January 2016

The feasibility of analysing food consumption combinations from overweight and obese participants of weight loss clinical trials

Vivienne Guan
University of Wollongong, xg885@uowmail.edu.au

Yasmine Probst
University of Wollongong, yasmine@uow.edu.au

Elizabeth Neale
University of Wollongong, elizan@uow.edu.au

Allison Martin
University of Wollongong, allisonh@uow.edu.au

Linda C. Tapsell
University of Wollongong, ltapsell@uow.edu.au

The feasibility of analysing food consumption combinations from overweight and obese participants of weight loss clinical trials

Abstract

Overweight and obesity is a global epidemic. Investigating food consumption combinations (FCCs) may offer useful insights into addressing eating behaviours to manage overweight and obesity. Using food intake data generated from a detailed dietary assessment method allows advanced analytical methods to be employed. Food intake data collected by a diet history interview appears to be more precise in terms of capturing the usual food intakes of individuals. Exploration of FCCs can be conducted using the Apriori algorithm, but this method is dependent on correct data preparation. Given the uncertainties related to collecting food intake data via diet history interviews, the aim of this study was to explore the feasibility of using food intake data derived from diet history interviews from three weight-loss clinical trials to investigate FCCs. A 10% random sample (n=62) of baseline paper-based diet history records, reflecting usual food intake by meal, from three registered clinical trials (n=617) were extracted. FCCs were assessed by considering the sum of single food items consumed at the same time or in the same occasion using the United States Department of Agriculture Food Combination Codes and the nested hierarchical food groups of the 2011-13 Australian Health Survey food classification system. FCCs were identified in all diet history data records at the major food group level. A proportion of FCCs for the dinner meal (n=13) were unable to be assessed at the specific food level due to limited detail for meat-containing FCCs. FCCs for the dinner meal created more challenges for accurately distinguishing and naming FCCs. Given the complexity of beverage reporting, combinations of foods and beverages were not revealed in the selected data set. In conclusion, despite a lack of meat-containing FCCs at dinner and food-beverage combinations, the food intake data collected using the diet history interview method can feasibly be used to investigate FCCs.

Disciplines

Medicine and Health Sciences

Publication Details

Guan, V., Probst, Y., Neale, E., Martin, A. & Tapsell, L. (2016). The feasibility of analysing food consumption combinations from overweight and obese participants of weight loss clinical trials. In K. Barbuto, L. Schaper & K. Verspoor (Eds.), Proceedings of the Scientific Stream at the Health Data Analytics Conference 2016 Brisbane, Australia (pp. 1-6). Germany: CEUR-WS.

The feasibility of analysing food consumption combinations from overweight and obese participants of weight loss clinical trials

Vivienne GUAN[a], Yasmine PROBST[a,b], Elizabeth NEALE[a,b], Allison MARTIN[a,b] and Linda TAPSELL[a,b]
[a]*School of Medicine, Faculty of Science, Medicine and Health, University of Wollongong, Australia*
[b]*Illawarra Health and Medical Research Institute, University of Wollongong, Australia*

Abstract. Overweight and obesity is a global epidemic. Investigating food consumption combinations (FCCs) may offer useful insights into addressing eating behaviours to manage overweight and obesity. Using food intake data generated from a detailed dietary assessment method allows advanced analytical methods to be employed. Food intake data collected by a diet history interview appears to be more precise in terms of capturing the usual food intakes of individuals. Exploration of FCCs can be conducted using the Apriori algorithm, but this method is dependent on correct data preparation. Given the uncertainties related to collecting food intake data via diet history interviews, the aim of this study was to explore the feasibility of using food intake data derived from diet history interviews from three weight-loss clinical trials to investigate FCCs. A 10% random sample (n=62) of baseline paper-based diet history records, reflecting usual food intake by meal, from three registered clinical trials (n=617) were extracted. FCCs were assessed by considering the sum of single food items consumed at the same time or in the same occasion using the United States Department of Agriculture Food Combination Codes and the nested hierarchical food groups of the 2011–13 Australian Health Survey food classification system. FCCs were identified in all diet history data records at the major food group level. A proportion of FCCs for the dinner meal (n=13) were unable to be assessed at the specific food level due to limited detail for meat-containing FCCs. FCCs for the dinner meal created more challenges for accurately distinguishing and naming FCCs. Given the complexity of beverage reporting, combinations of foods and beverages were not revealed in the selected data set. In conclusion, despite a lack of meat-containing FCCs at dinner and food-beverage combinations, the food intake data collected using the diet history interview method can feasibly be used to investigate FCCs.

