vegetables' with the addition of any two food groups of 'potatoes', 'other fruiting vegetables' or 'poultry and feathered game' to form a combination of four food groups. Furthermore, half of the participants reported having the combination of 'beef, sheep and pork, unprocessed', 'potatoes', 'other fruiting vegetables' and 'poultry and feathered game'. A total of sixty-seven closely related food groupings were identified (Fig. 2). When either 'fin fish' or 'leaf and stalk vegetables' or 'poultry and feathered game' was reported, 'beef, sheep and pork, unprocessed' was





**Fig. 2.** Parallel coordinates plot for closely related food groupings for the dinner meal occasion at the sub-major food group level showing sixty-seven items. Arrows represent closely related food groups and the relationship between individual food group or food group combination and its related food group. The width of the arrows represents the percentage of the records containing identified food group combination<sup>(22)</sup>, and the intensity of the colour (from light to dark colour) indicates the numerical percentage value of the time that a participant reported having closely related food groups<sup>(22)</sup>.



**Fig. 3.** Graph-based visualisation of items of closely related food groups for dinner at the minor food groups. Arrows represent closely related food group relationships between individual food groups or food group combinations and its related food group. The size of the sphere represents the percentage of the records containing the identified food group combination<sup>(22)</sup> and the intensity of the colour (from light to dark colour; darker colour indicates a higher value) indicates a numerical percentage value of the time that a participant reported having the item of closely related food groups<sup>(22)</sup>.

also reported 90% of the time. Moreover, 90% of the time where either ‘cabbage, cauliflower and similar brassica’ or ‘pea and beans’ was reported, ‘carrot and similar root vegetables’ was also reported.

At the minor food group level, the highest proportion of reported food groups was ‘chicken’ (76%) (online Supplementary Table S1). There were seventeen closely related food groupings identified, as shown in Fig. 3. When one of the above vegetables was reported, then ‘beef’ or ‘chicken’ was also reported approximately 80% of the time. Moreover, 90% of time that ‘tomato’ was reported, then ‘leaf vegetable’ was also reported.

### Other meal occasions

At the major food group level, the highest proportion of reported food groups was ‘cereal-based products and dishes’ (92%) (online Supplementary Table S1). A total of fourteen closely related food groupings were identified. When either ‘fruit products and dishes’, ‘snacks foods’ or ‘sugar products and dishes’ was reported, ‘confectionery and cereal/nut/fruit/seed bars’ was also reported 89% of the time. At the sub-group food level, the highest proportion of reported food groups was ‘chocolate and chocolate-based confectionery’ (75%) (online Supplementary Table S1). There was one closely related food grouping reported. The combination of ‘sweet biscuits’ and ‘chocolate and chocolate-based confectionery’ was reported by 43% of the participants, and 80% of time that ‘sweet biscuits’ was reported ‘chocolate and chocolate-based confectionery’ was also reported. At the minor food group level, approximately half of the participants reported having ‘chocolate’ (48%) or ‘potato crisps’ (47%). There were no closely related food groupings identified at the minor food group level.

### Discussion

The present work applied a descriptive data mining tool to expose food choices at meal occasions based on reported food item characteristics. The study has allowed identification of closely related food groups at meals based on the reported food item frequencies in screening dietary data in the context of a weight-loss trial. When applying more informative food groups data based on characteristics relating to animal/plant origins, plant family or processing methods rather than only the main nutrients, the dinner meal occasion appeared to be the most complex meal compared with breakfast, lunch and other meals (indicated by the number of reported food items and items of closely related food groups). Thus, focusing on the dinner meal may be an important consideration in dietary counselling for

weight loss in this context at the individual level, although the findings should be interpreted with caution.

The number and definitions of food groups used in the analyses are essential for revealing reported food choices within meals. Studies suggested that relatively similar food groups tend to be consumed at lunch and dinner<sup>(25,41)</sup>. The discrepancy between the present analysis and this may be due to the number and definition of food groups applied. Applying a limited number of food groups may be inadequate to appreciate the full range of food likely to be consumed within a food group. In this study, food choices between meals could be characterised when using the sub-major or minor food groups but not at higher levels. The broad nature of the major food groups means they are only based on the key nutrients or ingredients, resulting in groups such as fruit, vegetables and meat, which are a fairly blunt form of categorisation. Therefore, the precision of food groups may be important in a food choice analysis to develop easier and more practical dietary strategies.