Keywords. Food intake, food consumption combinations, diet history, association rule

Introduction

Overweight and obesity form a global epidemic with the prevalence of these conditions increasing since 1980 [1]. By 2010, overweight and obesity contributed to 23% of disability-adjusted life-years for ischaemic heart disease, but also to 3.4 million deaths

and 4% of years of life lost globally [2]. Although great effort has gone into managing overweight and obesity including dietary guidelines that use dietary models to target certain nutrients and foods [3], there is no reported successful case of turning the problem around [4]. Research on nutrients, single foods and food groups has been used as a basis for exploring diet-disease relationships that underpin obesity management regimens. However, most foods are consumed in combinations whether as meals or snacks [5]. For example, in Western diets, a savoury biscuit may be eaten with cheese; steamed vegetables might be consumed with roasted meat for dinner. Thus, investigations of food consumption combinations (FCCs) may offer an alternate strategy for examining eating behaviours to manage overweight and obesity.

A number of methods for the collection of dietary data exist, including the diet history interview, food record, and 24 hour recall. The diet history interview method employs an open-ended interviewer-administrated approach to collecting data about an individual's usual food intake over a defined time period [6]. During the interview, a trained interviewer asks the interviewee to describe food consumption generally from the start of the day, such as the first food item consumed after waking, through to the end of the day, before sleep. Based on the reported information, the interviewer applies probing questions to assist the interviewee to recall and report what had been consumed with the reported foods and in meals. Thus, data generated from a diet history interview is more precise in terms of capturing the usual food combination intakes of individuals than weighed food records and 24-hour recalls. However, the weakness of the diet history interview is that the interviewer asks the interviewee to make judgements about the food items and combinations through the types and timing of probing questions [7]. The effort and expertise of the interviewer, as well as the interaction between interviewer and interviewee can play significant roles in the information captured. For example, an experienced interviewer is able to ask further probing questions based on the interviewee's cues and responses to capture the 'actual' food consumption. This might imply that information captured through a diet history interview may not be compared in the same manner as other forms of dietary assessment.

With respect to the research trials themselves, different types of biases and errors have been identified in obesity research [8]. The study design and analysis methods used in obesity research have been criticised for their limited ability to translate to health outcomes [9]. For example, dietary research that focuses only on single nutrients may overlook the effects of combinations of foods or dietary patterns [10]. Conversely, analytical methods required to examine food combinations are complicated and less well explored than statistical methods [11]. This may imply that advanced analytical methods need to be employed to contribute robust evidence for the associations between body weight and food intakes, to provide meaningful insights towards more effective obesity management strategies.

Examining FCCs is one way of looking at food patterns in a trial. FCCs can be explored using association rules which are data mining tools used for identifying certain relationships or combinations in a large data set [12]. To date, there are only two published studies using a modified Apriori algorithm, which is one of algorithms of association rules that successfully identified FCCs within a meal [13, 14]. In the Apriori algorithm, two steps are conducted. The first step identifies frequent food item sets in a meal. The frequent item sets are supported by a pre-defined support level, which is the proportion of cases in the database containing the identified food item sets. The second step is to generate rules by using identified frequent food items sets. Apriori algorithms have previously been used in the literature to explore FCCs.

Woolhead et al only performed step 1 of the algorithm [14], while Burden et al conducted both steps to generate the FCCs [13]. Apart from the Woolhead study using pre-defined food groups, and Burden using specific food items, the discrepancies between methods might be due to differences in the study aims. Using identified FCCs to develop a generic meal code system to cover inter- and intra-variabilities of food consumptions, required the result to provide a wide coverage of the possible FCCs. On other hand, Burden et al aimed to use the outcome to develop software to assist dietary data collection, where users could select a food item from a drop-down menu. This required results to provide more accurate combination descriptions to improve the efficacy of the questioning scheme for the software. Therefore, it appears that the two steps of the Apriori algorithm may be used to investigate FCC behaviours to provide robust evidence on the association between body weight and food intake.

Data preparation is a critical step of data analyses using data mining tools, to accurately perform subsequent analyses [15]. Although FCCs were successfully investigated using food record [14] and 24 hour recall [13] dietary data, there are a number of uncertainties related to using food intake data collected by diet history interviews. This paper aimed to better understand the feasibility of using food intake data collected using a diet history interview method to examine the FCCs in an overweight and obese population. The paper will address two specific objectives: (1) to determine whether FCCs can be successfully identified using diet history interview records from pooled data pertaining to three clinical trials; (2) to examine challenges related to determining FCCs.

1. Methods

The basis of this work was diet history data from clinical trial participants. Hertzog has suggested that a 10-15% sample for a testing group is sufficient to test the feasibility of a study [16]. Thus, a 10% random sample (n=62) of baseline paper-based diet history records of participants from pooled analyses of three registered weight-loss clinical trials (n=617) were extracted as a pilot. Details of the trials have been described elsewhere [17-19]. In an open-ended interviewer-administrated interview, self-reported food intake data reflecting usual weekly (7 days) consumption was collected by Accredited Practising Dietitians (APDs), followed by a food list to systematically check for omitted food items. Meals, food items and their quantities were recorded on a paper-based diet history proforma.

FCCs were defined as the sum of single food items consumed together at the same time or in the same occasion, for example, toast, jam, and peanut butter reported as eaten together at the breakfast meal. The combination was noted as one event of a FCC. Firstly, FCC events of the extracted diet history records were identified and grouped by meal (breakfast, lunch, dinner, mid-meals and beverages). Secondly, the nested hierarchical food groups of the 2011–13 Australian Health Survey food classification system, including the major, sub-major and minor groups, were used to assist in assessing the food items of identified FCCs [20]. At the major food group level, there are 24 groups, in which foods are grouped on the basis of the main nutrient or ingredient (such as fruit products and dishes) [20, 21]. Foods are categorised at the sub-major food group level based on species, family, and cooking and/or preparation methods (for example citrus fruit) [20, 21]. Detailed or specific information related to a food is included at the minor food group level, in order to identify and distinguish

foods from each other (for example an orange versus a lemon) [20, 21]. Lastly, the United States Department of Agriculture (USDA) Food Combination Codes Scheme, hereafter referred to the USDA codes was used to guide the categorisation of the identified FCCs [22] (Table 1).

Table 1. The United States Department of Agriculture Food Combination Codes Scheme [22] and examples

Code	Description	Example
00	Non-combination	Chocolate consumed alone
01	Beverage with additions	Tea with milk and sugar
02	Cereal with additions	Ready-to-eat cereal (Weet-bix) with milk and banana
03	Bread/baked products with additions	Bread with margarine and jam
04	Salad	Lettuce, tomato, cucumber and avocado with dressing
05	Sandwiches	Bread, butter, ham, cheese, tomato, lettuce and mayonnaise
06	Soup	Pumpkin soup or ready-to-eat soup made by powder (liquid food)
07	Frozen meals	Lean Cuisine
08	Ice cream/frozen yoghurt with additions	Ice cream with chocolate sauce
09	Dried beans and vegetable with additions	Lentil curry (dried beans as the main ingredient for the combination)
10	Fruit with additions	Strawberry with yoghurt
11	Tortilla products	Taco
12	Meat, poultry, fish	Chicken and vegetable casserole
13	Lunchables	Vita-weat biscuits with canned tuna
90	Other mixtures	Omelette (eggs, cheese, ham and tomato)

2. Results

At the major food group level, FCCs were successfully identified in all extracted pilot diet history records, such as meat with vegetables and starchy foods (for example rice, pasta and potato products). Although FCCs of breakfast, lunch, mid-meals and beverages were successfully identified at the sub-major and the minor food group levels, FCCs were unable to be identified at the sub-major and minor level at the dinner meal in 13 (21%) diet history records. This occurred when variations in meat (beef, lamb, pork, and chicken) were recorded together. Thus, the specific meat item was unable to be matched with subsequent vegetable and starchy foods to articulate FCCs that were consumed together with the specific meat type.