The present exploratory results suggest that foods were chosen differently at main meals (e.g. breakfast, lunch and dinner). In our analysis, food choices at breakfast were consistent with the literature from other Western cultures, comprising cereal (e.g. bread or ready to eat breakfast cereal) and milk<sup>(24,42)</sup>. It is also consistent with literature suggesting that the increased BMI appears to be closely associated with consumption of eggs at breakfast<sup>(42,43)</sup>. However, skipping breakfast is also suggested to be related to increased BMI<sup>(45)</sup>, but only 2% of participants in the present sample reported skipping breakfast. Although different varieties of meat were chosen at lunch and dinner, at the lunch meal people were more likely to choose bread as the food at the centre of food combinations, whereas at dinner unprocessed meat took this position. Choosing meat in meals reflects the Australian tradition of eating habits, aligning with the nation's historical context aligned to meat consumption<sup>(44)</sup>.

The findings may also help in developing clinical dietetic strategies for better food choices and behaviour change at the individual level through understanding food choices at meal occasions, particularly for dinner meal occasions. To date, accumulated evidence has suggested that daily energy intake distribution of main meals plays an important role in the effectiveness of achieving weight-loss goals and adhering to a weight-loss plan in the long term<sup>(17)</sup>, particularly for reducing intake at dinner<sup>(45,46)</sup>. The daily energy intake at dinner appears to be positively associated with the risk of obesity<sup>(47)</sup>, contributed by the desynchronisation of the fasting/feeding cycles in the circadian system<sup>(48)</sup>. Therefore, effective strategies for reducing intake at dinner are required for weight loss. In addition, increasing the number of low-energy-dense food items consumed irrespective of intake of quantities is suggested as an effective strategy to lose weight<sup>(49)</sup>. Given that vegetables are a nutrient-dense food group but are low in energy, increasing vegetable consumption in the diet tends to have desirable effect on controlling energy intakes<sup>(50)</sup>. However, strategies used to promote consumption of vegetables tend to have limited influence on weight loss<sup>(51)</sup>. This may be owing to the wide variety of vegetable types and food forms available and a substantial effort is required to incorporate vegetables into a diet on a daily basis. Moreover, concurrent strategies on

restricting other foods when promoting the intake of vegetables is also required for weight loss<sup>(52)</sup>. This may imply that being more specific about the actual foods to be consumed is required to develop sustainable dietary strategies at the individual level. The present results suggest that participants chose a different variety of vegetables at lunch and dinner meal occasions, where a greater variety of vegetables was consumed with unprocessed meat at dinner. Thus, the current findings may help to provide more practical clinical dietary strategies to improve the intake of vegetables for weight loss, such as reducing the intake of unprocessed meat with increasing the specific vegetable intakes (e.g. carrot and broccoli).

To our knowledge, using a data mining method based on a nested hierarchical food grouping system has not been previously performed to explore food choices at meal occasions in a sample of overweight and obese participants. When applying data mining tools, a large database is commonly required because of the need to split the data set for training and testing purposes<sup>(22)</sup>. However, this may not be necessary when using descriptive data mining tools, such as the association rule algorithms<sup>(22)</sup>, providing an opportunity to analyse smaller data sets containing detailed data. Detailed food choices at meals reflecting an individual's habitual food consumption are required to explore food combinations. A diet history interview dietary assessment method uses an open-ended interview approach asking and probing participants to describe habitual food consumption generally from the first meal of the day through to the end of the day<sup>(53)</sup>. This may imply that food intake data generated through a diet history interview are more likely to provide food choices at self-defined meals. The richness of dietary intake data generated from a diet history interview at the screening phase of an intervention study also makes it possible to investigate food choices at meal occasions that are unlikely to be observed using other dietary assessment methods, such as the FFQ. Although studies using objective biomarkers of intake have demonstrated that overweight and obese populations tend to misreport energy intakes using self-reported dietary assessment tools<sup>(54)</sup>, the present study somewhat overcame this issue by focusing on food items reported within meals, rather than consumption quantities and frequencies.