Applying the USDA codes identified that 84% (n=52) of cases reported cereal with additions (such as milk, sugar and/or fruit) and 55% (n=34) reported bread/baked products with additions (such as spreads and eggs) at breakfast. A total of 92% (n=57) of cases reported sandwiches at lunch. The number of variations in FFCs for dinner was high (ranging from 1 to 9 combinations). However, the available USDA codes were unable to cover all FCCs from the extracted dataset, particularly for dinner. For example mixed dishes such as pasta dishes and shepherd's pie were often reported for dinner, but no USDA codes could be used to accurately reflect these FCCs.

The challenge identified for assessment of FCCs was in assessing the combination of foods and beverages. Beverages were found to be reported with food (n=30), alone (n=53), both with food and alone (n=13), and in the food frequency checklist at the end of the diet history interview proforma (n=30). Additionally, beverages of eight data records were reported with food, alone and in the food checklist. There was no

reporting trend for the characteristics of reporting non-alcoholic and alcoholic beverages. Therefore, the reported combination of food and beverages may be unable to be assessed using the available diet history data.

3. Discussion

The aim of this study was to determine the feasibility of using available food intake data collected by the diet history interview method to assess FCCs of overweight and obese participants in weight-loss clinical trials. Methodological issues associated with identifying FCCs and challenges relating to data preparation used to undertake the analyses were investigated. The results demonstrate that using diet history data provides sufficiently detailed information on FCCs of breakfast, lunch, mid-meals and beverages. However, the analysis process of this study has shown the complexity of preparing dietary intake data to investigate FCCs, due to the variation in food consumption between and within individuals. Specific challenges encountered related to determining FCCs, specifically meat-containing FCCs, at dinner and the consumption of beverages and food in combination.

The findings indicate that meat-containing FCCs at dinner could only be identified at the major food group level (meat food group), rather than the minor food group level (such as beef or chicken). This is due to red meat such as beef, lamb, pork, and veal and white meat such as chicken being reported as alternate FCC options, but being separated in the food classification system [20]. This may indicate that some eating behaviours can still be identified, such as meat with vegetables and starchy foods; however, the specific meat item consumption may not. Thus, given the high proportion of unsuccessful FCCs for the dinner meal at the minor group level, eating occasions such as main meals (breakfast, lunch and dinner) may need to be analysed separately, particularly for the meat-containing FCCs at dinner.

In addition, this pilot demonstrated that the USDA codes might be too simple to distinguish specific FCCs in different countries, suggesting that different eating habits of different countries may need to be taken into consideration. The USDA codes could only be used as a guide to categorise FCCs of overweight and obese participants of weight loss clinical trials. Specific FCCs aligned with particular cultural food codes may need to be created to accurately reflect true FCCs. Therefore, in order to accurately perform subsequent analyses using the Apriori algorithm, additional categories are required to accurately reflect true FCCs for an Australian overweight and obese population eating context, such as pasta dishes and pie dishes.

Furthermore, food and beverage consumption combinations could not be identified in the available data set. Although beverages may be consumed alone, such as coffee with milk at morning tea, the results indicated that the beverage reporting practice in the data set was not consistent. Due to the complexity of beverage reporting found here, further investigations may need to focus on foods only.

In conclusion, food intake data collected by diet history interview can be used successfully for investigating FCCs for breakfast, lunch, mid-meals and single beverages. Meat-containing FCCs at dinner and the combined consumption of beverages and food present challenges for identifying FCCs. To apply the present methods to investigate FCCs in future studies, such as the Apriori algorithm, the tool used to assess FCCs need to be modified or developed reflecting eating behaviours of targeted population.

References

[1] Finucane, M.M., et al., 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, *The Lancet* **377(9765)** (2011), 557-567.