However, there are several limitations for the present study. The study was a secondary analysis of screening data of a randomised controlled trial that was designed to answer a different question. The meals collected in the diet history were defined by participants, and specific names of other meals were not required during the dietary intake data collection and coding. A preliminary study in this sample found that the data coders clearly coded breakfast, lunch and dinner, but they applied different approaches to code other meal occasions<sup>(32)</sup>. For example, specific meal names such as morning tea or afternoon tea were used to describe other meal occasions. Snacks were also used in some circumstances to code other meal occasions together. It appeared that analysing the specific meal of other meal occasions was out of the scope of the primary aim of the clinical trial. There are also inherent limitations of self-reported dietary intake data derived by diet history interview, such as recall bias and social desirable reporting behaviour<sup>(55,56)</sup>. Moreover, the preliminary study of the present



analyses also demonstrated that the beverages were reported inconsistently<sup>(32)</sup>. For example, the beverages were reported alone, with food or with food and alone<sup>(32)</sup>. This might indicate that the present study was unable to offer any evidence on food combinations related to beverage intakes. The analysis conducted here was based on a small sample of overweight and obese individuals volunteering for a clinical trial, which may not be representative of the overweight and obese population. Reproducing this analysis in a larger sample is warranted to determine the applicability of the method and the findings. In addition, not all the mixed dishes were disaggregated into food groups in the present data set, as it was out of the scope of the primary aim of the clinical trial. The format and composition of mixed dishes may also be required to examine further in relation to food combination at meals. Therefore, these issues may be required to be addressed in future studies to offer more robust evidence for dietary counselling for weight loss. Notwithstanding this, the 142 closely related food groupings at meal occasions may provide examples on what foods are more likely to be chosen to help researchers develop easier and more practical dietary strategies to use in weight-loss counselling.

In conclusion, this study highlighted the food choices at main meal occasions made by overweight and obese study volunteers, which presented insights for counselling on dietary change to lose weight at the individual level. To date, few studies have investigated the food items consumed at meal occasions and their associations. The present study used a descriptive data mining tool (the Apriori algorithm of association rules) and screening usual dietary intake data from a clinical trial to expose these food choices at meal occasions at a much deeper level. Using a nested hierarchical food grouping system to examine food choices at meal occasions revealed that food choice analyses need to consider the level of detail encapsulated within the food groups. Exploring food choices at meal occasions may complement current dietary strategies on weight management and assist in the development of easier and more practical strategies for food choices to support weight loss.

### Acknowledgements

The authors thank all the participants and the HealthTrack study team.

The HealthTrack study was funded by the Illawarra Health and Medical Research Institute and California Walnut Commission. The funding bodies had no involvement in the data analysis, data interpretation or writing of the manuscript.

V. X. G., Y. C. P. and E. P. N. formulated the research question and designed the study. V. X. G. and M. J. B. analysed the data. V. X. G. wrote the manuscript. Y. C. P., E. P. N., M. J. B. and L. C. T. reviewed the manuscript. All authors read and approved the final manuscript.

None of the authors has any conflicts of interest to declare.

### Supplementary material

For supplementary material/s referred to in this article, please visit <https://doi.org/10.1017/S0007114518001587>

### References

1. Finucane MM, Stevens GA, Cowan MJ, *et al.* (2011) National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* **377**, 557–567.
2. Ng M, Fleming T, Robinson M, *et al.* (2013) Global, regional, and national prevalence of overweight and obesity in children and adults during, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **384**, 766–781.
3. World Health Organization (2016) Obesity and Overweight – Key Facts. Geneva: WHO. <http://www.who.int/mediacentre/factsheets/fs311/en/> (accessed August 2017).
4. Australian Bureau of Statistics (2015) *National Health Survey: First Results, 2014–15*. Canberra: ABS.
5. The Global Burden of Disease (GBD) 2015 Obesity Collaborators (2017) Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med* **377**, 13–27.
6. National Health and Medical Research Council (2013) *Clinical Practice Guidelines for the Management of Overweight and Obesity in Adults, Adolescents and Children in Australia*. Melbourne: NHMRC.
7. Grima M & Dixon J (2013) Obesity recommendations for management in general practice and beyond. *Aust Fam Physician* **42**, 532–541.
8. Sacks FM, Bray GA, Carey VJ, *et al.* (2009) Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N Engl J Med* **360**, 859–873.
9. Grafenauer SJ, Tapsell LC, Beck EJ, *et al.* (2013) Baseline dietary patterns are a significant consideration in correcting dietary exposure for weight loss. *Eur J Clin Nutr* **67**, 330–336.
10. Heymsfield SB & Wadden TA (2017) Mechanisms, pathophysiology, and management of obesity. *N Engl J Med* **376**, 254–266.
11. Barte JCM, Bogt NCWT, Bogers RP, *et al.* (2010) Maintenance of weight loss after lifestyle interventions for overweight and obesity, a systematic review. *Obes Rev* **11**, 899–906.
12. Looney SM & Raynor HA (2013) Behavioral lifestyle intervention in the treatment of obesity. *Health Serv Insights* **6**, HSI.S10474.
13. Gatenby SJ (1997) Eating frequency: methodological and dietary aspects. *Br J Nutr* **77**, S7–S20.
14. Gorgulho BM, Pot GK, Sarti FM, *et al.* (2016) Indices for the assessment of nutritional quality of meals: a systematic review. *Br J Nutr* **115**, 2017–2024.
15. Gorgulho BM, Pot GK, Sarti FM, *et al.* (2017) Main meal quality in Brazil and United Kingdom: similarities and differences. *Appetite* **111**, 151–157.
16. 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**, 1–21.
17. St-Onge M-P, Ard J, Baskin ML, *et al.* (2017) Meal timing and frequency: implications for cardiovascular disease prevention: a scientific statement from the American Heart Association. *Circulation* **135**, e96–e121.
18. Krebs-Smith SM, Cronin FJ, Haytowitz DB, *et al.* (1990) Contributions of food groups to intakes of energy, nutrients, cholesterol, and fiber in women's diets: effect of method of classifying food mixtures. *J Am Diet Assoc* **90**, 1541–1546.
19. Prynne CJ, Wagemakers JJMF, Stephen AM, *et al.* (2008) Meat consumption after disaggregation of meat dishes in a cohort of British adults in 1989 and 1999 in relation to diet quality. *Eur J Clin Nutr* **63**, 660–666.
20. 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 Compos Anal* **22**, 130–136.