[2] Lim, S.S., et al., A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2010, *The Lancet* **380(9859)** (2010), 2224-2260.

[3] Hession, M., et al., Systematic review of randomized controlled trials of low-carbohydrate vs. low-fat/low-calorie diets in the management of obesity and its comorbidities, *Obesity Reviews* **10(1)** (2009), 36-50.

[4] Ng, M., et al., Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013, *The Lancet* **384(9945)** (2013), 766-781.

[5] Scholderer, J., et al., Meal mapping, *Food Quality and Preference* **30(1)** (2013), 47-55.

[6] Martin, G.S., et al., Relative validity of a diet history interview in an intervention trial manipulating dietary fat in the management of Type II diabetes mellitus, *Preventive Medicine* **36**(4) (2003), 420-428.

[7] Thompson, F.E. and A. Subar, *Dietary assessment methodology*, in *Nutrition in the Prevention and Treatment of Disease*, A.M. Coulston, C.J. Boushey, and M.G. Feruzzi, Editors. Elsevier: Oxford, England, 2013, 5-46.

[8] Ioannidis, J.P.A., Biases in obesity research: Identify, correct, endorse, or abandon effort? *Obesity* **24(4)** (2016), 767-768.

[9] George, B.J., et al., Common scientific and statistical errors in obesity research, *Obesity* **24(4)** (2016) 781-790.

[10] Tapsell, L.C., et al., Foods, nutrients, and dietary patterns: Interconnections and implications for dietary guidelines, *Advances in Nutrition: An International Review Journal* **7(3)** (2016), 445-454.

[11] Leech, R.M., et al., Understanding meal patterns: definitions, methodology and impact on nutrient intake and diet quality, *Nutrition Research Reviews*, **28(01)** (2015), 1-21.

[12] Agrawal, R. and R. Srikant. Fast algorithms for mining association rules, *in Proc. 20th int. conf. very large data bases*, VLDB. 1994.

[13] Burden, S., et al., Identification of food groups for use in a self-administered, computer-assisted diet history interview for use in Australia, *Journal of Food Composition and Analysis* **22**(2) (2009), 130-136.

[14] Woolhead, C., et al., A generic coding approach for the examination of meal patterns, *The American Journal of Clinical Nutrition* **102(2)** (2015), 316-323.

[15] Zhang, S., C. Zhang, and Q. Yang, Data preparation for data mining, *Applied Artificial Intelligence* **17(5-6)** (2003), 375-381.

[16] Hertzog, M.A., Considerations in determining sample size for pilot studies, *Research in Nursing & Health* **31(2)** (2008), 180-191.

[17] Tapsell, L.C., et al., Foods, nutrients or whole diets: effects of targeting fish and LCn3PUFA consumption in a 12mo weight loss trial, *BMC Public Health* **13(1)** (2013), 1-11.

[18] Tapsell, L.C., et al., Weight loss effects from vegetable intake: a 12-month randomised controlled trial, *European Journal of Clinical Nutrition* **68(7)** (2014), 778-785.

[19] Tapsell, L.C., et al., Interdisciplinary lifestyle intervention for weight management in a community population (HealthTrack study): Study design and baseline sample characteristics, *Contemporary Clinical Trials* **45, Part B** (2015), 394-403.

[20] Food Standard Australia New Zealand. AUSNUT2011–13 food measures database file. Available from:http://www.foodstandards.gov.au/science/monitoringnutrients/ausnut/classificationofsupps/Pages/default.aspx. Accessed 29/02/16.

[21] Food Standard Australia New Zealand. AUSNUT2011–13 classification of foods and dietary supplements. Available from: http://www.foodstandards.gov.au/science/monitoringnutrients/ausnut/classificationofsupps/Pages /default.aspx. Accessed 08/07/16

[22] Centers for Disease Control and Prevention, Combination Code in Task 2: Key Concepts about the United States Department of Agriculture Food Coding Scheme. Available from: http://www.cdc.gov/nchs/tutorials/Dietary/SurveyOrientation/ResourceDietaryAnalysis/Info2.htm. Accessed 30 May 2016.