21. Probst YC, De Agnoli K, Batterham M, *et al.* (2009) Video-recorded participant behaviours: the association between food choices and observed behaviours from a web-based diet history interview. *J Hum Nutr Diet* **22**, 21–28.
22. Han J, Pei J & Kamber M (2011) *Data Mining: Concepts and Techniques*. Waltham, MA: Elsevier.
23. Hearty ÁP & 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**, 1632–1642.
24. Woolhead C, Gibney MJ, Walsh MC, *et al.* (2015) A generic coding approach for the examination of meal patterns. *Am J Clin Nutr* **102**, 316–323.
25. Murakami K, Livingstone MBE & Sasaki S (2017) Establishment of a meal coding system for the characterization of meal-based dietary patterns in Japan. *J Nutr* **147**, 2093–2101.
26. Agrawal R, Imieliski T & Swami A (1993) Mining association rules between sets of items in large databases. *SIGMOD Rec* **22**, 207–216.
27. Agrawal R & Srikant R (1994) Fast algorithms for mining association rules. *Proc 20th Int Conf Very Large Data Bases, VLDB* **1215**, 487–499.
28. Tapsell LC, Lonergan M, Martin A, *et al.* (2015) Interdisciplinary lifestyle intervention for weight management in a community population (HealthTrack study): study design and baseline sample characteristics. *Contemp Clin Trials* **45**, 394–403.
29. World Health Organization (1995) *Physical Status: The Use and Interpretation of Anthropometry. Report of a WHO Expert Committee. Technical Report Series* no. 854. Geneva: WHO.
30. Martin GS, Tapsell LC, Denmeade S, *et al.* (2003) Relative validity of a diet history interview in an intervention trial manipulating dietary fat in the management of type II diabetes mellitus. *Prev Med* **36**, 420–428.
31. Food Standards Australia New Zealand (2008) *AUSNUT 2007. Australian Food, Supplement and Nutrient Database for Estimation of Population Nutrient Intakes*. Canberra: FSANZ.
32. Guan V, Probst Y, Neale E, *et al.* (2016) The feasibility of analysing food consumption combinations from overweight and obese participants of weight loss clinical trials. *CEUR Workshop Proc* **1683**, 1–6.
33. Food Standards Australia New Zealand (2014) AUSNUT 2011–13 food measures database file. <http://www.foodstandards.gov.au/science/monitoringnutrients/ausnut/classificationofsupps/Pages/default.aspx> (accessed February 2016).
34. Neale EP, Probst YC & Tapsell LC (2016) Development of a matching file of Australian food composition databases (AUSNUT 2007 to 2011–13). *J Food Compos Anal* **50**, 30–35.
35. Food Standards Australia New Zealand (2014) *AUSNUT 2011–13. Australian Food, Supplement and Nutrient Database for Estimation of Population Nutrient Intakes*. Canberra: FSANZ.
36. R Core Team (2016) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>
37. Tapsell LC, Lonergan M, Batterham MJ, *et al.* (2017) Effect of interdisciplinary care on weight loss: a randomised controlled trial. *BMJ Open* **7**, e014533.
38. Brossette SE, Sprague AP, Hardin JM, *et al.* (1998) Association rules and data mining in hospital infection control and public health surveillance. *J Am Med Inform Assoc* **5**, 373–381.
39. Ramezankhani A, Pournik O, Shahrabi J, *et al.* (2015) An application of association rule mining to extract risk pattern for type 2 diabetes using Tehran Lipid and Glucose Study database. *Int J Endocrinol Metab* **13**, e25389.
40. Pasquier N, Taouil R, Bastide Y, *et al.* (2005) Generating a condensed representation for association rules. *J Intell Inform Syst* **24**, 29–60.
41. de Castro JM (2009) When, how much and what foods are eaten are related to total daily food intake. *Br J Nutr* **102**, 1228–1237.
42. Cho S, Dietrich M, Brown CJP, *et al.* (2003) The effect of breakfast type on total daily energy intake and body mass index: results from the Third National Health and Nutrition Examination Survey (NHANES III). *J Am Coll Nutr* **22**, 296–302.
43. Horikawa C, Kodama S, Yachi Y, *et al.* (2011) Skipping breakfast and prevalence of overweight and obesity in Asian and Pacific regions: a meta-analysis. *Prev Med* **53**, 260–267.
44. Sui Z, Raubenheimer D & Rangan A (2017) Exploratory analysis of meal composition in Australia: meat and accompanying foods. *Public Health Nutr* **20**, 2157–2165.
45. Madjid A, Taylor MA, Delavari A, *et al.* (2016) Beneficial effect of high energy intake at lunch rather than dinner on weight loss in healthy obese women in a weight-loss program: a randomized clinical trial. *Am J Clin Nutr* **104**, 982–989.
46. Jakubowicz D, Barnea M, Wainstein J, *et al.* (2013) High Caloric intake at breakfast vs. dinner differentially influences weight loss of overweight and obese women. *Obesity (Silver Spring)* **21**, 2504–2512.
47. Bo S, Musso G, Beccuti G, *et al.* (2014) Consuming more of daily caloric intake at dinner predisposes to obesity. A 6-year population-based prospective cohort study. *PLOS ONE* **9**, e108467.
48. Albrecht U (2017) The circadian clock, metabolism and obesity. *Obes Rev* **18**, 25–33.
49. Vadiveloo M, Parker H & Raynor H (2017) Increasing low-energy-dense foods and decreasing high-energy-dense foods differently influence weight loss trial outcomes. *Int J Obes (Lond)* **42**, 479.
50. Tapsell LC, Dunning A, Warensjo E, *et al.* (2014) Effects of vegetable consumption on weight loss: a review of the evidence with implications for design of randomized controlled trials. *Crit Rev Food Sci Nutr* **54**, 1529–1538.
51. Kaiser KA, Brown AW, Bohan Brown MM, *et al.* (2014) Increased fruit and vegetable intake has no discernible effect on weight loss: a systematic review and meta-analysis. *Am J Clin Nutr* **100**, 567–576.
52. Mytton OT, Nnoaham K, Eyles H, *et al.* (2014) Systematic review and meta-analysis of the effect of increased vegetable and fruit consumption on body weight and energy intake. *BMC Public Health* **14**, 886.
53. Burke BS (1947) The dietary history as a tool in research. *J Am Diet Assoc* **23**, 1041–1046.
54. Lichtman SW, Pisarska K, Berman ER, *et al.* (1992) Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *N Engl J Med* **327**, 1893–1898.
55. Smith AF (1993) Cognitive psychological issues of relevance to the validity of dietary reports. *Eur J Clin Nutr* **47**, Suppl. 2, S6–S18.
56. Hébert JR (2016) Social desirability trait: biaser or driver of self-reported dietary intake? *J Acad Nutr Diet* **116**, 1895–1898.

